EUROPEAN JOURNAL OF PARAPSYCHOLOGY

VOLUME 7

NUMBER 1

NOVEMBER 1987

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.

EUR.J.PARAPSYCHOL., V.7, 1987

A TEST ON POSSIBLE IMPLICATIONS OF THE OT'S FOR GANZFELD RESEARCH

D.J. Bierman
University of Amsterdam

Anecdotal evidence suggests that the mentation of subjects produced during the Ganzfeld might not only correspond to the target picture, but often contains elements of other pictures from the target set (Sondow, 1987). In one published Ganzfeld experiment weak post-hoc evidence was found suggestive for this phenomenon (Palmer at al, 1977). The fact that elements of the total target set are thought to be present in the mentation might be understood in terms of the Observational Theories (OT's). According to these theories GESP trials consist of two distinct processes: a) The partly random processes in the brain which result in the mentation of the subject and b) the observation of (the meaning) the result, i.e. the feedback of the target picture to the subject. The random brain processes are supposed to be biased contingent on the future observation.

What exactly constitutes an observation is still a question open to research (Weiner and Bierman, 1982). One could argue that the judging procedure where the target set is presented to the subject is a

---

Note: This paper fulfils the publication policy of this journal

partial observation. The space of possible targets which, as far as the subject is concerned, is unlimited until the judging procedure starts is suddenly reduced to only 4 possibilities. If this interpretation is correct it appears that it is consistent with the OT's to suppose a bias of the mentation report contingent upon the target set as a whole. When feedback of the real target occurs a second possibility of 'retroactive' biasing of the brain states during mentation arises, this time contingent on the target picture itself.

The present experiment has been set up to test empirically the hitherto mostly anecdotal evidence of psi on the target set rather than on the target picture. The above sketched 2-step model of GESP is also the background for the exploration of the time lapse between the end of the Ganzfeld stimulation and the start of the judging/feedback procedure. The underlying idea is that, although it may be argued that the state of consciousness produced by Ganzfeld is optimal with respect to the random brain processes it is not clear at all what state of consciousness might be optimal for the feedback part of the psi-process. Only one experiment has explicitly explored this question (v.d.Sijde et al, 1982) by direct manipulation of the state of consciousness right before feedback. Since in that experiment no evidence for psi was found, no conclusion could be drawn with regard to this manipulation. In the present experiment the state of consciousness is not manipulated but it is assumed that after finishing the Ganzfeld the subject recovers in continuous fashion to a normal waking state. Thus the larger the time lapse before feedback, the more normal the state would be on the moment of feedback.

It has been found (Schmeidler, 1982) that self-claimed psychics differ hardly from comparable control groups except for one trait, the time contraction trait. This trait appears to relate to dissociative states. After the Ganzfeld-stimulation but before feedback, subjects were given the time reproduction test in which they had to reproduce a time lapse of 10 seconds which was indicated by the experimenter knocking twice on a table. In our test the subjects were asked to do this 4 times in a row. If they have a perfect reproduction ability this would have resulted in a time of 40 seconds. Subjects who scored below the median were labeled as time contractors.

EXPERIMENTAL SET-UP

Experimenters

Six students of the regular Parapsychology course at the University of Amsterdam performed the role of experimenters in this experiment. To raise their interest and motivation they were asked to formulate a hypothesis of their own. A differential effect between trials with and trials without a 'sender' was predicted by these students. It should be stressed that this prediction was not made (nor supported) by the present author who supervised the experiment. There was no formal hypothesis concerning this effect included in the list of hypotheses submitted to the editor of the EJP (see Appendix) before the experiment and thus no analysis pertaining to this condition will be reported here.

Subjects

As subject volunteered 16 freshmen of the department of Psychology. The subjects were generally unknown to the experimenters. Ad-hoc combinations of 2 experimenters out of 6 were made for each subject in order to reduce the workload of the experimenters. Each experimenter therefore participated in only about 6 sessions. There were no more than 2 sessions per day.

Procedure

The procedure did not differ considerably from the generally accepted Ganzfeld-type experiments. Most recommendations given by Hyman and Honorton (Hyman and Honorton, 1986) were taken into account. Thus a two-experimenters protocol was adhered to and duplicate target sets were used for judging. Apart from these measures it was secured that all target sets in the experiment had never been observed by the experimenters nor by the current author. They were used only once. Randomization was done by the current author using a thoroughly tested RNG-II (Bierman, 1987) connected to a MacIntosh computer. Before each session the second experimenter selected the duplicate target set

according to the random set number and left this set in the subjects room. The subject-bound (first) experimenter received the subject and went straight to the subjects room and did not leave this room before the subject had indicated the final choice. The second experimenter, who went in the meanwhile into the 'senders' room, would select an envelope containing the target picture according to the random number produced by the RNG. This experimenter-agent would only open the envelope, at the moment that Ganzfeld-stimulation in the other room started, if a second random number indicated the trial was a Telepathy-condition trial. Otherwise the target picture was left on the table in its own envelope. The experimenter-agent who was the only one who could be aware of the target (in case of telepathy trials) stayed in the 'senders' room until subject's judging was over.

Judging was done by laying the duplicate pictures on a table in the order in which they were in the envelope (this order was unrelated to the target number since the sets had been prepared elsewhere). After the indication of the subject's choice on the scoring sheet the second experimenter would bring in the actual target picture.

Multiple analyses were prevented by submission of the planned analyses to the EJP before the experiment started (note 1).

In contrast to general accepted methodology no ping-pong balls were used but instead a large opale plexiglas hemisphere with a diameter of 50 cm. was placed over the subject's face and illuminated with red light. This method for Ganzfeld stimulation has been used with great success in physiological experiments (Lethonen and Lethinen, 1972).

Judging procedure by independent judges

Judges were volunteers generally picked up from the lab-canteen. The two sets (target and control) and the relevant subject protocol were laid down in the judging room before the judge entered the room. First the judge had to choose a set considering the set as a whole. Then the control set was removed and it was explained that the judging concerned a Ganzfeld experiment with a single target picture. Then they had to indicate the target picture from the target set. In order to keep the procedure not too complex the experimenter which handled the judge was not blind vis a vis the target picture number. Thus non-verbal sensory leakage was not completely excluded. However since

the current author was, unknowingly to this experimenter, interested mostly in the set-judging in the cases that the target was missed this was not considered to be a major problem.

RESULTS

Primary hypothesis

In table 1 the over-all results are presented in a condensed form.

TABLE 1
Subject's and judge's over-all direct hits

	Number trials	Direct hits	MCE	Rate	p-value(o.t)
Subjects	16	6	4	37.5%	0.19
Judges	16	8	4	50.0%	0.03

It can be concluded that there are marginal indications that psi was functioning in the present experiment. The subjects scoring rate is very close to the estimated true effect size of 33% (Rosenthal, 1986) and thus forms a replication of previous Ganzfeld experiments.

However our main hypothesis does not concern the evidence for psi as indicated by anomalous information transfer concerning the target picture but the amount of information transfer from the target set as a whole. In table 2 the scoring of the judges is given when they had to choose between the target set and a control set on the basis of the protocol of the subject. The results are split for cases where the judge ranked the target first or second (this is generally called a binary hit) or did rank the target third or fourth.

As can be seen from table 2 judges did correctly identify the target

TABLE 2
Setscoring

	Number trials	number hits	MCE	Rate	p-value(o.t)
High ranks (binary miss)	6	5	3	83%	0.11
Low ranks (binary hit)	10	5	5	50%	0.50
Total	16	10	8	62.5%	0.23

set in 62.5% (n.s.) of the cases. However the correct set also contains information which is in the target picture. Thus it can not be decided whether this slight above chance selection is due to the whole set or to the target picture. We could have removed the target picture in order to decide so. Instead we decided to analyse separately those sets where that target picture was not ranked as a binary hit and thus supposedly could not contribute the majority of information to the set. It can be seen that for these cases the judges identify 83% of the sets correctly. This result, a larger effect size for correct set-identification, is consistent with our hypothesis that the protocol of the subjects contains information pertaining to the target set as a whole especially in the case of a binary miss. However it should be stressed that the result is statistically non-significant.

Secondary planned analyses

Table 3 gives the judge data split for subjects who showed time contraction in the time-reproduction task and subjects who did not. Subjects who scored below the median were considered to be time contractors.

TABLE 3
Direct scoring as a function of the time perception of the subjects

	Number subjects	number hits	MCE	Rate	p-value
Contracters	10	7	2.5	77%	0.004
Others	6	1	1.5	17%	0.82

Fisher's exact p, corresponding to this difference is 0.059 so, although contracters score significantly as a group it can not be concluded that those subjects who tended to contract the time scored significantly better than those who did not.

The second planned analysis is given in table 4.

TABLE 4
Direct scoring as a function of time between end of Ganzfeld and feedback

	Number subjects	number hits	MCE	Rate	p-value
Short timespan	11	5	2.25	45%	0.11
Large timespan	5	3	1.25	60%	0.10

The corresponding Chi-square value is non-significant.

CONCLUSIONS

It appears that in the present study psi occurred marginally. The hypothesized information transfer from the set as a whole could not be demonstrated to a statistically satisfactory level. However the analysis pertaining to this effect was based upon only 6 trials. Thus the power was really low. The analysis can be applied to any properly kept old Ganzfeld-database and since the result might be interpreted in the framework of the OT's, maybe suggesting the relevance of concepts like 'partial observation', such analyses seem to be worthwhile.

The increase in scoring rate for subjects that got their feedback after a longer time delay (from the end of the Ganzfeld stimulation) might be a real effect but it is not necessarily related to the state of consciousness on the time of feedback. The timespan between end of Ganzfeld and feedback is strongly related with the amount of mentation and the effort that the experimenter spends to elaborate on this information. Thus this effect, if real, might also be correlated to amount of imagery.

Finally the scoring rate of 77% (MCE=25%) for the subjects who showed time contraction is larger than any known scoring rate of a group selected on the basis of a single measure. It suggests strongly that this measure is a relevant personality trait. It should be remarked that this predicted finding nicely agrees with earlier studies (Stanford et al, 1974; Palmer et al, 1977) which used a different measure and also found a positive relation between time contraction and scoring rate.

NOTES

1. It should be remarked here that on the original submission the indicated dependent measure was Sum of Ranks. However when visiting with Honorton, he strongly advised to use the Direct hit count instead. And it was decided to do so before the experiment was over and any result was known to this author.

ABSTRACT

In a 16-trial Ganzfeld experiment the hypothesis was tested that subjects not only mention elements of the target picture but also of other pictures of the target set as can be predicted on the basis of the Observational Theories. There were 8 trials with and 8 trials without a 'sender'. The direct hit rate of the subjects was 37.5% (N=16, MCE=25%, $p<0.19$) while the independent judges matched 50% of the protocols correctly with the target picture (N=16, MCE=25%, $p<0.03$). The judges identified in 62.5% the correct target set if they had to choose between the actual target set and a randomly selected control set (N=16, MCE=50%, $p<0.23$). The correct identification rate of the target set, when the target picture was judged to be a binary miss was 83% (N=6, MCE=50%, $p<0.11$). This result is only suggestive for the main hypothesis of psi on the whole set. Two planned explorations were done. Subjects that showed time contraction in the time reproduction test tend to score better (77%, MCE=25%, $p<0.004$) than those subjects who showed no time contraction (77% versus 17%, Fisher's exact $p<0.059$). Secondly the time was measured between the end of the Ganzfeld stimulation and the moment of feedback of the actual target. No relation was found between this time lapse and psi performance.

REFERENCES

Bierman, D.J. 'Explorations of some theoretical frameworks using a new PK-test environment', Proceedings of the 30th Annual PA Convention, 1987, 31-51.

Hyman, R. and Honorton, C. 'A joint communique: The psi Ganzfeld controversy', J.o.P., 1986, 50, 351-364.

Lethonen, J.B. and Lethinen, I. 'Alpha rythm and uniformal visual field in man', Electroencephalography and Clinical Neurophysiology, 1972, 32, 139-147.

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

Rosenthal, R. ´Meta-analytic procedures and the nature of replication: The Ganzfeld debate´, J.o.P., 1986, 50, 315-336.

Schmeidler, G. ´A possible commonality among gifted subjects´, J. A.S.P.R., 1982, 76, 53-59.

Sondow, N. Private communication, 1987.

Sijde, P.C.v.d. and Wesseling, P. ´State of consciousness during feedback: Exploring the two step model of psi´, E.J.P., 1983, 5, 5-18.

Weiner, D.H. and Bierman, D.J. ´Toward a definition of observation: A test of the effects of information specificity and meaningfulness on PK´. In W.G. Roll, R.L. Morris and R.A. White (eds.) Research in parapsychology 1981, Scarecrow Press, Metuchen, 1982, 134-135.

Dr. D.J. Bierman
Psychology Laboratory
University of Amsterdam
Weesperplein 8
1018 XA Amsterdam
The Netherlands

APPENDIX

Literal transcription of the description of the experiment submitted to the editors of the E.J.P. in May 1987.

A test on possible implications of the OT's for Ganzfeld research

D.J. Bierman

NB. All recommendations by Hyman and Honorton (J.o.P., 50, 1987, in press) are taken into account.

Procedure:
- standard GF (using a large field)
- 15' relaxation before GF
- double-blind

Material
- 32 sets of 4 pictures with duplicate pool for judging (sets are prepared elsewhere and unknown to E's)

Subjects
- 16 volunteers, freshmen psychology

Randomization
- List produced by thoroughly tested hardware RNG
- Once used target set is never used again

Independent variables
- TC: time contraction trait of subject
- MFB: moment of feedback, minutes after end of 'noise'
- TEL/CLV: sender condition

Dependent variables
- SRss: Sum of ranks of targets (subject scores)
- BHjd: Binary hits of judges comparing sets!
- SRjd: Sum of ranks of targets (judges scores)

Hypotheses
- SRss > 16 * 2.5 to be tested in standard SoR way
- BHjd > 8 t-test
- SRjd > 16 * 2.5
- Scores are dependent on MFB

EUR.J.PARAPSYCHOL., V.7, 1987

TIME REVERSED INTERFERENCE: A REPLICATION STUDY

Gert Camfferman
University of Utrecht

In perception research interference phenomena are extensively studied. Interference implies that the perception of a stimulus affects the perception of another stimulus when the two stimuli are presented either at the same time or within a short time interval. The interference becomes apparent for example from differences in the time to react to a stimulus contingent on the presence or absence of the other stimulus. The effects become in general stronger when the time interval between the presentation of the two stimuli is shorter and disappears when the interstimulus time is greater than the time needed to respond to the first stimulus. In cognition research, interference is observed as a function of the interstimulus time, as well as of the meaning of the stimuli with respect to each other, again expressing itself in changes in the time to react to the second stimulus.

Klintman published a study (1983, study 1) in which he measured the reaction times of a subject to two successive stimuli: a COLOR and a colorNAME. He presented these stimuli on a screen with an interstimulus time of .85 sec. The subject was instructed to call out the name of the COLOR or the colorNAME as quickly as possible after perceiving it and the reaction time was registered. Due to cognitive

interference of the first stimulus (S1, the COLOR) with the second stimulus (S2, the NAME), the reaction time to S2 (RT2) changed with the so called ´Congruence´ of the stimuli. The stimuli are called Congruent when the colorNAME is the name of the COLOR (e.g. S1=RED, S2=´RED´ - quotationmarks indicate a written NAME), otherwise they are called Incongruent (e.g. S1=BLUE, S2=´YELLOW´). According to Klintman, for some subjects RT2 is found to be shorter for Congruent stimuli than for Incongruent ones. For other subjects the reverse is found.

In addition to this interference effect Klintman noticed a tendency in the subject´s reaction times to S1 (RT1) to show similar changes as RT2 as a function of the Congruence or Incongruence of the stimuli: If RT2 tended to get shorter under Congruent conditions than RT1 tended to get shorter as well. This is quite remarkable, since the response to S1 was already completed before the onset of S2, and the subject could not know whether the second stimulus would be Congruent or Incongruent. For this reason and since no other explanation could be found for this observation Klintman considered a paranormal explanation which he called Time Reversed Interference (TRI). TRI then might operate as follows:
By precognition the meaning of S2 is known by the subject before the onset of S1. This has the same effect as if S2 is presented before S1 and gives rise to cognitive interference of S2 with S1, resulting in changes in RT1 similar to the changes in RT2 (i.e. when RT2 is shorter in Congruent than in Incongruent situations so will RT1 and vice versa).

This experiment has some interesting features:
- The occurrence of TRI is totally unintentional: The subject is blind as to the paranormal aspect of the study and no intentional choice has to be made by him.
- The experiment consists of a simple and short test, and hence lends itself for applications in other ESP studies.
- According to Klintman´s hypothesis TRI is related to the ´normal´ interference, so the occurrence of an established psychological effect may be used to predict the direction in the reaction time changes of the paranormal effect.
- The test can be completely automated by the use of a computer and thus might offer a standardised test for research in parapsychology.

The primary aim of the study presented here was to replicate Klintman´s Study I.

Since in TRI a supposed reversed interference plays a role (one in which in the first presentation the precognitively observed NAME of the second presentation interferes with the COLOR) in the present study an additional NAME-COLOR test was introduced to study the normal interference of the NAME with the COLOR stimulus.

METHOD

Each subject participated in a session which consisted of two runs of 35 trials each. Each trial, which was preceded by a visual warning signal of 2 seconds duration, consisted of the presentation of two stimuli, presented sequentially in time with an interstimulus time of .8 sec. (onset stimulus 1 till onset stimulus 2). The visual warning signal was made up of five white stars on a dark background, centered near the bottom of the monitor. The presentation of each stimulus was terminated at the moment the subject responded. Stimulus 1 (S1) consisted of a COLOR and stimulus 2 (S2) consisted of a colorNAME, presented on a monitor in front of the subject. They were taken from the following sets:

COLOR (RED , BLUE , GREEN , YELLOW , WHITE)
NAME ('RED','BLUE','GREEN','YELLOW','WHITE')

The subject had to react to each stimulus by calling the name of the COLOR or the colorNAME as quickly as possible. The computer measured the reaction times RT1 and RT2 i.e.:
-RT1, the time elapsed since the onset of S1 till the onset of the verbal reaction of the subject.
-RT2, the time elapsed since the onset of S2 till the onset of the verbal reaction of the subject.
A trial was called Congruent (C) if S1 and S2 had the same meaning (e.g. S1:GREEN and S2:'GREEN'), otherwise it was called Incongruent (I). These two values (C and I) define a (binary) variable called the 'Congruence' of a trial, which served as the independent variable. RT1 and RT2 are considered to be the dependent variables.

The two runs in which each subject participated differed in one respect: In run 1 the stimuli were presented in the order COLOR-NAME. In run 2 the stimuli were presented in the order NAME-COLOR. Half of the subjects started with run 1 first and the other half with run 2 first.

Because we had planned the experiment to be carried out by four experimenters each running ten subjects, ten sequences of 35 COLORs serving as S1's for the COLOR-NAME run were generated beforehand with the aid of an RNG (Random Number Generator, Bierman type), with the restrictions that each COLOR for run 1 (or colorNAME for run 2) appeared 7 times in 35 trials, and that each COLOR for run 1 (or colorNAME for run 2) appeared at the most two times in succession. These sequences were stored on a diskette.

With the same RNG the Congruence of each trial was determined just before each trial, in such a way that each trial had a p=.5 to be Congruent. In the case the trial was Incongruent, four possibilities remain to serve as S2. In this case the value of S2 was determined by means of the RNG, with equal probabilities for each possibility.

APPARATUS AND ENVIRONMENT

The equipment consisted of a voice key and a color monitor (30x40 cm) connected to an APPLE II+ computer. The stimuli, COLOR and colorNAME, were presented on the monitor. The colorNAME consisted of big white characters on a dark background. The equipment measured the reaction time as follows: The voice key registered the verbal reaction of the subject, while the computer counted the number of clockticks from the onset of the stimulus till the moment of verbal reaction. One clocktick equaled 1/256 sec., hence the resolution of the reaction time measurement is about 4 milliseconds. The experiment was carried out in two adjacent rooms: an experimental room and a computer room. During the experiment the experimenter stayed in the computer room. The experimental room was dimly lighted and located in a quiet environment, to avoid disturbing noises as much as possible. The subjects were seated on a normal chair, which provided an alert sitting posture. The chair was placed at a distance of 1.5 meters from the monitor. An intercom was present for registration of irregular sounds (e.g. coughing or other disturbing noises, which might cause artifacts in the voice key registration). Trials in which artifacts of this kind occurred, were marked. These trials were not used for analysis.

SUBJECTS AND EXPERIMENTERS

The forty subjects, participating in this study, were psychology students or acquaintances of the experimenters, between 18 and 40 years old. The experimenters were four psychology students, participating in a course on parapsychology. For this reason they could not be kept completely blind to the fact that the purpose of the experiment was to study a possible paranormal effect. However, they did not know in which data to expect a paranormal effect in the experiment.

EXPERIMENTAL PROCEDURE

In each session first the subject´s (S) name and the session number were fed into the computer. S´s with odd session numbers started with the COLOR-NAME run, subjects with even session numbers started with the NAME-COLOR run. It was explained to the subjects that they participated in a ´reaction time´ experiment, with the aim to study interference phenomena, and that the experiment consisted of two runs, a COLOR-NAME run and a NAME-COLOR run. Further explanation depended on the first run planned for the subject. E.g. for the COLOR-NAME run it was explained that the monitor would show a COLOR, to which S had to react as quickly as possible by naming it. They were told that the microphone would pick up the response and that the response time would be automatically registered. Then a colorNAME would appear on the screen, to which they had to react again as quickly as possible by reading it out loud. Five stars near the bottom of the screen would indicate that a trial would start in a few seconds. The S´s were informed that 35 trials were to be completed and that after completion a short break would follow, in order to prepare the equipment for the next run. The experimenter then showed the five COLORs and NAMEs one by one. The subject had to name them all to see whether he could perceive the COLORs correctly. After answering questions the experimenter showed S a button with which S could start the run. He was instructed to push this button only after the experimenter had left the experimental room. In the computer room the experimenter could watch the trials appearing on a black and white monitor, and could hear the responses of the subject through the intercom. When obvious artifacts occurred like coughing, that might trigger the voice key erroneously, a notation was made.

As soon as the run was completed the experimenter went to the experimental room and told the S that he was going to prepare the equipment for the next run. The same procedure was then followed as described above, except that the explanation was kept shorter and the roles of COLOR and NAME were reversed.

Each of the experimenters would run 10 subjects, in ten sessions numbered 1 to 10. Dependent on the session number the computer loaded a COLOR sequence from the ten sequences stored before on the experimenter's diskette.

ANALYSES AND HYPOTHESIS

Because the study was intended as a replication study, the analysis was the same as the analysis Klintman carried out. For this analysis the following variables were used:

- RT1, reaction time to the first stimulus in milliseconds
- RT2, reaction time to the second stimulus in milliseconds
- Congruence, C (Congruent) or I (Incongruent), see METHOD for the definition).

Each trial yielded a value for RT1, RT2 and Congruence. Hence for each run in total 35 values for RT1, RT2 and Congruence, were obtained.

Klintman operationalized the TRI in the following hypothesis:

> "If in a given subject there is a tendency to identify the second stimulus (S2) faster under Congruent than under Incongruent 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."

In order to evaluate this hypothesis Klintman defined the following differences:

D1=RT1(C)-RT1(I) (always subtract the Incongruent
D2=RT2(C)-RT2(I) RT value from the Congruent one)

RT1(C) is the reaction time to stimulus 1 for a Congruent trial.
RT1(I) idem for an Incongruent trial.
RT2(C) is the reaction time to stimulus 2 for a Congruent trial.
RT2(I) idem for an Incongruent trial.

Having defined D1 and D2, we are able to formulate the TRI hypothesis in terms of these differences:

If D2>0 then also D1>0 and if D2<0 then also D1<0

Klintman observed the coherence between D1 and D2 when the evaluation of D1 and D2 was based on four RT values only. These four values RT1(C), RT1(I), RT2(C), and RT2(I) were selected by Klintman from the 35 trials of the run according to the following procedure:

1- Take the very first trial, and take the values of RT1 and Congruence of this trial.
2- Find among the remaining trials the first one with the same COLOR (read NAME for the NAME-COLOR run) but with a different value for Congruence, and take the value of RT1 from this trial.

To obtain the difference D1 subtract the RT1 value of the Incongruent trial from the RT1 value of the Congruent trial.

3- Take the second trial (provided this trial was not used in step 2, otherwise take trial 3) and take the values of RT2 and Congruence from this trial.
4- Find among the remaining trials the first one, not used in one of the previous steps with the same COLOR (read NAME for the NAME-COLOR run) but with a value for Congruence differing from that found in 3.

To obtain the difference D2 subtract the RT2 value of the Incongruent trial from the RT2 value of the Congruent one.

The results are scored in a two by two matrix, as in table 1, and Fisher's exact probability test will be used to test the hypothesis. The hypothesis implies that the results will not be randomly distributed over the four cells of the matrix but that higher frequencies will be observed in the diagonal.

NOTES: - RT1's exceeding the interstimulus time of 800 milliseconds are eliminated: In this case subjects were either very slow, or they called the names so softly that the voice key was not triggered.
- RT2's shorter than 200 milliseconds are left out as well: They can be considered as artifacts, arising for instance in the case in which the end of the response to S1 overlaps the onset of S2, thereby improperly triggering the voice key which consequently yields very short RT2 values.
- RT2's longer than 1200 milliseconds are omitted, because they are also obviously artifacts, again mostly due to the fact that the subject called the NAME of the color too softly.

RESULTS

The results for the COLOR-NAME run are shown in table 1, for the NAME-COLOR run in table 2. In these tables each cell presents the number of subjects for which the indicated D1,D2 combination applies. The total number of subjects tested was 40. Due to a mistake of one of the experimenters the data of two subjects for the COLOR-NAME run and of 4 subjects for the NAME-COLOR run were lost. Applying the rules to construct D1 and D2 for each subject and applying the restrictions given in the notes above, 26 combinations of D1 and D2 were obtained which could be entered in table 1 and 28 combinations were obtained which could be entered in table 2. It appears from the Fisher test that for the two runs Klintman's hypothesis could not be confirmed.

FURTHER ANALYSES

I. USING MEAN VALUES OF RT'S

In the analysis above, following Klintman's method to determine D1 and D2, only four out of the 35 trials for each subject were used to obtain these values. In order to study if the rest of the data might yield evidence of a TRI effect, an analysis was carried out in which for each subject D1 was calculated from the difference of the mean RT1(C) and the mean RT1(I):

TABLE 1
Test of the relationship between D1 and D2 for the COLOR-NAME run. Analysis according to Klintman.

	D1>0	D1<0
D2>0	3	4
D2<0	7	12

Fisher's exact: n.s. (one-tailed test)

TABLE 2
Test of the relationship between D1 and D2 for the NAME-COLOR run. Analysis according to Klintman

	D1>0	D1<0
D2>0	2	2
D2<0	11	13

Fisher's exact: n.s. (one-tailed test)

D1=meanRT1(C)-meanRT1(I) and likewise
D2=meanRT2(C)-meanRT2(I).

The results are presented in table 3 (COLOR-NAME run) and table 4 (NAME-COLOR run). Here a marginally significant result appears but only for the COLOR-NAME run. Thus this analysis supports Klintman's finding of a TRI effect. For the COLOR-NAME run we observe that D2 is significantly more often negative (24 subjects showed D2<0 against 9

with D2>0, p<.01). This can be explained by cognitive interference, which aids most subjects to react faster in reading the colorNAME under Congruent conditions.

II. ANALYSIS OF THE NAME-COLOR RUN

The NAME-COLOR run yields an interesting result. Although in this condition the hypothesis is not confirmed, all 35 subjects in the NAME-COLOR run turn out to have a D2<0: they all name the COLOR faster in Congruent than in Incongruent trials. This means that reading the words (colorNAMES) from the screen very strongly interferes with naming the subsequent COLORS. This outcome is consistent with the Stroop COLOR-WORD test in which subjects name e.g. the COLOR of the WORD ´blue´ written in the color red, significantly slower than the COLOR of the WORD ´blue´ written in blue (Stroop, 1935). Thus ´normal´ interference is much stronger in the NAME-COLOR condition than in COLOR-NAME condition.

III. CONSIDERING THE PRECOGNITION EXPLANATION OF TRI

The strong interference from reading colorNAMES with naming subsequent COLORS, observed above, raises the question whether in the COLOR-NAME run a similar effect is present: If Time Reversed Interference is present in a COLOR-NAME run, it acts as if the NAME is presented (precognitively) before the onset of the COLOR stimulus, introducing a NAME-COLOR sequence, which we expect (on the basis of the above findings) to interfere strongly with the naming of the COLOR and thus to influence the reaction times for naming the COLORs. Hence we expect that this interference will reveal itself in that most subject will show a D1<0.

Looking at table 3 we notice that 51.5% (=17/33) of the D1 values are negative: a result which we expect when no effect is present. The RNG randomly attributed Congruence or Incongruence to a trial. Hence the Congruent trials and Incongruent trials constitute two randomly chosen samples from the 35 RT1 values of one subject. When the Congruence or Incongruence of the trial exerts no influence on RT1, we expect the means of these two samples not to differ i.e. the difference of their means (which is D1) is expected to be 0, and the

distribution of these D1 values will be very near to a normal distribution, since we are dealing with mean values. So we expect about 50% of the D1´s to be negative, and that is what we actually observe here (z=.08, p=n.s.).

Thus we can say that there is no support for a hypothesis of precognitive knowledge of the second stimulus, before the onset of the first stimulus. Secondly in the COLOR-NAME run the Congruence of the trial does not exert any influence on the task of naming of the COLOR stimulus, whereas it does influence the reading of the NAME stimulus. But the observation Klintman made was that (Klintman, 1983, p.20): "In addition to the perception of S2 being affected by 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 characteristic changes in the RT to S1 as a function of the meaning of the subsequent S2".

In the following I present arguments that RT1 is not to be considered as a dependent variable, a ´function´ of the meaning of S2 (i.e. a function of the Congruence of the trial), but that it merely shows a correlation with RT2, and that rather RT2 can be considered a function of the Congruence of the trial and of RT1.

IV. COHERENCE BETWEEN D1 AND D2

When we consider table 3 again, we notice that of the subjects with D1<0, 15 out of 17 subjects have a D2<0. However, from the subjects with D1>0, only 9 out of 16 subjects have a D2<0. There is a significant difference between these proportions (p=.019, one-tailed). The high proportion of subjects with D1<0 and D2<0 of 15/17, supports the TRI hypothesis. However, in the following I would like to present another, more plausible, explanation for this observation. In order to do so we will first consider D1 and D2 in terms of the reaction times themselves.

Since D1=meanRT1(C)-meanRT1(I), D1<0 means: meanRT1(C) < meanRT1(I), But meanRT1(C) and meanRT1(I) are two partial means, so we can derive the total mean from them, which must lie somewhere in between. In other words these partial means must lay on either side of the mean. Hence there are only two possibilities:
1) either meanRT1(C) < meanRT1 < meanRT1(I), which is equivalent to

TABLE 3
Test of the relationship between D1 and D2, where D1 and D2 are the mean differences in RT for Congruent and Incongruent trials. COLOR-NAME run.

	D1>0	D1<0	
D2>0	7	2	9
D2<0	9	15	24
total	16	17	33

Fisher's exact P=.046 (one-tailed test)

TABLE 4
Test of the relationship between D1 and D2, where D1 and D2 are the mean differences in RT for Congruent and Incongruent trials. NAME-COLOR run.

	D1>0	D1<0	
D2>0	0	0	0
D2<0	15	20	35

Fisher's exact: n.s. (one-tailed test)

D1<0. This will be called a 'fast' meanRT1(C) in the following. Or 2) meanRT1(C) > meanRT1 > meanRT1(I), which is equivalent to D1>0. This will be called a 'slow' meanRT1(C) in the following.

TABLE 5
Frequency of slow and fast reaction times RT2
as a function of congruence and RT1 for all trials.
COLOR-NAME run

	Congruent trials			relative frequencies		
	RT1 slow	RT1 fast	tot.	RT1 slow	RT1 fast	tot.
RT2 slow	7	2	9	.44	.12	.27
RT2 fast	9	15	24	.56	.88	.73
tot. RT1	16	17	33	.48	.52	1.00

With this new definition we have: D1<0 is equivalent to a 'fast' meanRT1(C) and D1>0 is equivalent to a 'slow' meanRT1(C). This way we can rewrite table 3 in terms of 'fast' and 'slow' RT's, which results in table 5.

We observe from table 5 that each subject has equal chances to fall into one of two categories: Those who on the moments that they reacted fast to stimulus 1 and had a Congruent trial assigned to them by the RNG (52% of the subjects), and those who reacted slow and had a Congruent trial assigned to them (the other 48%). We observe that overall, by cognitive interference 73% of the subjects was aided to react fast to stimulus 2. For the subjects of the first group mentioned above, however, i.e. those who were already fast in their reaction to S1, it appears that 88% reacts fast on S2. The subjects of the second group, i.e. those who were already slow in their reaction to S1, do not reach this percentage of 88%, but only 56% of these subjects were fast in their reaction to S2. As already mentioned above, there is a significant difference between these two proportions ($z=2.06$, $p=.019$, one-tailed).

In view of these observations it seems plausible to interpret the data as follows: RT1 can be seen as a sort of measure of 'alertness'.

When for a given trial, the subject has a fast RT1 this means he is alert and when the computer presents a Congruent second stimulus in this situation, the subject is in addition aided by the cognitive interference to be fast in his reaction to stimulus 2. Thus if one considers a subject who's RT1 is fast and the trial is a Congruent one, then there is a high probability of finding a fast RT2. To reverse this statement: In Congruent trials you will find that most subjects with fast RT2's have fast RT1's. Considering Congruent trials again, one will also find for subjects with a slow RT2 that this is associated with a slow RT1 (their slow start). Considering Klintman's TRI hypothesis we conclude that indeed it holds that (Klintman, 1983, p.23): "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)". This is the first part of the TRI hypothesis and is in agreement with the data. The second part fits equally well: "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".

For the NAME-COLOR run in table 4 we observe that all D2<0. This is equivalent to the statement that all meanRT2(C) are fast. This shows how strong the effect of cognitive interference is in a NAME-COLOR run. But here again there are not significantly more subjects with D1<0 (meanRT1 fast). Congruence is neatly distributed over all trials and apparently RT1 does not depend on the Congruence of the trial.

To sum up the results so far:
1) The observation above, i.e. many subjects having D2<0, fits very well with the model of cognitive interference. Using the value of reaction time RT1 we can predict the probability of finding a fast or a slow RT2 for the subject.
2) The contingency tables for RT1 and RT2 (or D1 and D2) obtained in this way satisfy the TRI hypothesis as well, but we don't need a paranormal explanation to understand their properties.
3) The results of the analysis using the mean values for D1 and D2 (table 3) shows the same tendency as the analysis according to Klintman (table 1), though the latter must show more variance, since this analysis uses a (small) sample of all trials
4) From the data of the NAME-COLOR run presented in table 4, which yielded only negative values for D2 for all subjects regardless of the value of D1, it was according to the TRI hypothesis plausible if precognition is present to expect the same tendency for D1 in the

COLOR-NAME run. Since we do not observe this, the precognition assumption is not supported by these observations.

ANALYSES BASED ON ALL TRIALS

In the analyses above we considered differences of the means of RT's, and we entered the number of subjects who had certain values for these mean RT's in a contingency table. Since interference and 'alertness' phenomena reveal itself within trials, I wondered whether I could find a direct association between RT1 and RT2, within each trial, for each subject. In order to study this all trials of a run were pooled and for each trial, each RT1 and each RT2 was labeled 'fast' or 'slow' as follows: If RT1 for a run is below the median RT1 for that subject than RT1 is labeled 'fast', otherwise it is called 'slow'. Likewise RT2 is labeled 'fast' or 'slow'. Since we are now considering individual RT's, who's distribution is skewed, the median is used in this analysis, which ensures that exactly 50% of the RT's of each subject will be fast and 50% will be slow. A classification problem arose when it appeared that one or more RT's coincide with the median. In that case the trial was omitted and the median was calculated again. With this classification I counted all trials within the four categories: fast RT1 and fast RT2, fast RT1 and a slow RT2, slow RT1 and fast RT2, and slow RT1 and slow RT2. Table 6 presents the results.

RESULTS

Table 6 presents the contingency tables for the COLOR-NAME run and for the NAME-COLOR run (the figures are relative frequencies). The coherence between RT1 and RT2, due to the effect which was called 'alertness' above, is clearly present in the two runs: Fast RT1's go significantly more often together with fast RT2's (COLOR-NAME run, $Z=2.06$, $p=.02$, NAME-COLOR run, $Z=2.2$, $p=.014$, one-tailed). From comparing the COLOR-NAME and the NAME-COLOR results, it appears that the strength of this effect is independent of the order in which the stimuli are presented, even though the interference is much stronger in the NAME-COLOR than COLOR-NAME condition (see also analysis I). Hence we conclude that the strength of the 'alertness' effect is independent of the strenght of the interference phenomenon.

TABLE 6
Freqency of slow and fast reaction times RT2
as a function of RT1 for all trials pooled.

	COLOR-NAME run		NAME-COLOR run	
	RT2 slow	RT2 fast	RT2 slow	RT2 fast
RT1 slow	.556	.443	.555	.445
RT1 fast	.443	.556	.445	.555
		N=343		N=441

When we differentiate the data in table 6 with respect to Congruence and Incongruence we arrive at table 7. We observe from the totals in this table that RT2 depends on the Congruence, hence that cognitive interference is present: For the COLOR-NAME run 56% of the RT2's are fast in Congruent trials against 44% in Incongruent trials (Zdiff=3.1, p=.001, one-tailed). The interference for the NAME-COLOR run is much stronger (Zdiff=12.6), which is consistent with our findings above.

Secondly we observe that RT1 does not significantly depend on the Congruence of the trial, neither for the COLOR-NAME (.53 for Congruent trials against .47 for Incongruent trials, Zdiff=1.56, p=.06, one-tailed), nor for the NAME-COLOR run (Zdiff=.6, p=.26, one-tailed). Since the Congruence does not influence RT1, there are no "characteristic changes in the RT to S1 as a function of the meaning of the subsequent S2", (Klintman, 1983, p.20)

Thirdly as we can observe in table 7, in Congruent trials of the COLOR-NAME run a fast RT1 is significantly more often accompanied by a fast RT2 (Z=3.9) and for Incongruent trials a slow RT1 is significantly more often accompanied by a slow RT2 (Z=2.5). Here we see clearly that when considering the data on the level of individual subjects that fast RT2,s for Congruent trials will be accompanied more often by fast RT1 values, giving rise to a TRI hypothesis.

TABLE 7
Freqency of slow and fast reaction times RT2
as a function of Congruence and RT1
for all trials pooled.

	COLOR-NAME run				NAME-COLOR run		
	RT2 slow	RT2 fast	tot.		RT2 slow	RT2 fast	tot.
Congruent trials				Congruent trials			
RT1 slow	.52	.48	.53	slow	.38	.62	.49
fast	.34	.66	.47	fast	.21	.79	.51
Tot.	.44	.56			.29	.71	
			N=333				N=434
Incongruent trials				Incongruent trials			
RT1 slow	.60	.40	.47	slow	.74	.26	.51
fast	.53	.47	.53	fast	.69	.31	.49
Tot.	.56	.44			.71	.29	
			N=353				N=441

The conclusion of this analysis is consistent with our earlier conclusions: interference and 'alertness' can explain the behavior of the RT data.

ABSTRACT

The present study is a replication of a study carried out by H.

Klintman (1983). Klintman reported an effect which he observed in the data of a cognitive interference study, and which he called Time Reversed Interference (TRI). The present study consists of two runs of 35 trials, in which each trial consists of two stimuli presented to the subject on a video monitor. In the first run a COLOR is presented and 800 milliseconds later a colorNAME. In the second run the colorNAME is presented first, followed by a COLOR. The subject has to react to the stimuli by naming them as quickly as possible and the reaction times are registered automatically by a computer equipped with a voicekey. The interference and TRI effects are revealed in changes in the reaction times as a function of the so called Congruence of the trial. A trial is called Congruent if the colorNAME is the name of the COLOR, otherwise it is called Incongruent. The reaction time to stimulus 2 (RT2) tends on the average to be shorter for Congruent than for Incongruent trials. TRI is the (hypothesized) tendency of RT1 (the reaction time to stimulus 1) to show similar changes as RT2 for a given subject. RT1 tends to 'follow' the changes in RT2. Klintman interpreted TRI as a possible precognitive phenomenon.

The data of the present study does not yield support for the TRI hypothesis. A further analysis however indicates that there is an association between RT1 and RT2. For Klintman this observed association gave rise to the TRI hypothesis. It is demonstrated that RT1 is an independent variable, and that RT2 is the dependent variable. RT2 depends on the Congruence of the trial and on RT1. An alternative interpretation is presented in which a fast RT1 reflects an alert state of the subject, favoring a fast RT2. It is shown that the strength of this dependency is independent of the strength of the cognitive interference. Interpreting RT1 as a dependent variable gives rise to a TRI hypothesis, which then of course also is satisfied by the data.

A strong interference effect was found for the NAME-COLOR run: All subjects reacted on the average faster in naming the COLOR in Congruent trials than in Incongruent trials. This effect is consistent with the principles on which the Stroop COLOR-WORD test is based.

REFERENCES

Klintman, H. 'Is there a paranormal (precognitive) influence in certain types of perceptual sequences? Part I', E.J.P., 1983, 5, 19-50.

Stroop, J.R. 'Studies of interference in serial verbal reactions', Journal of Experimental Psychology, 1935, XVIII, 6, 649-651.

G. Camfferman
Parapsychology Laboratory
University of Utrecht
Sorbonnelaan 16
3584 CA Utrecht
The Netherlands

EUR.J.PAPAPSYCHOLOGY, V.7, 1987

A STUDY OF PARANORMAL IMPRESSIONS OF PSYCHICS
PART VI. LONG-TERM VARIATION IN THE BEHAVIOUR OF PSYCHICS

Hendrik G. Boerenkamp
University of Utrecht

In a previous paper the design of an experimental study of paranormal impressions of psychics was described (Boerenkamp, 1985a). Psychics are defined as persons who believe themselves able to obtain paranormal impressions at will. Usually, psychics are consulted by clients about problems related to themselves and to persons in their environment. The entire study consists of a predetermined number of series of sessions with a group of psychics and two groups of non-psychics in which the subjects gave their impressions about persons unknown to them.
A series of sessions means that each psychic gave his or her impressions about one target person. The series with the psychics were divided in three standard series, plus two subgroups of five experimental series each. Two of the three standard series were held at the start of the investigation (standard series A and B) and one at the end of the investigation (standard series C). Among other reasons, the standard series A and B were included to study the short-term variation in the behaviour of psychics and to provide a description of the content of sessions when psychics apply their assumed ESP abilities under conditions which resemble as much as possible the daily circumstances of sessions with clients (for the results, see Boerenkamp, 1985b). A further aim of including these standard series A

and B was to compare the content of these sessions with the content of the sessions of each of the experimental series, which were held to study the effect of a number of variables on the content of the sessions, i.e. the statements of the psychics about target persons. The results of the experimental series were discussed in Boerenkamp, 1985c and 1986a. The content of the sessions of standard series A and B was also compared with the content of similar sessions with two groups of non-psychics (see Boerenkamp, 1986b, for the results).

In this paper the results of standard series C are discussed and compared with the results of standard series A and B in order to study possible long-term variation in the behaviour of the psychics. The period between the sessions of standard series A and B on the one hand and standard series C on the other hand was about five years.

This paper is the last one concerning the entire study of paranormal impressions of psychics. In addition to the planned series of sessions one post-hoc series was carried out. The results of this series were not interesting enough to discuss them extensively, but some remarks about the post-hoc series are made at the end of this paper.

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

The sessions are described in terms of the number of statements of the psychics and the number of informative actions (feedback) of the sitters. Also the results of the informational, structural and interactional analyses are presented. For a detailed discussion of these concepts, see Boerenkamp, 1985a.

From the previous study (Boerenkamp, 1985b) it appeared that the behaviour of the psychics in standard series B was consistent with their behaviour in standard series A. Hence, the data of standard

series A and B were combined in order to minimize the role of accidental influences on the content of these standard series and the combined series are further denoted as standard series A+B. Three of the six target persons of series A and B were the target persons in standard series C.

The two standard series A and B were carried out with 12 psychics. Standard series C was conducted with 8 of the 12 psychics. (As discussed in Boerenkamp, 1985a, for different reasons the contribution of of the other psychics was lacking in the series at the end of the investigation.) The comparison between the data of standard series A+B and the data of standard series C is based on the data of the psychics involved in both series. Thus, for the comparison the data of P3, P8, P9 and P12 were excluded from the standard series A+B. (In this Chapter standard series A+B without the contribution of P3, P8, P9 and P12 and standard series C will be referred to as just series A+B and series C).

Number of statements and informative actions

The total number of statements made by the 8 psychics in series A+B was 1561 and in series C 646. The number of informative actions in series A+B was 1004 and in series C 446. The informative actions in series A+B consisted of 408 extended informative actions (41%) and of 596 confined ones, and in series C of 137 extended informative actions (31%) and 309 confined ones.

Thus, it appeared that the psychics made less statements in the series at the end of the investigation (Wilcoxon matched-pairs signed-ranks test: $T=0.00$, $N=7$, $p<.02$; the Wilcoxon matched-pairs signed-ranks test will further in this paper be denoted with WXT; table G of Siegel's Nonparametric Statistics, McGraw-Hill, 1956, is consulted to find the significance level of the observed T-value). However, the behaviour of the sitters probably contributed to this difference. Altough they gave relatively the same number of feedback responses (1004 responses to 1561 statements in series A+B and 446 to 646 statements in series C; WXT: $T=17.00$, $N=8$, n.s.), they gave it relatively more often in the form of extended informative actions in series A+B than in series C (WXT: $T=3.00$, $N=8$, $p<.05$). This implies that the sitters had a tendency to give less extended feedback at the end of the investigation. It appeared from the experimental series

that the number of statements from the psychics is directly related to the amount of feedback and therefore the difference between the number of statements in series A+B and series C probably has to be attributed to the behaviour of the sitters.

It further appeared that psychics were consistent in the number of statements they made in the sessions of series A and series B. The correlation between the length of the sessions of the same psychics in series A and B was positive to a significant degree (Spearman rank correlation coefficient: $r_s=.86$, $t=5.29$, $df=10$, $p<.001$; all correlations applied in this paper are Spearman correlations and are further denoted with r_s; table P of Siegel's Nonparametric Statistics, McGraw-Hill, 1956, is consulted to find the significance level of the observed correlation). The psychics were still consistent in this respect at the end of the investigation ($r_s=.95$, $df=6$, $p<.01$).

As was the case in series A+B in series C the number of informative actions from the sitter depended on the number of statements from the psychics ($r_s=.77$, $df=6$, $p<.05$).

The informational analysis

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

The inter-rater-reliability for two judges in series A+B and in series C was comparable to the inter-rater-reliability obtained in the previous study. In series A+B the judges agreed on 84% of the statements and in series C on 80% of the statements. The distributions of the scores of potential paranormal value, based on the combined scores of two judges rating each statement on both probability of

correspondence and degree of spontaneity on scale ranges from 1 to 4, are presented in table 1.

TABLE 1
Distribution of scores of potential paranormal value

	Categories of potential paranormal value									
	low	(%)	low-med	(%)	medium	(%)	med-high	(%)	high	(%)
series A+B	814	(52)	373	(24)	221	(14)	96	(6)	57	(4)
series C	312	(48)	164	(25)	117	(18)	32	(5)	21	(3)

Applying a cut-off criterion between the medium and medium-high categories, series A+B yielded about 10% and series C about 8% statements with potential paranormal value (ppv). The percentages of statements with potential paranormal value in series A+B and series C are not significantly different (WXT: T=18.00, N=8, n.s.).

Thus, the psychics did not take more or less 'risk' by making more or less spontaneous and specific statements at the end of the investigation than at the start of the investigation. It appeared further that each psychic was rather consistent in the percentage of statements with ppv in series A+B and in series C (r_s=.78, N=8, $p<.05$).

Of the 53 statements with ppv of series C, 29 were made in the first halves and 24 were made in the second halves of the sessions. The difference is not significant. Comparable results were observed in series A+B.

The second step in the informational analysis was to establish how many of the statements with ppv also fulfilled the criterion of 'sufficient degree of correspondence'. The statements with ppv were rated on a four point scale as being true or untrue. Statements which received a score of either 1 or 2 were called statements of negative paranormal value (untrue statements) and statements which received a score of 3 or 4 were called statements of positive paranormal value (true statements).
In series A+B 18 out of the 153 statements with ppv (12%) and in

series C 7 out of the 53 statements with ppv (13%) had positive paranormal value. The difference is not significant.

The structural analysis

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

(1) Topics discussed in the statements

The total number of topics discussed in the 1561 statements of series A+B was 2096 and in the 646 statements of series C 856. The distribution of the topics in the series is presented in table 2. (For some categories the description is abbreviated because of the limited space in the table; for a complete description, see Boerenkamp 1984).

The distributions of the topics discussed in series A+B and in series C are rather similar. The correlation between the frequency of the topics (subcategories 10-45) discussed in series A+B and in series C is positive to a significant degree ($r_s=.78$, $t=4.70$, $df=14$, $p<.001$). Further, in the set of statements with ppv compared to the set of statements without ppv of series A+B fewer descriptions of psychological characteristics and more statements about physical characteristics and specific topics were observed. The same pattern was also observed in series C (WXT: e.g. psychological characteristics: $T=3.00$, $N=8$, $p<.05$).

(2) Person discussed in the topics

Out of the 2096 topics discussed in the statements of series A+B a total of 311 (15%) concerned a person related to the target person. In series C the equivalent number is 118 out of 856 topics (14%). The

TABLE 2
Distribution of topics discussed in the statements

Topics	A+B	C	A+B%	C%
A Physical characteristics	366	184	17%	21%
10 Sex	6	3	0%	0%
11 Age	25	12	1%	1%
12 Appearance	96	45	5%	5%
13 Bodily health	184	90	9%	11%
14 Being alive or dead	55	34	3%	4%
B Psychological characteristics	909	307	43%	36%
21 Personality traits	536	187	26%	22%
22 Psychological circumstances	308	101	15%	12%
23 Religious orientation	65	19	3%	2%
C Relations	353	106	17%	12%
31 Relations with family members	143	56	7%	7%
32 Relations with friends	160	29	8%	3%
33 Relation with sitter	50	21	2%	2%
D Specific topics	468	259	22%	30%
41 Civil status etc	50	33	2%	4%
42 Circumstances in work	157	58	7%	7%
43 Circumstances in living	82	40	4%	5%
44 Leisure activity	96	27	6%	3%
45 Specific name, event	83	101	4%	12%

difference is not significant (WXT: T=12.50, N=7, n.s.).
Further, it appeared in series A+B that, compared to statements without ppv, statements with ppv involved relatively the same proportion of topics concerning a person related to the target person. The same pattern was observed in series C (WXT: T=9.50, N=7, n.s.).

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

For the different series the distributions of statements about past, present and future are presented in table 3.

TABLE 3
Distribution of statements about past, present and future

	past		present		future	
	FH	SH	FH	SH	FH	SH
series A+B	126	151	598	544	57	85
series C	40	61	252	193	31	69
series A+B total	277		1142		142	
series C total	101		445		100	
series A+B %	18%		73%		9%	
series C %	16%		69%		15%	

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

The distributions of statements about past, present and future in series A+B and in series C are significantly different. In the series at the end of the investigation the psychics made more statements about the future of the target persons (WXT: future: $T=3.00$, $N=8$, $p<.05$).

After splitting up the sessions into first and second halves, it appeared in series A+B that there was a predominance of statements about the future in the second halves of the sessions. The same holds for series C (WXT: $T=0.00$, $N=6$, $p<.05$).
In series A+B the set of statements with ppv contained relatively more statements about the past and relatively fewer statements about the present and the future compared to the set of statements without ppv. This difference was also found in series C (WXT: $T=1.50$, $N=7$, $p<.05$). $p<.02$).

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

The distributions of statements about favourable, neutral or

unfavourable states of affairs are presented in table 4.

TABLE 4
Distribution of statements concerning favourable, neutral or unfavourable states of affairs

	favourable		neutral		unfavourable	
	FH	SH	FH	SH	FH	SH
series A+B	128	144	288	303	368	330
series C	39	58	120	114	164	151
series A+B total	272		591		698	
series C total	97		234		315	
series A+B %	17%		38%		45%	
series C %	15%		36%		49%	

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

The distributions in series A+B and series C are not significantly different (WXT: e.g. unfavourable: T=11.00, N=8, n.s.). It appeared that statements about favourable and unfavourable states of affairs were equally distributed over both halves of the session in series A+B. The same holds for series C (WXT: e.g. unfavourable: T=7.50, N=7, n.s.).
The set of statements with ppv in series A+B had the same distribution with respect to this characteristic as the set of statements without ppv. The same holds for series C.

(5) Number of statements in the form of advice

In series A+B 95 statements (6%) involved advice, in series C 65 statements (10%). The difference is not significant (WXT: T=11.00, N=7, n.s.).
It appeared that the set of statements with ppv contained significantly fewer statements involving advice than the set of statements without ppv in series A+B. The same holds for series C (WXT: T=2.00, N=7, $p<.05$).

(6) Number of statements preceded by a silence

In series A+B 267 statements (17%) were preceded by a silence of 3 seconds or more. In series C this number was 87 (13%). The difference is significant (WXT: T=3.00, N=8, $p<.05$). Thus, in the series at the end of the investigation the statements of the psychics were less often preceded by a silence.
Further, it appeared that with respect to this characteristic in series A+B the set of statements with ppv did not significantly differ from the set of statements without ppv. However, in series C the two sets were significantly different (WXT: T=1.00, N=7, $p<.05$). The set of statements with ppv contained relatively more statements preceded by a silence in series C.

(7) Number of positive and rhetorical statements

Rhetorical statements are statements in which the psychic asks for immediate feedback in contrast with positive statements in which the psychic does not ask for immediate feedback. The distributions of positive and rhetorical statements for the different series are presented in table 5.

TABLE 5
Distribution of positive and rhetorical statements

	positive		rhetorical	
	FH	SH	FH	SH
series A+B	522	571	259	209
series C	208	235	115	88
series A+B total	1093		468	
series C total	443		203	
series A+B %	70%		30%	
series C %	69%		31%	

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

The distributions of positive and rhetorical statements in series A+B and series C are not significantly different (WXT: T=10.50, N=8, n.s.).

When first and second halves of the sessions of series A+B were compared, a significant difference appeared between the two halves of the sessions as regards the proportion of rhetorical statements. More rhetorical statements were found in the first halves of the sessions. The predominance of rhetorical statements in the first halves of the sessions was also found in series C (WXT: T=3.50, N=8, $p<.05$). The set of statements with ppv in series A+B contained significantly more rhetorical statements compared to the set of statements without ppv. This was also found in series C (WXT: T=0.00, N=7, $p<.02$).

(8) Number of statements preceding an informative action

Since all sorts of statements might invite feedback it was studied how many statements the psychic made before the sitter reacted with an informative action. If the psychic made a statement and the sitter reacted directly with one or more informative actions, it is indicated by P1S in table 6. If the psychic made two statements before the sitter reacted, it is indicated by P2S, etcetera.

TABLE 6
Distribution of statements preceding an informative action

	statements	P1S%	P1S	P2S	P3S	P4S	P>4S
series A+B	1651	39%	649	210	63	28	31
series C	646	49%	314	70	27	9	9

Series A+B is not significantly different from series C with respect to the proportion of P1S interactions (WXT: T=9.00, N=8, n.s.). In series A+B the set of statements with ppv contained an equal proportion of statements which were part of P1S interactions as the set of statements without ppv. A similar result was found in series C.

The interactional analysis

The objective of this analysis is to study the actions taken by the psychics after receiving a denial to one of their statements. In

standard series A+B it was found that the psychics mainly use four types of responses to a denial:

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

In table 7 the different responses of the psychics for the different series are presented.

TABLE 7
Distribution of types of responses to a denial

	AD	%	NI	%	T>S	%	I=S	%
series A+B	17	14%	73	62%	17	14%	11	9%
series C	12	20%	33	56%	4	7%	10	17%

It appeared that in series A+B as well as in series C the psychics used 76% 'acceptance of failure' (AD and NI) reactions (WXT: T=12.00, N=7, n.s.).

The post-hoc series

In addition to the planned 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 whether telephone sessions would differ from face-to-face sessions.

It appeared that two of the eight psychics approached did not engage themselves in telephone sessions. The potential paranormal value and the positive paranormal value of the statements of the remaining six psychics was equally low as in the face-to-face sessions. Further, it appeared that in the telephone session all psychics asked for as much information as available before they made statements.

DISCUSSION

As regards the length of the sessions, it seems safe to conclude that there is not much long-term variation in the behaviour of the psychics. The difference in length in the series at the start of the investigation and in the series at the end of the investigation is probably mainly caused by a difference in the behaviour of the sitters. They gave less extended feedback in the series at the end of the investigation, resulting in less statements from the psychics. Possibly, the tendency of the sitters to give less extended feedback, reflects their low expectations at the end of the investigation that psychics might be able to give paranormal impressions.

In the standard series run at the start of the investigation it was found that approximately only one out of ten statements could be considered as meeting the criterion of 'being sufficiently spontaneous and specific', i.e. could be assigned potential paranormal value. Also, that approximately only one out of ten of the statements with potential paranormal value could be considered to meet the criterion of 'being sufficiently correct', i.e. could be assigned positive paranormal value.
About the same number of statements with potential paranormal value and the same number of statements with positive paranormal value were found in the standard series at the end of the investigation. Therefore the conclusions from the previous studies, that it is not necessary to assume a specific paranormal ability of psychics and that therefore research with psychics does not offer a more promising research method than the methods currently employed in parapsychology, are confirmed by the data of the last series of the investigation.

It further appeared that there is also hardly long-term variation in the behaviour of the psychics in terms of the characteristics from the structural and interactional analyses.
Topics are discussed in the same proportion, the same proportion of topics concerns persons related to the target person, the same proportion of statements about favourable or neutral or unfavourable states of affairs is made, the same proportion of statements involving advice and the same proportion of rhetorical statements is made. The psychics react in the same way to denials.
Some minor differences in their behaviour were that they made somewhat more statements about the future, and that their statements were somewhat less often preceded by a silence. This might reflect what is

happening if psychics know a sitter for a longer period of time. He or she possibly feels more confident to 'see' the future of the target person.

ABSTRACT

The aim of the analyses presented in this paper was to study the possible long-term variation in the behaviour of psychics. Psychics made statements about persons unknown to them in similar standard conditions at the start and at the end of a long-term 'process-oriented' investigation of paranormal impressions of psychics. The period between the start and the end of the entire study was 5 years. The psychics produced in both conditions about the same proportion of statements with positive paranormal value, i.e. spontaneous and specific statements which apply to the target person. However, only about 1 percent of all statements in both conditions were spontaneous, specific and correct, and that can be expected by the chance-hypothesis. The data of the present study confirm the conclusion from previous studies, i.e. that research with psychics does not offer a more promising research method than the methods employed usually in parapsychology, employing non-psychics as subjects.
The structure of the behaviour of psychics is rather constant over a long period of time. Some minor differences in the structure of their behaviour might reflect that they feel more confident to 'see' the future of the target person if they know the sitter for a longer time. Unhappily, the results of the entire investigation indicate that this feeling is unjustified as the psychics did not show paranormal abilities.

REFERENCES

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

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

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

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

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

Boerenkamp, H.G. 'A study of paranormal impressions of psychics. Part IV. The second group of experimental series', E.J.P., 6, 2, 1986a, 107-128.

Boerenkamp, H.G. 'A study of paranormal impressions of psychics. Part V. The group of control series with non-psychics', E.J.P., 6, 3, 1986b, 259-284.

ACKNOWLEDGMENT

The author gratefully acknowledges his indebtedness to the 120 students who were involved in the investigation in order to itemize and score all statements of the psychics and the non-psychics on potential paranormal value.

EUR.J.PARAPSYCHOL., V.7, 1987

A 30-YEAR 'EXPERIMENT WITH TIME'
EVALUATION OF AN INDIVIDUAL CASE STUDY OF PRECOGNITIVE DREAMS

Friederike Schriever
W.G.F.P.

AN EXPERIMENT IN 'EXPECTANT OBSERVATION'

In 1953, 40-year-old actress Mrs. M. got in touch with Hans Bender's 'Institute for Border Areas of Psychology and Mental Hygiene' (Freiburg, West Germany) (note 1) and reported that some of her dreams apparently anticipated future events. She volunteered as a subject for a quasi-experimental study dealing with precognition in dreams, and she regularly (every fortnight) sent reports on her dreams with her interpretations. Whenever she had the impression that her dream had come true, she wrote down the experienced real event that she assumed to be related to her dream and provided, if possible, witness testimonies. These were also sent to Bender. The reports on the dreams and their probable confirmation were collected in the institute's archives. From time to time, the material was discussed in detail with the subject, which was particularly helpful in clarifying her actual life situation as well as to elucidate her way of handling the dreams.

Mrs. M. conducted this 'expectant observation' for 28 years - up to her death in 1982. Thus, the study covered an important part of her life. A total of 3000 dreams were filed. This long-term study has become known for the very detailed analysis of the 'Case of

Gotenhafen', a series of dreams which is part of this material, carried out by Bender & Mischo (1960/61, 1961/62). That analysis dealt with a series of 12 dreams which refer to specific events in connection with the film 'Night Fell Upon Gotenhafen', in which Mrs. M. played a part. Four months to five years before the actual events, they appear to have been parts of her dreams.

THE PROBLEM

The objective of this study is to give both a quantitative and qualitative survey of the complete material. Therefore, nearly 1700 dreams from the first 11 years and 3 month of regular dream recording (1 January 1954 - 27 March 1965) were analyzed under two different aspects:

(1) Can precognitive elements, as assumed by the subject, be confirmed by objective means, for example by comparing them to the reported real situation ?
(2) Do these presumably precognitive dreams provide typical patterns such as have been reported in the literature ?

As to the first question, it has to be checked whether all possibly precognitive dreams, i.e. those which were confirmed by the subject, meet the requirements for precognitive dreams. In the ideal case of a precognition:

(1) the precognitive dream has to be reported or recorded before its fulfilment to guarantee objectiveness;
(2) in order to avoid coincidental correspondence the dream must include specific details that are identical in the report and its confirmation;
(3) the prediction of the future event must not be inferable from known facts;
(4) the percipient must in no way, not even by suggestion, be able to cause the fulfilment (i.e. self fulfilling prophecy has to be excluded).

Today, that is in most cases 30 years after the dreams, we could only adhere strictly to the date of receipt of the dreams at the institute in order to meet the first requirement. Thus, it was checked whether each dream had reached the institute before its presumed confirmation. Only those dreams were examined that had the date of receipt clearly stated or for which that date could be reconstructed

on the basis of the correspondence. Unfortunately, Mrs. M.´s letters, that had been marked by the date and that had been attached to the submitted dreams, were at some point separated from the dreams. Therefore, the date of receipt of a great part of the material cannot be established anymore. Since the subject had not always sent her dreams in chronological order, a reconstruction based on known dates of receipt of other dreams is not possible either. Although it is likely that the dreams have arrived ´in time´ at the institute, we did not include these dreams in the analysis in order to comply with one of the basic requirements for the documentation of cases of precognition.

DESCRIPTION OF THE MATERIAL

The original material consists of 1691 dreams. Of these, 726 dreams (43%) do not confirm to the required reliability of documentation.

Number of dreams with confirmation:

687 dreams of the total material,
395 dreams of the documented dreams,

of these:

269 dreams are insufficiently documentated (i.e. the dream arrived at the institute only after its presumed confirmation (n=221); a written confirmation (n=12) or the date of the day of the fulfilment (n=36) is missing)

Consequently there are 126 dreams left (115 with known date of receipt; for 11 other dreams the date could be deducted). Thus, 126 dreams (note 2) are available for an examination in terms of the first of the questions mentioned above, i.e. for analyzing

(a) the fulfilment of requirements (2) through (4)
and
(b) the quality of the correspondence between dream and presumed confirmation.

Ad (a):
For the following reasons it is difficult to check these fundamental requirements:
-The subject is no longer alive and available for interviews about her possible previous knowledge of individual events. Thus, we are dependent on the information provided by her comments and by the recorded explorations.
-Since most of the dreams and their confirming events took place 30 years ago, it is almost futile to try to investigate the circumstances of the confirmation, in particular the more important details. It is no longer possible to do interviews with witnesses as they were done by Bender & Mischo (1960/61, 1961/62) in the ´Gotenhafen case´.
From this, it follows that:
-in some cases it is difficult to assess whether Mrs. M. could have had any influence, whether conscious or unconscious, on the dream fulfilment.
Therefore we have every reason to be very critical of the present material.

Ad (b):
The quality of the confirmation provided by the subject has been differentiated into five categories. We distinguish between:
-unconfirmed dreams: no correspondence whatsoever between dream and confirmation;
-insufficiently confirmed dreams: correspondence between dream and confirmation is only found for irrelevant details and is not specific enough; and/or knowing the subject´s situation there are several alternatives for explaining the dream;
-relatively well confirmed dreams: distinct reference points between dream and the related event are identified. A correspondence is given for at least some relevant motives;
-well confirmed dreams: on the whole, the theme of the dream is confirmed, elements of dream and confirmation show evident parallels, correspondence is obvious;
-very well confirmed dreams: compared to the well confirmed dreams, correspondences here are even more precise and outstanding. They seem to be more unrelated to coincidence, i.e. they are marked by even greater specificity.

For the determination of the probable precognitive content of the dreams each dream was subdivided into qualitative, meaningful units, called elements. After that, the confirmations of the single dreams were checked for the same content elements. If this was the case, the

elements of the confirmations and those of the dreams were juxtaposed in a table, so that the number of the confirmed elements of the dreams and their quality could be determined. The quality, however, always depends on the subjective criteria of evidence of the evaluating person, especially since the elements of dreams differ. There are elements that can clearly be recognized (realistically or symbolically) in the later real situation and others that only give hints or allow speculations.

To illustrate this, it may be useful to describe a specific example:

Dream no. 1109 of 20 August, 1961:

"This very long dream certainly is about the dissolution of the ´Young Theatre´. Cases in an enormous disorder, with workers present. Brochures and programmes in many boxes. I leaf through a brochure with pictures of previous seasons. A certain Ilse Langen, which in my dream I mistake for the actress Inge Langen, in Biedermeier style poke bonnet. Yet, she does not at all look like the girl from Munich and I think that she might have had a nose operation. Actually, she then steals a handcart full of boxes. In all this mess, I am watching a rehearsal or performance. Very good and serious, with Vika K. Suddenly it turns into a weird, but skilful comedy and I laugh till tears come. The actor K.M. enters with a ladder, leans it against a rock and steps on one of the lower steps and falls over with the ladder, which falls on him. He is lying there motionless with a face distorted with pain, but very funny. Then he holds awfully funny masks in front of his face while standing on the ladder again - a perfect clown. The ladder fell over because it was standing on the lid of a gully which apparently had been slightly opened."

Comment by the subject:

"Months ago, I heard that the ´Young Theatre´, where I act since the 17th (of August, 1961; F.S.), will move into a new building. At the moment, I have heard nothing about it, but it is possible that they are thinking of moving. At the beginning of the season, we all received brochures with old pictures in them on the occasion of the theatre´s 10th anniversary. Probably, the Ilse Langen of my dream is my colleague Ilse Laux, with whom I play every night. The actor K.M. is not a good clown, to be sure, he is not a good actor at all and I can´t imagine how he got into my mind. Maybe future will show."

Confirmation (by the subject) October, 1961:

"As mentioned in the comment: maybe future will show. Now I am playing Molière for the first time, with both the actors K.M. and Ilse L. of my dream (here, she refers to Ilse Langen; F.S.) in the ´Young Theatre´, and she does not wear a Biedermeier costume, but something similar. K.M. enters and leaves via a ladder - as some of us do as well - (from the lower floor into the stage). K.M. - heaven knows, he is not a comedian - raised a laugh in all of us. I happened to watch the final scene, where all colleagues have funny masks in front of their faces. The elastic of his mask was torn and he had to press the mask against his nose for the whole time and spoke his part through his nose. The atmosphere of departure on the theatre and the packing workers are certainly related to our planned excursion to Wuppertal, which I heard of only now, after the contract for the Christmas fairy tale had been signed, that is end of October."

Second confirmation (by the actor K.M.) of 31 October, 1961:

"I hereby confirm that I had to desperately hold my mask in front of my face in the last scene in one of the last performances of ´The Imaginary Invalid´ because the elastic unhooked and there was insufficient time to repair it. I must have looked very funny because I raised a laugh in my colleagues."

Comparison of dream and confirmation elements:

Dream of August 20, 1961	Confirmation of October, 1961 and of October 31, 1961
Time distance: no more than 10 weeks	
(1) Dissolution of the "Young Theatre"	(1) (Mrs. M. acts in play by Molière in the "Young Theatre")
(2) Cases and disorder	
(3) Brochures and programmes in boxes	
(4) Brochures with pictures of previous seasons	
(5) Colleague Ilse Langen in Biedermeier poke bonnet	(5) Colleague Ilse Langen wearing a similar costume in the play by Molière
(6) Mistaking Ilse Langen for Inge Langen	

(7) Question about nose operation of the colleague	
(8) Colleague steals boxes	
(9) Subject watches performance or rehearsal	(9) Mrs. M. watches final scene of a Molière performance
(10) Serious play with Vika K.	
(11) Sudden turn into a comedy	(11) Unexpected turn into a comedy due to the mishap
(12) Subject is laughing her head off	(12) Colleagues are laughing their head off
(13) Actor K. M.	(13) K. M. plays a part
(14) K. M. with a ladder	(14 & 18) K. M. and others enter the stage via a ladder
(15) Ladder falls over	
(16) K. M.´s face is distorted with pain, but looks funny	(16 & 17) K. M. desperately holds mask on his nose, since the elastic unhooked
(17) K. M. holds funny mask in front of his face	
(18) K. M. on the ladder again	
(19) K. M. - a clown	(19) K.M. looks funny
(20) Ladder fell over, since lid of gully had been opened	

In this dream nearly all elements can be either reduced to ´rests of the day´ or relate to a future theatre play. As mentioned in the comment, dream elements (DEs) 1-3 probably refer as assumed by the subject, to the move of the ´Young Theatre´ to a new house. DE 4 finds its determination in the special brochure for the 10th anniversary of the opening of the same theatre (at the time of the dream). As to DE 5, a twofold reference can be found: On the one hand, around the time of the dream Mrs. M. is on stage together with her colleague Ilse Laux every night (in the dream a shift from Ilse Laux to Ilse Langen); on the other hand, approximately 10 weeks later, Ilse Langen the woman she dreamt of plays a part in Molière´s play ´The Imaginary Invalid´, which 9 other DEs seem to refer to. Actually, there is a relation to one particular performance only. Congruence of the text of the dream and its confirmation is so evident and realistic that the comparison of the elements of dream and confirmation is impressive as such.

Even the aspect of the unexpected turn from seriousness into a comic effect in a particular performance of this play - compare actor K. M.´s confirmation: "(...) that I had to desperately hold my mask in front of my face in the last scene in one of the last performances of ´The Imaginary Invalid´ because the elastic unhooked and there was

insufficient time to repair it." - is expressed in the dream: "(...) a rehearsal or performance (...). Very good and serious, with Vika K. Suddenly it turns into a weird, but skilful comedy (...)". Since it is a production by the ´Young Theatre´, there is a relation to DE 1. However, this relationship is insignificant, since Mrs. M. was engaged at this theatre at the time of the dream. Congruence of the 10 DEs mentioned above, which show relevant details, is so obvious that we count this dream among the very well confirmed ones.

In this way, all 115 dreams were analyzed and classified according to their quality taking into consideration, if possible, dream psychological models and standard psychological alternatives for explanation.

The rating procedure produces the following dream classification:

a) 28 dreams are not confirmed
b) 39 dreams are insufficiently confirmed
c) 34 dreams are relatively well confirmed
d) 12 dreams are well confirmed
e) 2 dreams are very well confirmed

c+d+e: 48 dreams = 5% of the 965 well-documented dreams.
d+e: 14 dreams = 1.45% of the 965 well-documented dreams.

(Taking into consideration the ´well´ and ´very well´ confirmed ´Gotenhafen dreams´, the numbers increase to 2.1% or 5.6% respectively).

Thus, we consider the 48 best-matched dreams as presumably precognitive and the remaining 67 badly-matched ones as presumably non-precognitive. There is no direct evidence to support precognition in the well-matched dreams. In none of the cases we can exclude coincidence, because there is no way of calculating the expected frequency of obtaining the observed number of close matches between dreams and events. The rating conveys the strongly subjective impression that precognitive dreams are characterized by high specificity. A special rating for that purpose has not been carried out, because we have not found an independent variable in the material, pointing to this specificity (in the sense of the ´tracer hypothesis´, put forward by W. von Lucadou 1986).

HYPOTHESES

According to our second guiding question, these 48 presumably precognitive dreams will be examined with regard to particular structures reported in previous work. Typical patterns in precognitive experiences were observed over and over again. Various previous investigations of collections of precognitive experiences have shown, that the subject of these experiences are very often affective events, which are negatively loaded: 61 % according to Stevenson (1970), 60% according to Saltmarsh (1938). These involve particularly ´serious and shocking events´ such as death, accident and illness. The remaining precognitive experiences concern ´trivial incidents´, namely 24% according to Saltmarsh, and positive events, but for only 4%. According to Rush (1986), nearly all spontaneous precognitive experiences refer to personal misfortune or large disasters, rarely to happy events. Most authors assume that precognition of events of emotional significance is more frequent than precognition of events which are emotionally neutral (Dean 1974, Greenhouse 1972, Hastings 1977, Saltmarsh 1938, Stevenson 1970).
Therefore, our first hypothesis is:

H1: Precognitive dreams have specific contents or qualities of contents in common; affective events, which are negatively loaded, and emotional significance predominate in well-matched dreams.

Some studies (Priestley, 1964; Saltmarsh, 1938; Stevenson, 1970) point out the remarkable vividness of precognitive dreams or, at least, their qualitatively felt difference.
Therefore, we hypothesize:

H2: Well-matched dreams stand out for their vividness.

Closely related to this hypothesis is the problem of the recall frequency of precognitive dreams. Following Saltmarsh (1938) and Priestley (1964) we set up the following hypothesis:

H3: Well-matched dreams are recalled more frequently than badly-matched dreams.

If precognitive dreams are distinguished only by the qualities mentioned above, the question arises how the dreamer becomes aware of the significance of the dream. Are dreams being recognized as

precognitive before they are, due to a confirmation, retrospectively defined as ´precognitive´? In spite of Saltmarsh´s conclusion that this is not the case, we put up the following hypothesis:

H4: The percipient has a paranormal awareness of the significance of precognitive dreams (´subjective confidence or conviction´).

In support of our assumption, we may point out that we have had the unusual chance to have Mrs. M. as a subject, who precisely observed and recorded her dreams over a long period of time. Therefore, we are able to profit from her unusually wide experience in dealing with her dreams, so that it seems worthwhile to consider that question again. Bender & Mischo (1960/61, 1961/62) point out a possible relationship between the momentary significance of the dreams at the time of the dream and the later situation, when the dreams seem to become reality. In their investigation of the ´Gotenhafen dreams´ they observed that Mrs. M. had precognitive dreams in situations of fear or stress, i.e. when she was in a special situation.
Thus, we arrive at the assumption:

H5: Well-matched dreams mainly occur in particularly affective, significant life-situations.

The data concerning the time interval between the precognitive dream and its confirmation differ. On the one hand, with increasing time distance between precognition and its related future event, a decline effect has been observed. With two subjects, maximal effects were registered within a few days between dream and event; the number of precognitive effects drops significantly after a period of time longer than 4 days (Tenhaeff, 1976). Dean (1974), Orme (1974) and Stevenson (1970) report similar observations. On the other hand, examples have been described where the precognition occurred years before the corresponding event. Palmer (1978) did not find any time dependence in his experimental precognition studies. These differing results lead to our hypothesis:

H6: There will be no significant difference between the time distribution in well-matched and badly-matched dreams.

The present data will also be checked for Stevenson´s (1970) and Saltmarsh´s (1938) findings that precognitive experiences mostly contain realistic, and hardly ever symbolic, representations.

H7: Well-matched dreams as opposed to badly-matched ones mainly contain realistic (not symbolic) dream images.

However, it should be taken into consideration that symbolic foreseeing might not be recognized as such if the percipient is not aware of the symbolism.

In order to verify these seven hypotheses, a qualitative content analysis was conducted ('structured content analysis' after Mayring 1983). The establishment of categories was done following Hall & Van de Castle (1966), since these authors created a system for classification that attempts to cover frequent psychologically significant contents. In addition to the consideration of the content analysis, all personal comments and all interviews with Mrs. M. have been evaluated (in particular with regard to questions of dream recall frequency and of the subject's life-situation at the time of the dreams, since these points are left aside in the content analysis).

In addition to these predominantly qualitative analyses chi-square tests were carried out, in which the characteristics of the 48, presumably precognitive, well-matched dreams were compared with those of the 67, presumably non-precognitive, badly-matched dreams. Only the (unweighted) frequencies of the characteristics were taken into account. We are aware of a great loss of information by performing this statistical analysis. However, our data are too incomplete to justify any more in-depth analysis.

RESULTS

The qualitative and quantitative analyses yield the following results. Hypothesis H1 according to which affective events, which are negatively loaded, are more frequently foreseen in precognitive dreams, has to be rejected. Mrs. M. only describes mishaps in two of the 48 well-matched dreams (the illness of a daughter and a catastrophe). Most dreams deal with themes such as 'theatre' and 'family'; 52% of these dreams concern her professional life and 60% refer to her family - her relatives, however, often act in connection with theatre, since Mrs. M.'s husband and daughters are actors as well. The dream contents do not contain that many situations that express emotional involvement of the subject - this is the case in only 8% of the well-matched dreams. However, affective events, which

are negatively loaded, are mentioned in 44% of these dreams. Mostly, these concern negative moods, feelings, sorrow and fears, which are in no way related to so-called ´serious and shocking events´. An unexpected number of positive affects, e.g. joy and sense of happiness, - in 25% of the dreams - have been observed. Again, this contradicts our expectations. If we compare these results with those concerning the badly-matched dreams, we find similar percentages: 42% of these dreams refer to the theatre; 49% concern family matters; ´serious and shocking events´ are observed in only 9%; negative affective events in 36%; and positive affective events in 12%. In 7% of the cases the subject expresses emotional involvement. Chi-square tests show that well-matched and badly-matched dreams do not differ in their content or content-quality. Thus we have to reject H1 (see tables 1 to 6).

Both well- and badly-matched dreams contain themes that are related to specific life situations of Mrs. M., to themes, which concern her personally. Looking in particular at those future situations the well-matched dreams are referring to, 10% positively loaded affective events are found and 17% negatively loaded ones (illness, accident, death, unexpected loss of a part). The latter seem to have considerable emotional significance for the subject, but are only reflected in the dreams in two cases. Thus, we have to conclude that only a small part of well-matched dreams refer to an emotionally significant future event, that is negatively loaded. This number is too small to enable us to consider such events characteristic of precognitive dreams.

We cannot take a decision with regard to the hypothesis about particular vividness in precognitive dreams (H2). On the one hand, Mrs. M. mentioned vividness only in 10% of the well-matched cases; on the other hand, in 58% of these cases she sent very lively reports, which could lead to the assumption that these dreams were vividly experienced. There were also 4% non-confirmed dreams reported by her that express this vividness and 55% badly-matched dreams were sent in with lively reports. Chi-square tests compel us to reject this hypothesis, too. Well-matched and badly-matched dreams do not differ in their vividness or livelyness of their dream reports (see tables 7 and 8).

Concerning the recall frequency of precognitive dreams (H3), there

TABLE 1
Dream content in regard to profession
(theatre setting, objects, actions) (note 3)

	Well-matched	Badly-matched	Sum
Yes	25	28	53
No	23	39	62
Sum	48	67	115

Chi-square=1.19 < Chi-square(0.05;1)=3.84, df=1

TABLE 2
Dream content in regard to family

	Well-matched	Badly-matched	Sum
Yes	29	33	62
No	19	34	53
Sum	48	67	115

Chi-square=1.40 < Chi-square(0.05;1)=3.84, df=1

TABLE 3
Dream content in regard to 'serious and shocking events'

	Well-matched	Badly-matched	Sum
Yes	2	6	8
No	46	61	107
Sum	48	67	115

Cells with expected frequency <5: 2 of 4 (50%)
Chi-square=0.99 < Chi-square(0.05;1)=3.84, df=1

Table 4
Dream content in regard to affective events which are positively loaded

	Well-matched	Badly-matched	Sum
Yes	12	8	20
No	36	59	95
Sum	48	67	115

Chi-square=3.32 < Chi-square(0.05;1)=3.84, df=1

TABLE 5
Dream content in regard to affective events which are negatively loaded

	Well-matched	Badly-matched	Sum
Yes	21	24	45
No	27	43	95
Sum	48	67	115

Chi-square=2.00 < Chi-square(0.05;1)=3.84, df=1

TABLE 6
Dream content in regard to emotional significance

	Well-matched	Badly-matched	Sum
Yes	4	5	9
No	44	62	106
Sum	48	67	115

Cells with expected frequency < 5: 1 of 4 (25%)
Chi-square=0.03 < Chi-square(0.05;1)=3.84, df=1

TABLE 7
Vividness

	Well-matched	Badly-matched	Sum
Yes	5	3	8
No	43	64	107
Sum	48	67	115

Cells with expected frequency < 5: 2 of 4 (50%)
Chi-square=1.52 < Chi-square(0.05;1)=3.84, df=1

TABLE 8
Livelyness of dream reports

	Well-matched	Badly-matched	Sum
Yes	28	37	65
No	20	30	50
Sum	48	67	115

Chi-square=0.11 < Chi-square(0.05;1)=3.84, df=1

TABLE 9
Recall frequency

	Well-matched	Badly-matched	Sum
Yes	5	4	9
No	43	63	106
Sum	48	67	115

Cells with expected frequency < 5: 1 of 4 (25%)
Chi-square=0.77 < Chi-square(0.05;1)=3.84, df=1

are contradicting subjective and objective experiences. In personal discussions, Mrs. M. affirmed again and again that she remembered precognitive dreams better than others, which she soon forgot. Looking at her comments and her confirmations of the dreams, she mentioned a good recollection that seemed obvious to her, only in 10% of the well-matched cases and in 6% of the badly-matched ones. Examination by a chi-square test shows that there is no difference between well-matched and badly-matched dreams as far as their recall frequency is concerned (see table 9). In addition, the recall frequency hypothesis is unsupported by the material because we found that she only identified many dreams to be precognitive after having studied them again, i.e. post hoc in relation to later real situations; for this reason, she read her notes again and again. (This applies to the whole ´Gotenhafen series´). Thus, we have to reject this hypothesis.

Paranormal awareness of the dreams´ significance immediately after the dream (H4), i.e. before its probable fulfilment, is not supported by the data. For five months, Mrs. M. classified her dreams every morning, according to quality of rememberance, mood, impression, significance in her actual situation and paranormal awareness of the significance. For 42 dreams she stated her conviction that they were psychic dreams. Out of these 42 dreams, only one is counted among our well-matched dreams (2.4%). Out of the 48 well-matched dreams (omitting the one just mentioned, there are 47 dreams left) she had a paranormal awareness of significance for only 2 other dreams (4.3%). Thus, she was convinced of the paranormality of her dreams in three out of 89 cases (3.4%). That means that the number of well-matched dreams, which are classified by the subject to be precognitive, is less than the number of possibly well-matched dreams out of this sample (48 of 965 = 5%). In addition, she was convinced of a precognitive character in 20 of 42 dreams (48%) which cannot be matched at all, because she did not sent any confirmation for them. Thus, we can assume that she herself had no criterion for distinguishing between precognitive and non-precognitive dreams before she found confirmative events. Thus, we could confirm Saltmarsh´s (1938) findings.

Because today it is hardly possible to reconstruct the situation of the subject at the time of the individual dreams, and because too little information can be drawn from her comments and the interviews, we are unable to reach any conclusion with regard to the hypothesis concerning the affective situation at the time of the dreams (H5). If we look at the later situations that were anticipated in well-matched

dreams, it becomes obvious that 27% affected her deeply, sometimes very deeply (10% pos., 17% neg.). We cannot decide whether these affective events are somehow related to her situation at the time of her dreams. There is no information in the data about her life-situation at the time the badly-matched dreams took place. Therefore we cannot compare well- and badly-matched dreams in this respect.

The time interval hypothesis (H6) can be accepted. The time interval between dream and alleged confirmation does not differ between well-matched and badly-matched dreams. In both cases there is a wide range (from a few days up to 10 years). Possibly, this result is flawed by the above mentioned deficiency in the documentary material. Unfortunately, the dreams that, according to Mrs. M., became reality within a few days, arrived too late at the institute. Thus, we cannot decide about the hypothesis concerning the decline effect for the first-week confirmations. A t-test does not show any difference between the time-distortion of well- and badly-matched dreams (see table 10), when the means of the time intervals between dreams and their related events are compared.
The evaluation of the time hypothesis involves a fundamental problem, because we do not have any data available about the time-range distribution of matches between random dreams and random events.

Hypothesis H7, concerning the quality of the dream images, can be accepted, since the chi-square test shows a significant difference between well-matched and badly-matched dreams for realistic versus nonrealistic dream images (see table 11).

The present total material does contain many symbolic images - Mrs. M. even has developed her own interesting symbolisms - , but these do not belong to the well-matched ones. One reason for this result may be that non-realistic (i.e. symbolic) images convey a lesser degree of correspondence and that realistic images are therefore preferred by the researcher who interprets the data. The problem is that there is no independent criterion to judge about realistic or symbolic dream images.

Now, let us have a look at the results in respect to our questions. When comparing the number of dreams considered as 'confirmed dreams' by, respectively, the subject and ourselves (395 vs. 48), the question about the reasons for this large difference becomes inevitable. Obviously, the main cause is the deficient documentation of the

TABLE 10
Comparison of means of the time-range in well- and badly-matched dreams (in weeks)

Well-matched dreams:	n = 50 (*)
	M = 58.58
	Var = 12221
Badly-matched dreams:	n = 71 (*)
	M = 79.92
	Var = 23210

t=0.845
df=119
t(0.05;119;two-tailed)=1.98
t=0.845 < t(crit.)

*: In a few cases, single events in a dream were matched to different future events at different points of time. Therefore n is larger than the number of well- and badly-matched dreams.

dreams, since 269 of the dreams either arrived at the institute after their confirmation or the confirmations are incomplete, so that we had to exclude them from further analyses. As to the other 126 or 115 dreams, respectively, the reason for the poor correspondence between dream and probable confirmation is the different ratings of the confirmations given by the subject. Mrs. M., in a way, was highly motivated under the pressure to provide possible confirmations of her dreams by getting herself into the role of a test person in an ´experiment in expectant observation´. Often she related dreams with later realistic situations, which cannot easily be reconstructed and accepted without reservation by an outside observer. Moreover, she often considered the relation to the future as the only explanation for her dreams and did not at all take any alternative explanation into account. Finally, today it is difficult to check for possible ´previous knowledge of the related event´ and for the ´possibility to

TABLE 11
Realistic dream images

	Well-matched	Badly-matched	Sum
Yes	37	35	72
No	5	23	28
Sum	42	58	100 *

Chi-square=9.31 < Chi-square(0.01;1)=6.64, df=1

* In a few cases there were both realistic and non-realistic dream images. Those cases were excluded here.

influence the fulfilment'. Of course we have to take into account that the actual dream is only known to Mrs. M. - already reporting represents a distortion - and that we cannot estimate the value that Mrs. M. attributed to an experience in the case of a confirmation. On the whole, we tried to classify the available dreams and their possible confirmations both in a critical and understanding mind.

DISCUSSION

How, then, can it be explained that our results concerning the structure of precognitive dreams differ from those described in previous work on precognitive experiences? First of all, there are fundamental differences in material. The precognitive dreams other investigations refer to are mainly single precognitive dreams taken from a large population, each dream from a different individual, whereas this study is based on a series of dreams of one single subject. Furthermore, the foreseeing of a mishap is, of course, particularly impressive and the person that experienced the precognition will, unless he or she observed the dreams as intensively as the subject of this investigation did, tend to report this distinct event instead of a previously dreamt event of everyday life - if the

latter is noticed at all. Of course, such outstanding experiences are more spectacular and therefore better memorized (see Dunne 1935). So it is not surprising that impressive events (such as mishaps and catastrophes) predominate in collections of spontaneous precognition.

This important difference in the source material seems to be responsible for the different results, in particular those concerning the hypotheses of contents, dream recall frequency and vividness. Here, the underlying dreams - however, the small number should not be overlooked (n=48) - offer a broader range of contents than the individually selected spontaneous reports on precognition, since the former are taken from continuous recordings of dreams. Thus, our observations would suggest that the traditional image of precognitive dreams has to be expanded or revised. At least, the source of the precognitive dreams involved needs to be described more precisely.

Also of psychological significance are the results regarding the contents of precognitive dreams. It seems ´to make sense´ that precognitive events are of significance only for the percipient and that they are closely related to very private, recurrent themes in his or her life. This corresponds to the general conception of dreams as defined by Adler, who considers a dream as "the product of life-style" (Ansbacher & Ansbacher, 1975, p.225). According to this conception, a precognitive dream is to be regarded as a special case of a ´normal´ dream. This, however, does not mean that catastrophes or other mishaps, that are not related to the percipient, cannot be sensed precognitively. An example can be found in the present material - the foreseeing of an accident in a mine. But all the other affective events, which are negatively loaded, are closely related to the subject. It seems as if the observer of a continuous series of dreams is more aware of precognitive correspondences and it might be that he or she would never have recognized them without this continuous observation. This confirms Dunne´s assumption (1935). Probably, only the outstanding events (e.g. ´serious and shocking events´) would be memorized by individuals who do not observe their dreams and possible related future events so systematically. However, without equivalent material, this remains a mere hypothesis.

The present survey emphasizes the problem as to how one should evaluate subjective impressions by objective methods. Statistical analysis does not yield positive results and, therefore, cannot support the assumption that presumably precognitive dreams are characterized by specific patterns.

Since the subjective impression of precognitive elements in the dreams is very strong - not only the subject was really convinced of her ability of precognition, but also the researchers who judged her dreams were persuaded in many cases - the question arises, why there is no possibility to gain ´objective´ evidence? The main problem lies in the unavailability of a sample of ordinary, random, non-psychic dreams and possible random confirmatory events. These data are not available. As we know such events are mostly not public but private and, usually, selected by the percipient. Independent blind judging can hardly be performed.

Thus, we can formulate four remaining questions:

(1) Are our ´usual´ statistical methods only of limited value with regard to such phenomena?
(2) Is the subjective impression misleading?
(3) Are obvious correspondences only generated by psychological mechanisms which lead to the erroneous impression of precognition?
(4) Are precognitive dreams, in spite of the results of this study, nevertheless characterized by specific patterns (or ´tracers´) which allow for a differentiation between precognition and ´pure´ coincidence?

At the moment at least we have to agree with Rush, whose general statement on psi-experiences is: "Evidently no sharp boundary can be drawn between ordinary subjective experiences and those involving a psi component" (1986, page 67).

NOTES

1) The research work was done during the author´s stay at the ´Institute for Border Areas of Psychology and Mental Hygiene´ in Freiburg i.Br. (1985-1986). A complete presentation of this work with the dream material can be found in Schriever (1987).

2) We did not take the ´Gotenhafen dreams´ into account, since they have a different quality of confirmation. Bender & Mischo (1960/61, 1961/62) carried out extensive interviews with witnesses. They had much more material for each dream than we have. The 11 ´Gotenhafen dreams´ among our 126 dreams all belong to the categories of ´well´ and ´very well´ confirmed dreams as mentioned below. So there are 115

dreams left for our analysis.

3) This and all the following tests are based on the unweighted numbers of the relevant characteristics which occur in the well- and badly-matched dreams.

ABSTRACT

In the present study an extensive selection of 1700 dreams, reported by a single subject, were analyzed with the objective to find precognitive elements. The subject claims to have observed precognitive elements in a number of her dreams. A small number of presumably precognitive dreams were found by means of a qualitative categorization. These dreams show a detailed correspondence with their confirmation (i.e. they are well-matched). Based on work in this field that suggest common structures in precognitive dreams, these 48 well-matched dreams were analyzed with regard to the following characteristics: Contents, vividness, dream recall frequency, the dreamer´s awareness of their significance, life situation of the subject, quality of the dream images, time interval between precognition and its supposed confirmation. The well-matched dreams were compared as regards these characteristics to 67 badly-matched dreams which presumably are not precognitive.

No specific patterns for the presumably precognitive dreams were found. The dreams considered as precognitive, as well as those which were not considered precognitive, mainly refer to the subject´s private spheres of life, which are only of significance to herself. In most cases the dreams concern matter of daily life. Rarely do they involve ´serious and shocking events´. Consequently, the author feels that the hitherto common image of precognitive dreams, in particular with regard to the contents of events experienced in precognition, should be revised.

Since no significant difference between well-matched and badly-matched dreams could be found, except for the quality of dream images, further research needs to be done, especially with regard to the problem as to how to objectify such subjectively found correspondences.

ACKNOWLEDGMENT

This paper was presented at the 30th Annual Convention of the Parapsychological Association, August, 1987, Edinburgh, U.K.

REFERENCES

Ansbacher, H.L. and Ansbacher, R.R. (Eds) ´Alfred Adlers Individualpsychologie. Eine systematische Darstellung seiner Lehre in Auszuegen aus seinen Schriften´, Muenchen/Basel, Reinhardt, 1975.

Bender, H. ´The Gotenhafen case of correspondence between dreams and future events: A study of motivation´, International Journal of Neuropsychiatry, 1966, 2, 398-407.

Bender, H. and Mischo, J. ´´Praekognition´ in Traumserien. Dokumentation und Strukturanalyse sinnvoller Koinzidenzen im ´Fall Gotenhafen´, Zeitschrift fuer Parapsychologie und Grenzgebiete der Psychologie, 1960/61, 4, 144-198, and 1961/62, 5, 10-47.

Dean, E.D. ´Precognition and retrocognition´, in Mitchell, E.D. and White, J. (Eds): Psychic exploration: A challenge for science, New York, Putnam´s Sons, 1974, 153-177.

Dunne, J.W. ´An experiment with time´, London, Faber and Faber, 1935.

Greenhouse, H.G. ´Premonitions: A leap into the future´, London, Turnstone Press, 1972.

Hall, C.S. and Van de Castle, R.L. ´The content analysis of dreams´, New York, Appleton-Century-Crofts, 1966.

Hastings, A.C. ´Dreams of future events: Precognition and perpectives´, Journal of the American Society of Psychosomatic Dentistry and Medicine, 1977, 24, 2, 51-60.

Lucadou, W. v. ´Experimentelle Untersuchungen zur Beeinflussbarkeit von stochastischen quantenphysikalischen Systemen durch den Beobachter´, Frankfurt a.M., Haag und Herchen, 1986.

Mayring, P. ´Qualitative Inhaltsanalyse. Grundlagen und Techniken´, Weinheim/Basel, Beltz, 1983.

Orme, J.E. ´Precognition and time´, J. S.P.R., 1974, 47, 351-365.

Palmer, J. ´Extrasensory Perception: Research findings´, in Krippner, S. (Ed): Advances in parapsychological research, Vol.2: Extrasensory Perception, New York/London, Plenum Press, 1978, 59-243.

Priestly, J.B. ´Man and time´, London, Aldus Books, 1964.

Rush, J.H. ´Spontaneous psi phenomena: Case studies and field investigations´, in Edge, H.L. et al (Eds): Foundations of parapsychology, Boston/London/Henley, Routledge, Kegan and Paul, 1986, 47-69.

Saltmarsh, H.F. ´Foreknowledge´, London, Bell and Sons, 1938.

Schriever, F. ´Zur Untersuchung praekognitiver Trauminhalte´ (Investigation of precognitive dream contents), Diplom-Arbeit im Fach Psychologie, Universitaet Bielefeld, 1987 (unpublished).

Stevenson, I. ´Precognition of disasters´, J. A.S.P.R., 1970, 64, 187-210.

Tenhaeff, W.H.C. ´Der Blick in die Zukunft. Praekognition´, Berlin, Universitas, 1976.

Dipl.-Psych. Friederike Schriever,
Wissenschaftliche Gesellschaft
zur Foerderung der Parapsychologie (WGFP).
c/o Hinbeersteig 64,
D-1000 Berlin 38
West Germany.

EUR.J.PARAPSYCHOL., V.7, 1987

DULLING OCCAM'S RAZOR: THE ROLE OF COHERENCE IN ASSESSING SCIENTIFIC KNOWLEDGE CLAIMS

John Palmer
Institute for Parapsychology

The psi controversy can profitably be considered in the context of the larger debate in philosophy between empiricism and rationalism. Science inevitably involves both. Scientific inquiry always begins with empirical observations, yet these observations must be submitted to interpretation based upon logic as well as the application of certain principles or rules that can be subsumed under the heading 'scientific method'.

Although science involves both empirical and rationalistic elements, their ratio can vary widely. In a radical empiricism, observations are taken pretty much at face value and interpretation is kept to a minimum. In a radical rationalism, observations are interpreted in the light of some prevailing theory or set of principles, irrespective of their face value. To a radical empiricist, facts are sacred; to a radical rationalist, theories are sacred. In practice, these positions are two ends of a continuum, and most of us fall somewhere between the two extremes.

PARSIMONY AND COHERENCE

It is noteworthy that most conventionalist critics of parapsychology call themselves rationalists. At least when it comes to parapsychology, I think many of them could be called radical rationalists. In any event, they appeal strongly to a priori principles to support their claim that psychic phenomena can be adequately explained conventionally. Specifically, they often invoke a version of the parsimony principle (also called Occam's Razor) which requires that when two explanations are offered for a given effect, one of which is consistent with already existing scientific principles and the other not, the former should be preferred.

In its more general form, parsimony states that the simpler of two competing explanations should be preferred, but the term simpler is inherently ambiguous. In the particular application at issue, simpler is defined as 'most consistent with established principles' Elsewhere (Palmer, 1986b), I have labeled this subspecies of the parsimony principle as the coherence principle, and I will retain that term here.

In a recent issue of the Skeptical Inquirer, no less than three major articles were devoted to explicit discussions or applications of the coherence principle (Flew, 1986; Klass, 1986; Shneour, 1986). Of these, the most relevant to parapsychology was an article by the philosopher Antony Flew, who uses the success of conventional science to justify a strong application of coherence in adjudicating scientific knowledge claims. He also objects to the characterization of his position as a priori:

> ... it is simply grotesque to complain, in the absence of ... decisive falsifying evidence, that these appeals to (C.D. Broad's Basic Limiting Principles) and the named laws of established physics are exercises in apriori dogmatism. For what 'apriori' means is: prior to and independent of experience. But in both of these kinds of cases we have an enormous mass of experience supporting our present beliefs and our present incredulities. (p.322)

I certainly agree with Flew that the BLPs and the present laws of physics are supported by an "enormous mass of experience" and in that sense are not a priori. However, the fundamental principle at issue is neither the laws of physics nor (strictly speaking) the BLPs, but rather the principle of induction, which (as used here) says that

because the laws of physics and the BLPs do indeed apply to an enormous mass of experience, we therefore are entitled to assume that they are universal. Whatever one thinks the status of induction should be in science, it is clearly an a priori, not an a posteriori, principle.

However, there is a more fundamental problem with this line of argument. Assuming that we grant to induction a role in science, surely we would require more than an "enormous mass of experience" in favor of a principle or law in order to justify a universal inductive inference from it. I think we would also insist that there be no known experiences that the principle or law does not account for.

Are there in fact no experiences that the laws of physics (incorporating the BLPs) do not explain? What about psychic experiences? Numerous surveys have recorded that such experiences not only occur, but that they are quite common (e.g., Greely, 1975; Palmer, 1979; Prasad and Stevenson, 1968). Even if we concede that a large percentage of psi experiences are attributable to coincidence, delusions, etc., sufficient cases remain to exclude them from the category of exceedingly rare events. Moreover, as Stanford (1974) has cogently argued, potentially paranormal interactions that are recognized as such are likely just a small percentage of those that actually occur. Clearly, if the induction underlying the coherence principle is to be maintained, it must be argued that these experiences fall within the domain of the BLPs and the present laws of physics, that is, that they have conventional explanations. But how is this latter conclusion arrived at? Not by empirical research (which the radical rationalist considers superflous for this purpose) but by appeal to the coherence principle itself: The preferred interpretation is the one consistent with already established principles.

We have now uncovered the fatal flaw in using coherence to validate knowledge claims: It is patently circular. How do we know that conventional scientific theory explains psychic experiences? By applying the coherence principle. Why are we justified in applying the coherence principle? Because conventional science explains all of nature (which includes psychic experiences).

One might object to the above argument by maintaining that even though psychic experiences have not been shown to require conventional explanations, neither have they been shown to require paranormal explanations. Thus, we have yet to prove any exceptions to our current

physical laws and, therefore, the basis for assuming their universality remains.

Although the premise is true, the existence of phenomena that appear to violate physical laws should be sufficient to raise enough doubts about their universality to require that the phenomena be subjected to empirical test and shown to be conventional before the universality is assumed. Classically, the induction principle is used to make predictions based upon universally consistent past experience. For example, all crows we have seen in the past have been black, and, therefore, we can predict that all crows we see in the future will be black. But this is not the application of induction that I am challenging. To represent the application at issue, which perhaps could be more precisely labeled as generalization, we would have to modify the example as follows: Most crows are black, but a small percentage appear to be white; nonetheless, since the vast majority are black, it is safe to assume that the white crows are not really white, but that the observations were inaccurate, someone captured a few crows and dyed their feathers white to trick us, or there is some other ´rational´ explanation we haven´t thought of yet. I submit that this is not a proper use of induction or generalization.

THE EXTRAORDINARY PROOF MAXIM

To make a philosophical point, I have been attacking an extreme form of the conventionalist position. Not even Flew, for example, is willing to entirely eschew empirical evidence in favor of the coherence principle. The more popular view nowadays among conventionalists is that the coherence principle should be used to define the degree of empirical evidence needed to support a given claim. In other words, claims that seem unlikely on a priori grounds require more empirical support than do more likely claims. In Flew´s case, for example, the requirement is repeatability on demand. This larger view is often represented semi-colloquially by the maxim ´Extraordinary claims require extraordinary proof´, where ´extraordinary claims´ is defined with reference to the coherence principle. I will call this proposition the extraordinary proof maxim, or EPM.

Although the EPM is a weak version of the position I have just critiqued, it suffers from the same logical difficulties. Its

application in the scientific decision-making process could produce in principle (and may already have produced in practice) serious biases in the scientific research literature.

Perhaps my point can be explained by an example. Let's say we have two theories, A and B, and that Theory B is superior, in the sense that it better explains all the data ultimately assembled. Let's assume further that the two theories predict the same primary effects but different secondary effects. Each theory has a group of adherents who try to promote and verify it. The adherents to Theory A get a head start and establish an impressive rate of success for their theory in confirming primary effects. Because of this track record, Theory A comes to be highly regarded by conventional science and its truth is considered 'a priori likely.'

Adherents to Theory B, noting that their theory predicts the same primary effects but different secondary effects than does Theory A, later initiate a research program to test their predictions of such effects. Let's assume that the adherents to Theory A do likewise. (If they were radical rationalists they would probably ignore this step, but I don't need this attribution to make my point.) The research of the two sides is equally competent.

Because of the EPM, more demanding methodological standards are applied to the research of the Theory B adherents than to the research of the Theory A adherents. As a result, the research supporting Theory A is more frequently accepted for publication in the respectable literature than is the research supporting Theory B, which is passed off as 'sloppy'. As this process continues, the gap between the accepted evidence for the two theories widens, thus causing Theory B to be seen as even more extraordinary, and the methodological standards it must meet even more demanding. In other words, a snowball effect occurs, and it eventually annihilates Theory B.

Just the opposite would have happened had it not been for the EPM. If the research supporting Theory B had been judged by the same standards as that supporting Theory A, Theory B would have eventually prevailed, because (as stipulated for the example) it in fact predicts the secondary effects better than does Theory A.

I do not claim that this example fits parapsychology in all respects, although I think it fits it in many respects. My point, rather, is that the EPM is logically flawed and should not be applied

to any scientific problem. Its consequences are most damaging, of course, when the actual circumstances approximate those of the example.

SOME QUALIFICATIONS

Does the coherence principle have any place at all in evaluating competing scientific claims? Let's change our example to say that Theory A and Theory B explain the secondary effects equally well, and this is reflected in the fact that after thorough testing, without the EPM, neither theory emerges as clearly superior. At this point, it would be quite appropriate to resort to a priori conventions to declare a winner. Thus if Theory A fits in better with the rest of science than does Theory B, or it is more 'elegant' in requiring fewer core assumptions, it should be preferred.

Inductions based on the coherence principle are also appropriate in a vast number of trivial cases where the observations to be explained are not anomalous with respect to the BLPs or established physical laws. This is true even in the absence of empirical research bearing directly on their interpretation. To be considered anomalous, observations must appear to violate the relevant established principles when taken at face value, and the (generally ad hoc) conventional interpretations offered to account for them must be scientifically inadequate. I have discussed my definition of the term anomalous more fully elsewhere (Palmer, in press).

One implication of this last qualification is that a priori principles can be used to reject totally ad hoc theories. Consider, for example, the old controversy between dualistic and materialistic conceptions of behavior. The contrast between these two conceptions can be schematically represented as follows:

Dualism:	Mind	-->	Brain	-->	Behavior
Materialism:			Brain	-->	Behavior

To begin with, note that both conceptions predict the same thing about the relationship between behavior and the brain. Thus, the success of the brain sciences in demonstrating the dependence of behavior upon brain states and brain activity cannot be used to support materialism

or refute dualism. However, one can use the parsimony principle to reject dualism, on the grounds that it postulates a construct that is superfluous (viz., mind), but only if it can be maintained that the postulation of mind is strictly ad hoc -- i.e., there exist no data that are anomalous with respect to materialism. If, on the other hand, such anomalous data do exist, and they are potentially explainable by a testable dualistic theory or model, the resolution of the controversy should depend upon the empirical success of the dualistic theory or model in accounting for the anomalies. Some, most notably Beloff (1980), would argue that psi by its very nature cannot be explained materialistically, but I will not try to address that thorny issue here.

Finally, my critique of a priori probabilities applies only to their use in dictating the kind of evidence necessary to verify scientific knowledge claims. I do not deny that such probabilities exist, and I acknowledge that they may have some value in helping one decide which hypotheses are most likely to yield payoffs if they are pursued. Thus, I can sympathize with scientists who might not wish to pursue paranormal hypotheses as explanations of psi anomalies on the basis of such considerations, although I think they also should take into account the failure of conventional theories to provide adequate explanations of such anomalies in many instances. As I have argued elsewhere (Palmer, 1986a), the ideal approach is to pursue psi anomalies from multiple theoretical perspectives.

THE HISTORICAL APPROACH

An objection to my whole line of argument might be raised by some who advocate the recently popular historical approach to the philosophy of science. Is it not true, they might argue, that, historically, science has embraced the EPM, and is it not also true that science has been one of the most dramatically successful enterprises in history? Yes on both counts, but has science prospered because of its conservatism (of which the EPM is one manifestation) or in spite of it? The value of conservatism in science cannot be assessed by comparing the accomplishments of science to the accomplishments of non-scientific approaches to knowledge acquisition, but only by comparing them to the hypothetical accomplishments of a less conservative science. The problem, of course, is that such a comparison cannot be made in practice, because we have no way of

knowing what the less conservative science would have accomplished in the same amount of time. But is it not at least plausible, for example, to suggest that a less conservative science would have more quickly acknowledged the limitations of, say, the phlogiston theory than did the more conservative science? I will leave it to those more knowledgeable than myself to debate this example. My point is simply this: The crucial comparison that must be made if the EPM is to be defended on strictly historical grounds cannot in fact be made, and therefore a strictly historical approach to the problem is not fruitful.

CONCLUSION

The position for which I am pleading is that standards of evidence should be as uniform as possible throughout science. I recognize that it may be unrealistic to demand that the same standards be applied in, say, psychology as in chemistry, because the phenomena studied by the latter are inherently more reliable than those studied by the former. However, the principle can be and should be applied to theories that purport to explain the same data. This in no way implies that the standards should not be rigorous. It simply means, for example, that if certain standards are applied to research designed to confirm a paranormal interpretation of psychic experiences, the same standards should be applied to research designed to confirm a conventional interpretation of such experiences -- and vice-versa.

ABSTRACT

Conventionalist critics of parapsychology often appeal to what I call the coherence principle, a subspecies of the parsimony principle, to support their claim that ostensible psychic events have adequate conventional explanations. In this paper, I have argued that this application is based on circular logic. A weaker version of the application is embodied in the often-heard maxim 'Extraordinary claims require extraordinary proof'. An example was used to illustrate how application of this maxim can in principle cause serious biases in the scientific research literature. Nonetheless, the coherence principle and a priori probabilities do have a role to play in science. Finally, it was argued that a strictly historical approach to defending the

maxim is not fruitful.

NOTE

This is a revised version of a paper presented at the 30th Annual Convention of the Parapsychological Association, Edinburgh, Scotland, August, 1987.

REFERENCES

Beloff, J. 'Could there be a physical explanation for psi?', J. S.P.R., 50, 1980, 263-272.

Greely, A.M. 'The sociology of the paranormal: A reconnaisance', Beverly Hills, Calif., Sage, 1975.

Flew, A. 'Parapsychology, miracles, and repeatability', Skeptical Inquirer, 10, 1986, 319-325.

Klass, P.J. 'The Condon UFO study: A trick or a conspiracy?, Skeptical Inquirer, 10, 1986, 328-341.

Palmer, J. 'A community mail survey of psychic experiences', J. A.S.P.R., 73, 1979, 221-251.

Palmer, J. Letter in J.o.P., 50, 1986a, 302-303.

Palmer, J. 'Progressive skepticism: A critical approach to the psi controversy', J.o.P., 50, 1986b, 29-42.

Palmer, J. 'Conceptualizing the psi controversy', Parapsychology Review, in press.

Prasad, J. and Stevenson, I. 'A survey of spontaneous psychical experiences in school children in Uttar Pradesh, India', International Journal of Parapsychology, 10, 1968, 241-261.

Shneour, E.A. 'Occam's razor', Skeptical Inquirer, 10, 1986, 310-313.

Stanford, R.G. 'An experimentally testable model for spontaneous psi events. I. Extrasensory events', J. A.S.P.R., 68, 1974, 34-57.

J. Palmer
Institute for Parapsychology
F.R.N.M.
College Station
Durham, N.C. 27708
U.S.A.

EUR.J.PARAPSYCHOL., V.7, 1987

WHEN WILL WE BEGIN TO REDUCE ALPHA - AND BETA - ERRORS IN STATISTICAL PSI EXPERIMENTS ?

Ulrich Timm
Institut für Grenzgebiete der Psychologie und Psychohygiene
Freiburg
West Germany

I.

For over twenty years I have been intensely involved with the methodological problems of planning and evaluating psi experiments and have presented a large number of articles in that area, though most of them are unfortunately written in German. My primary concern may be summarized as follows: How can the structure and statistical analysis of psi experiments be so designed as to minimize interpretational and statistical errors while still maximizing interpretability and statistical efficiency? In the process of so doing I had to take a good look at the current practice of statistical evaluation and attained, even years ago, some fairly provocative conclusions:
1) In many psi experiments some rather grave (especially statistical) mistakes are made, after whose correction, the initial statistical significance often disappears. Thus we get an increase of the type I errors, more simply called alpha-errors.
2) That does not necessarily mean, however, that real psi effects do not exist at all, since the usual methods, if utilized correctly, are often so ineffective - with regard to the rareness, instability, and

inconsistency of psi effects - that they can only seldom lead to statistical significance. This inefficiency of statistical methods yields an unnecessary increase of type II errors or beta-errors.
3) In the course of many decades, suggestions made to decrease the alpha- and beta-errors by replacing the current procedures with more correct and efficient ones, have proven virtually unsuccessful. I have personally experienced this mainly in Germany, but I know that, outside of Germany, others have had the same experience.
4) The main reason for the notorious emphasis of parapsychologists on dubious statistical methods, is that these methods are strongly sanctioned by the tradition of the behavioural sciences and are recognized and criticized by only few scientists even outside the field of parapsychology.

Naturally one should expect that I would precisely substantiate and clarify this severe criticism. Unfortunately, due to the difficulty of the statistical material and the short length of time, it is here only partly possible. However, after the presentation of some smaller papers (1980, 1982), I gave an extensive discussion in a German article which appeared in 1983 in the ´Zeitschrift fur Parapsychologie und Grenzgebiete der Psychology´; and I hope to soon publish an improved version in English. This article has nothing to do with elementary statistical mistakes usually made by beginners. It focuses instead on a higher order of alpha-errors which I regard as especially dangerous because they are partially made also by experienced scientists and have become, in the course of time, the common property of whole scientific disciplines. I call this class of errors ´statistical selection errors´ and, for that reason, the title of my article was ´statistical selection errors in parapsychology and other empirical sciences´. Although these selection errors are made everywhere, they are, in my opinion, so serious in parapsychology that for years I have been fighting an intensive, but futile, battle against them. My situation vaguely resembles that of Don Quixote fighting against the wing of the windmill.

I will stress, however, that in no way do I want to invalidate the whole of parapsychology as typical psi skeptics do. True, I contend that many psi experiments prove to be statistically insignificant after the alpha-errors are corrected. But, since I am battling the beta-errors as well, I further contend that, by choosing more efficient statistical methods, one could detect even very small and inconsistent psi effects in a flawless manner. Thus I devoted a considerable part of my article to the description of more effective

evaluation methods, which would make the renunciation of well loved selection errors more tolerable for parapsychologists. In short, my objective is not only the reduction of alpha-errors and the related decrease of spurious significances, but also the reduction of beta-errors and the related increase of real significances.

II.

Following these introductory comments, I first want to give an overview of the statistical selection errors in parapsychology. I will limit the concept of selection error to the point that it will always overestimate the significance of a result or underestimate the probability alpha of the type I error. Since in parapsychology the significance level alpha is seldom fixed in advance, but instead the P-value of a test statistic (e.g. of the CR) is given, we can here, correspondingly, speak of an underestimation of the P-value.

Such a statistical selection error shows, simply stated, the following three qualities:
1) From a set of statistical results a single result is selected and evaluated by some significance test.
2) The selection is not performed randomly, but according to a criterion which is related to the level of the single results in that it directly or indirectly favors the positive results.
3) Despite this success-dependent selection, the significance test is carried out and interpreted in the usual manner without any correction.

Following this simple recipe it is almost always possible, even in investigations whose results are purely random, to find some kind of ´significant effects´. In practice it means that one picks out apparent significances from several subsamples of subjects, from several sections of an experiment, from several replications of an experiment, from various experimental variables, from various significance tests etc., and finally holds that these particular selected data show a significant effect while the rest is insignificant. If one finds, for example, among twenty independent statistical results, one single one which is in excess of the 5% significance limit, then one should conclude that this corresponds exactly to the chance expectation. One may, however, single out that particular result and declare it as significant: then one will have

made an exemplary selection error! In contrast, the correct evaluation would consist in a statistical analysis of the total result. Through such a global significance test every statistical selection error will automatically be avoided. But one can also apply a correction formula to individual selections.

A look at experimental parapsychology immediately shows that it supplies virtually unlimited possibilities for making such selection errors. This is already true for simple standard experiments containing only one variable, for example a number of hits which is evaluated with a CR. Here, at the very least, the following selection errors are possible:

A: Intraexperimental selections

1) The selection of single temporal sections of an experiment, for example single ´runs´, single ´sessions´, etc. These can usually be additionally differentiated according to a multitude of coincidental conditions (such as change of supervision, change in group composition, change in time of day etc.).

2) The selection of single subjects from the total group.

3) The selection of single significance tests from several tests which, e.g., may respond differently to score fluctuations, to special scoring patterns etc.

B: Interexperimental selections

4) The selection of single experiments from the total number of all replications of an experiment.

5) The selection of single kinds of experiments from the total number of all psi experiments.

These or similar selection errors appear with various frequency and with various arguments in most of the experimental papers from parapsychology. But however intelligent the arguments might be: they are always questionable from a statistical point of view. True, there seems to be a plausible argument that one would be allowed to separately test the significance of single experimental sections, single subjects, single experiments etc. in parapsychology. One says, namely, that the separate results are not homogeneous due to the great intra- and interexperimental variability of psi performance. Heterogeneous results, one says further, need not be combined since each time one tests a different hypothesis. I took an especially good look at this argumentation and can not, unfortunately, accept it without objection: The significance test of a statistical experiment always refers to the null hypothesis; and, in the case of complex experiments, which can be broken down into a number of parts, there

usually exists a whole hierarchy of null hypotheses. Then, any subordinate null hypothesis is to be interpreted as a special case of a superordinate null hypothesis and can only be rejected if the superordinate null hypothesis has already been rejected. Correspondingly, the subordinate results, in reference to all superordinate null hypotheses, are to be classified homogeneous and can only be separately tested when all of the superordinate results have become significant.

In parapsychology, one can even formulate such a general null hypothesis that it is superordinate to each and every psi experiment. It simply states that psi phenomena do not exist at all. Thus, to evade selection errors, one had to combine all of the psi experiments up to that point and let them undergo a global significance test before one is allowed to interpret them separately. Even if one assumes that, meanwhile, the existence of psi has been established, one must, in any case, test the total result of every single experiment, since the psi effect is said to vary between experiments and consequently must not necessarily appear in each of them. Only if the total result is significant, is one allowed, then, to test the significance of results from separate subjects, and only if it is ensured, may one finally ask whether they were, perhaps, significant in single sections of the experiment.

Until now, the talk was only of simple standard experiments, which exclusively serve the purpose of giving evidence for the existence of psi. However, the same possibilities of error exist also in the case of differential or correlational psi experiments, which examine differences between various experimental conditions or correlations between psi variables and other variables. Here, one was previously content with rather simple experimental designs testing, for example, the significance of the difference between 'sheep' and 'goats' or the correlation between the number of hits and a score of extraversion. However, in recent years, the interest has turned to multivariate designs containing many experimental conditions, personality- and psi-variables. Here, the same principle of hierarchy as discussed previously is valid: wherever a meaningful superordinate null hypothesis exists, it must be rejected before separate experimental effects, correlations etc. are allowed to undergo a normal significance test. One could, for instance, formulate the superordinate null hypothesis that psi scores do not correlate with any other variables. Even if today one rejects this general hypothesis, one must still, from similar considerations made above,

demand the calculation of global significance tests for almost all correlational experiments. This can be done, for instance, through a multiple or canonical correlation, in which the psi variables serve as criteria and the other variables as predictors. If one abstains from this, they will find in every larger set of predictor variables some significant correlations with some psi variables; but one is not allowed to single them out and interpret them in the usual manner. If done, they will make a selection error and could possibly fall victim to a statistical artifact. If the apparently discovered effect is not replicated in the next experiment, this corresponds with the statistical expectation and naturally has nothing to do with the ´non-repeatability´ of psi.

One may object to this discussion that sophisticated experiments are carried out in a much more refined manner. Here, in advance, one formulates certain hypotheses which correspond to expected correlations or differences within the results. In the evaluation, one limits himself to these hypotheses. In this case selection errors are said to be excluded and only possible if one tests post-hoc-hypotheses. Unfortunately, this argument is also not correct. It is true, that one limits the evaluation possibilities through these preformulated hypotheses, which is very recommendable. However, if one has formulated enough hypotheses, they still have, amongst these hypotheses, enough possibilities for selection. One must, for that reason, here also carry out a global significance test for such single hypotheses to which a superordinate null hypothesis can be assigned. Otherwise selection errors are also possible and the difference from the testing of post-hoc-hypotheses is only a matter of degree.

Although many other selection errors could be described, I hope to have said the essential part. Above all, it should be clear that by performing global significance tests many psi experiments must lose their significance. I remember, though, that I also mentioned the interexperimental selection above, to whose avoidance, at the least, all similar psi experiments should be combined and submitted to a global significance test. Through this combination, on the other hand, the significance may increase so that the single experiment loses part of its meaning. A few years ago, there would have been more to say about the worthiness and practical problems of combining separate experiments. In the meantime, however, this technique has already established itself under the key term ´meta-analysis´, even though the statistical combination methods favoured in this area are not in full agreement with those proposed by me.

III.

Thus I want to turn to my second theme, that is the reduction of beta-errors in the statistical evaluation of psi experiments. This point is complementary to, and directly combines with, my explanation of statistical selection errors. The problem is to increase the statistical efficiency (or power) of the significance tests in such a way that -despite the avoidance of selection errors- minimal psi effects can be statistically detected. In doing so, I confine myself to two different questions both of which are of considerable importance to the practice. The first question is: which are the statistically optimal methods for correcting a given selection or for combining single results which are to undergo a global significance test?

Here, first it can be answered that for any selection of a single result there is a simple statistical correction possible which replaces the global significance test. An approximate formula for this purpose requires that one multiplies the P-value of the selected result by the number of the given results. Naturally, in this manner, the P-value will be strongly increased so that the statistical significance will, in most cases, disappear, as in the case of a global significance test. There are, however, certain exceptional psi experiments in which sheerly 'astronomically' low P-values have been obtained. Thus one could, theoretically, pick out one from these extremely significant experiments and multiply its P-value of perhaps .000000001 with the estimated number of all (published and unpublished) psi experiments conducted up to this day. The significance would then certainly not be lost, even if the P-value ascended to, say, .00001. However, this method is not especially recommended for giving general evidence for the existence of psi because everything is then based on the reliability of one single experiment which could possibly become falsified. Such a thing emerged drastically when Dr. Thouless applied this formula to an experiment of Dr. Soal in 1963. With selections from within an experiment this formula can, however, be used as is, wherein its efficiency is dependent on various factors which cannot be described here. In any case, this is a universal and very simple method of correcting statistical selection. It is astonishing that it is so seldom used in the field of parapsychology.

Most of the other methods consist in weighted combinations or

non-linear functions of the single results to attain a most efficient global significance test. In the case of standard psi experiments that seems trivial because one needs only to add the different hits, whose sum can be evaluated with a CR just as well as the separate results. However, an analysis if intra- and interindividual distributions of psi scores shows that the simple addition of hits is one of the statistically least efficient methods, even for the aggregation of small experimental units such as individual runs. The reason for this lies in the strong variability of psi scores, which can even vary in a bipolar fashion between psi-hitting and psi-missing so that the hit deviations cancel out each other. Although this fact had been known for a long time, no one before, astonishing enough, has derived decisive consequences from it.

However, at the latest then, when one wants to combine all of the psi scores of an experiment (or even several experiments) to avoid selection errors, one should derive these consequences and introduce statistically more efficient measures of psi performance. I have occupied myself for years with this problem and derived several usefull measures. Finally, following the method of the likelihood quotient, I came to a measure which is statistically most efficient for strongly varying psi scores. To calculate this, one must split up the trials of an experiment according to empirical criteria into a number of k sections and then sum up the squared CRs from these sections. This sum, in which the single CRs are weighted according to their size, is chi-square distributed (with df=k) and a linear function of the well known ´run-score variance´. It may be extended over any number of experiments to attain an interexperimental significance test. Naturally one must establish the optimal section length before the calculation. One can even use individual subjects or whole experiments as units. By the way, I had already used this measure successfully for many years and always advocated the abolition of the misleading indication ´variance´, which is only correct under the null hypothesis. Several other measures are also more suited to combination in a global significance test than the usual number of hits; for instance the logarithm of the one-tailed P-value of each section (note 1). Therefore, and since one should not introduce too many psi variables at one time (which could lead to new selection errors), I have suggested to exclude the number of hits completely from the valid statistical evaluation.

The second question which I would like to briefly address refers to the problem whether or not there are permissible forms of selection,

which one could use to increase the statistical efficiency. This question should clearly be answered with a yes, and the relevant possibilities should without a question be taken advantage of. For example, the above definition of selection error allows one to exclude any partial result from the global significance test of an experiment, if the exclusion ensues according to a criterion that, under the null hypothesis, is independent of the respective results. If one, in this way, discovers certain clues that particular experimental situations, certain subjects, certain variables etc. could be unsuccessful, one is allowed to eliminate them as is. This can be a great advantage because every insignificant partial result reduces the significance of the total result. By the way, one should allow variables about whose suitability one is not certain, to 'run with' in an experiment without valid statistical evaluation, or perhaps use them to carry out a pilot study whose statistical evaluation one declares invalid in advance. If one does not have this possibility, they can (following the Edinburgh example) separate the experiment, before evaluation, into two halves, whereby the first half will be used exclusively to select variables or to build hypotheses and the second will be submitted to the global significance test.

In the global statistical evaluation of a multivariate experiment, one should, further, reduce correlated criterion or predictor variables to a smaller number of factors by performing a factor analysis, because the statistical efficiency, in the case of correlated variables, decreases with the number of variables. Finally, the so-called extreme-group method should be mentioned, according to which one is allowed to eliminate the middle cases of the distribution of a variable when calculating correlations. For example, one could eliminate all the chance-scoring subjects of a correlational study, if enough psi-hitters and psi-missers remain. The correlations between psi variables and other variables could, in that way, become much more significant. In all of these procedures we are dealing with permitted selections. Amazingly enough, though, they are less frequently used than the non-permitted selections which I have described above as selection errors.

IV.

Finally, I want to say that I am in a rather pessimistic mood: I am afraid my explanations will not lead to a decisive change in the

statistical methods of parapsychologists. When I pointed to the problem of statistical selection errors at the P.A. Convention in Reykjavik (1980), it also did not show any considerable effect. One must, apparently, turn to the psi skeptics to attain such effects. By the way, I have had the experience that there we also have a psychological problem: Many interlocutors exhibit surprising understanding for selection errors in other investigations while they do not recognize their own errors. From psychological research it is known that human perception and cognition also function in a highly selective manner. Thus, statistical selection errors may serve the general psychological tendency to synchronize the given empirical data with one's own expectations regarding reality. -Therefore, the final demand can only be to answer one's own ways of acting with increased self-criticism, even in such an objective area as mathematical statistics. Otherwise those cynics will be confirmed who always have contended that, with statistics, one can prove everything.

NOTES

1) The expression - Sum ln P is also chi-square distributed, but with df=2k. This measure is recommended, if we expect varying psi scores of known (positive or negative) direction. Therefore one-tailed P-values should be used in this case.

ABSTRACT

In the first part of this paper, some so-called statistical selection errors are discussed, which lead to an increase of the type I error. These are, in the author's opinion, so serious in parapsychology that they can be found in most statistical investigations published until today. After correcting those errors, the claimed statistical significance frequently disappears. This especially applies to many complex multivariate studies. -In the second part some evaluation methods are mentioned which, by increasing the statistical efficiency, are able to detect even small psi effects though avoiding selection errors.

ACKNOWLEDGMENT

This paper has been presented at the 30th Annual Convention of the Parapsychological Association, Edinburgh, August, 1987.

REFERENCES

Timm, U. ´Statistical selection errors in parapsychology´. Unpublished summary presented in a workshop at the 23rd Annual Convention of the Parapsychological Association, 1980.

Timm, U. ´Methodologische Probleme bei der Planung und Auswertung differentieller Psi-Experimente´. Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 1982, 24, 140-160.

Timm, U. ´Statistische Selektionsfehler in der Parapsychologie und in anderen empirischen Wissenschaften´, Zeitschrift für Parapsychologie u. Grenzgebiete der Psychologie, 25, 1983, 195-229.

Dr. Ulrich Timm
Auwaldstrasse 98
78 Freiburg (i.Br.)
B.R.D.

EUROPEAN JOURNAL OF PARAPSYCHOLOGY

VOLUME 7

NUMBER 2-4

1988 - 1989

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.

EUR.J.PARAPSYCHOL., V.7, 1988-1989

THE END OF THE PARAPSYCHOLOGY LABORATORY OF THE UNIVERSITY OF UTRECHT

Sybo A. Schouten
University of Utrecht

Due to considerable reductions in financial support in 1987/88 the Faculty of Social Sciences of the State University of Utrecht, to which the Parapsychology Laboratory is organizationally attached, was forced to implement a major re-organization of its activities. The aim was to adapt its tasks and means to new requirements and support levels. A first re-organization also aimed at reducing expenses and staff had taken place in 1982/83. The Parapsychology Laboratory did survive that one although not without damage: it lost one of its two positions for appointing guest researchers. The last re-organization, however, took a more heavy toll. The Faculty was forced to reduce its staff from 315 to 240 full-time positions. However, from this reduced number of positions about 136 had to be destined for new tasks, among others to enable appointments for temporarily research positions. All in all these measurements implied that the existing staff was reduced by about 2/3 rds, although some staff who lost their position found employment in some of the new tasks.

In this process, unfortunately, the Parapsychology Laboratory became one of the (many) victims, and the Faculty decided to terminate the Lab's activities. It is of interest to note that this decision was not based on negative feelings with regards to the relevance of the

subject matter of our research or on misgivings as regards our functioning. On the contrary, it was explicitly stated by the Faculty that they had no doubt as to the positive quality of the educational and research activities of the Parapsychology Laboratory.

Parapsychology became officially part of the curriculum of the University of Utrecht with the appointment in 1953 of Dr. W.H.C. Tenhaeff as Special Professor in Parapsychology on behalf of the Dutch Society for Psychical Research. Tenhaeff was strongly in favour of what he called the anthropological approach, which opposes the reductionistic tendencies found among others in experimental psychology and which as regards its methods gives preference to depth-psychological and phenomenological descriptions over experimental research.

A Special Professorship at Dutch universities is in many respects different from the regular professorship. It is a joint appointment by an organization and the university in order to promote at the university the field the organization represents. The Special Professorship involves only teaching and the only obligation the university has towards the Special Professor is to provide facilities to lecture. As a consequence the special professorship is financially, for all its activities, entirely dependent on the financial support provided by the organization involved. Needless to say that in the case of parapsychology such support has been virtually non-existent.

In the early sixties the University of Utrecht obtained a regular professorship in parapsychology. A regular professorship is financially supported by the state and involves all scientific activities which normally goes with a professorship, especially research and teaching, and of course the unavoidable bureaucratic and management duties. After a complicated process coloured by strong differences in opinion between Prof. Tenhaeff and the Faculty of Social Sciences a decision was taken to the effect, that the regular chair should devote itself to experimental research in parapsychology, in line with the developments in experimental psychology. As a consequence the Faculty looked for a candidate which supported the integration of experimental psychological methodology in parapsychology and found one in the person of Martin Johnson. Johnson started working in Utrecht at the Psychology Department of the Faculty of Social Sciences in 1971, together with Sybo A. Schouten who had been a staff member of the Psychology Department for parapsychology since 1966. The official appointment of Johnson as Professor in

Parapsychology did not come until 1973.

Because of the strong opposition of Prof. Tenhaeff against the appointment of Martin Johnson and the different view on parapsychology Tenhaeff held the University granted the request of the Dutch S.P.R. to maintain a Special Professorship at the university. After Tenhaeff, H. van Praag, a philosopher who viewed parapsychology as a field which could provide important contributions to a new vision of mankind, became the successor to Tenhaeff's chair. Prof. van Praag retired in 1987. The Faculty considered the two chairs as complementary, each covering different aspects of the field. The regular chair represented parapsychology as a scientific area of research emphasizing empirical and experimental research, the special chair represented Tenhaeff's qualitative and anthropological approach. By closing the Parapsychology Laboratory the regular chair in parapsychology is lost and as a consequence in 1988 Prof. Johnson retired and returned to Sweden. Hence only the special chair in parapsychology, which is at present vacant, remains but because of the complementary approaches of the two chairs its continuance is by no means certain. However, if the Faculty decides to maintain a special chair in parapsychology then according to the Faculty its task must change and the Special Professor has to represent the entire field of parapsychology. In addition, the Faculty wants the chair to fit into the social-psychological background and scientific philosophy of the Faculty.

After Johnson's appointment and the establishment of the Parapsychology Laboratory in the beginning of the Seventies, Johnson and Schouten agreed on the policy the Laboratory would adopt. In view of the conflict, the associated negative press coverage which accompanied Johnson's appointment and the resulting damage to the university and the name of the field it was decided to keep a low profile and to avoid at all cost new public conflicts. Therefore public appearances in Holland was left to the special professorship. For the same reason the Laboratory refrained from playing a role in Dutch parapsychological activities. Instead the following targets were set for the Laboratory's activities: The development of an experimental research program and the acquisition of facilities and equipment for laboratory research; to make research in parapsychology a scientifically acceptable endeavour in Dutch academic circles; and to concentrate on international rather than on national collaboration and recognition.

The organizational structure and the activities of the Laboratory were adapted to these goals. In view of the small size of the Laboratory the main tasks were divided over the permanent staff members and each member could carry out his tasks without consulting with the others as long as the general lab policy was followed. This policy saved a lot of time, not in the least because it eliminated the need for frequent meetings. Johnson was responsible for international contacts and research and educational policy, Schouten for all organizational contacts especially within the university, and Boerenkamp for dealing with persons who turned to the Laboratory for assistance regarding problems related to personal paranormal experiences. Later on Camfferman became a part time staff member, in charge of the laboratory equipment, and he also participated in various research projects.

Another consequence of the adopted policy was the reservation of two research positions for temporarily appointments of foreign guest researchers. This policy seemed to us beneficial in several respects. First it allowed us to benefit from the experience of skilled researchers in the field, an important advantage for a new laboratory with no tradition. In addition it strengthened our international ties. Another consideration was that this policy would counteract the possible isolationistic tendencies which sometimes threatens small research groups. Parapsychology is a small field with few institutes which are generally at a large distance of each other. In addition as a rule parapsychologists work at best in a neutral but often in a rather hostile environment. Because of these circumstances there can be a tendency in parapsychology to develop a defensive and isolationistic attitude which hampers the integration of parapsychology into mainstream science. A constant stream of guest researchers not only from parapsychology but also from other disciplines implies a continuous confrontation with new ideas and different opinions, which stimulates an open-minded attitude. And last but not least, guest researchers are important to ensure a regular research production. This is a necessity, because in a university setting, permanent staff members tend to become clogged with many organizational, educational, and other duties which have to be fulfilled mainly at the expense of research activities. As a result of this scheme, since 1974 twenty guest researchers from all over the world have been employed at the Laboratory. In hindsight one can conclude that they fulfilled all the expectations we had of this policy.

Because it appeared to be the most valid scientific approach, and therefore the best way to ensure integration of research in parapsychology into the academic world, a research policy was adopted based on a two-track approach. In the process-oriented two-track approach both psychological (coupled with biological, physical and other known processes) and parapsychological explanations for paranormal experiences and phenomena are considered. We felt that paranormal experiences should be studied like other human experiences and that such experiences can not be separated from other psychological processes going on in the individual. Paranormal experiences are not isolated incidents but are often highly meaningful to the experient in the context of her or his life. Not only are psychological variables and circumstances relevant for understanding paranormal experiences, but it is also to be expected that some ostensible paranormal experiences do not require a paranormal explanation and can satisfactorily be explained by familiar psychological processes. Thus psychological variables play a role in or might explain at least part of the phenomena we study. On the other hand the empirical and experimental data base in parapsychology seems to us strong enough not to reject a priori the possibility of a paranormal explanation. Hence psychological and parapsychological explanations may both contribute to an understanding of paranormal experiences and can be considered as supplementary or even as compatitive models. Research will eventually disclose which model contributes most to the explanation of these experiences.

In our experience this research approach to the study of paranormal experiences and phenomena does not provoke the usual hostile attitude in the academic environment and seems acceptable even to scientists who do not believe at all in possible paranormal effects. In addition, the study of paranormal experiences is our field and it seems illogical to leave parts of the field, in this case the study of the applicability of models of a non-paranormal character like psychological models, to skeptics or other non-parapsychologists (that is to say, scientists who do not specialize in the study of paranormal experiences).

Another activity which followed naturally from the goals set for the laboratory was the inception of the European Journal of Parapsychology. The European Journal of Parapsychology was primarily intended to be a strictly technical journal on parapsychological research which did not cater for a large public. As parapsychological activities in European countries often are not known in other

countries due to language barriers, we feel that a need remains for an English written European technical journal on parapsychological research which can serve as a bridge between these other language areas and the English dominated American and British professional literature.

But there were more reasons to start such a journal. One was the reason mentioned above: the journal was to serve as a bridge between other language areas and the English dominated international literature in our field. Another important reason was to implement Johnson's view on publication policies for scientific journals. In his inaugural speech as Professor of Parapsychology at the State University of Utrecht, held on November the 8th, 1974, entitled 'Models of Control and Control of Bias in Experimental Parapsychology' Johnson drew attention to the distorting influence traditional editorial publication policies can exert on a field of science. He illustrated it with an example from experimental psychology, drawing attention to the tendency in the social sciences to bury negative or non-significant results and to publish only results which 'turned out'. Needless to say that similar problems are not unknown in our field. So we formulated a different editorial policy for the European Journal of Parapsychology. It's main element was that selective reporting would be avoided by giving priority to the publication of research of which the acceptance of the manuscript had taken place prior to the phase in which the experimental data became collected. Hence acceptance or rejection could be based only on the quality of the design and methodology and not on the significance level of the outcome. In addition, this way we also ensured that the rationale of the study and the hypotheses on which the study was originally based could not afterwards become modified in the direction of the outcome.

An important element of our editorial policy was that we felt that authors should be left as free as possible in the way and the style they wanted to present their work. In general editors take an active part in a tidying up process of manuscripts. A research report can only provide a limited amount of (formally required) information. However, there are many other aspects associated with a study which are never discussed on paper but about which the reader tries to form an opinion in order to evaluate for her or himself the value of the study. For instance, is the experimenter a careful or a sloppy person? How qualified is the experimenter as regards the techniques which were applied in the study (randomization tests, using psychophysiological equipment). How accurately does the report depict the real course of

events in the experiment? In general the report itself is the only source from which an impression can be obtained as regards these questions. Hence one can argue that the tidying up process of editors sometimes serves more the interest and prestige of the journal than that it serves the reader. There are of course limits but as a consequence it seemed to us desirable that the report itself reflects as much as possible the personal style of the author.

It is of interest to note that recently the problem of selective reporting has been taken up in experimental psychology too. Kupfersmid discussed the 'statistical significance' problem in manuscript acceptance in the August 1988 issue of the 'American Psychologist'. This publication stimulated David J. Weiss, editor of the journal 'Applied Psychological Measurement' to devote an editorial (An Experiment in Publication: Advance Publication Review, APM, V.13, Nr.1, March 1989, p.1-7) to his idea, for the first time proposed by him in 1986, of the 'Advance Review Option', an evaluation procedure roughly similar to the publication policy of our journal. It should be noted, though, that Weiss' concern was not so much the possible distortion due to selective reporting but rather the considerable waste in research effort reflected in the large number of manuscripts turned down for publication. His 'Advance Review Option' might strongly improve the acceptance rate of manuscripts by forcing researchers to revise and improve their study prior to the data collection phase according to the reviewers comments. So clearly an editorial policy based on advanced reviewing of manuscripts serves different goals. It not only reduces the problem of selective reporting but it also serves to increase the quality of research and hence the research productivity. Therefore I hope that this editorial policy, which perhaps originated in our field, will be taken up and continued by other journals in parapsychology. Especially in a disputable area such as parapsychology it is important to face all kinds of challenges with an open mind and to introduce new and better ways of doing things. We might well compensate for our in many eyes suspicious subject matter by setting new and better standards in areas of the scientific process, as for instance methodology or editorial policies.

The European Journal of Parapsychology has always been closely related to the Parapsychology Laboratory and therefore the demise of the Parapsychology Laboratory has consequences for the publication of the European Journal of Parapsychology. One unfortunate consequence has been a serious delay in printing the journal. Fortunately the

Koestler Chair of Parapsychology of the University of Edinburgh has offered cooperation and support to the journal. Therefore it will be possible to continue with its printing.

The educational programmes offered by the Laboratory have always been part of the curriculum for students of the Social Sciences. Hence students received the usual academic credits for participating in our courses. In the Seventies and the beginning of the Eighties the study in the Social Sciences required 5 years to obtain the M.Sc. In that period students could participate in different courses in parapsychology, which ranged from introductory courses for freshmen to specialisation programmes which took about one year. Also supervision for Ph.D. research was provided.

Topics of teaching involved the developments of research in parapsychology, the experiences and phenomena we study, and the relationship between our field and the other disciplines of the Social Sciences, especially with psychology. In addition philosophy of science and methodological issues were often discussed. Since many students entered our courses with strongly inflated ideas about the paranormal, a critical attitude was encouraged, both with regards to inflated claims as well as towards skeptical prejudices. We favoured courses which remained not restricted to parapsychological issues but which had more general use to students, for instance, courses in which students were trained to design and carry out experimental studies. In our opinion methodological standards in parapsychology are on the whole at a reasonable level, at least when compared to the norm in the Social Sciences, if only because as a consequence of the critical attitudes towards our work researchers in parapsychology are in general rather sensitive to possible weaknesses in their studies. Hence the combination of methodology and parapsychology lends itself well to constructing meaningful courses of general relevance for students.

Because in the Netherlands parapsychology was taught only at the University of Utrecht, we also offered supervision for students from other universities. Unfortunately in recent years the formal study duration in the Social Sciences was reduced to 4 years for the M.Sc. Consequently the opportunities for following courses in parapsychology became strongly limited to those wishing to do some project in the field. Although participation varied it is estimated that, until the reduction in study duration became effective, on average each year about 50 students participated in one or more of our courses.

The research carried out in the Parapsychology Laboratory has been rather diverse and reflected the various main approaches in the field. The philosophy on which our research was based has been the above mentioned two-track approach, in which both parapsychological and psychological explanations for paranormal experiences are studied. An important contribution to the research effort has been made by our guest researchers. Guest researchers were invited on the basis of the person's qualities and general research interests. Only rarely were people appointed on the basis of a detailed research proposal. In general, prior to the appointment, a rather loose agreement was made about the line of research the candidate would pursue when employed by the laboratory, mainly to ensure that the work would fit into the broad research interests of the laboratory and that the facilities for that line of research were available.

Every research laboratory with limited capacities is faced with the somewhat basic decision whether to concentrate all efforts on one line of promising research (for instance Ganzfeld or psychophysiological studies), or whether to spread out the research effort over various approaches and topics. Especially in view of the still modest degree of our knowledge about parapsychological experiences and phenomena, and because of the at the time lack of a clearly successful experimental approach, we felt more comfortable with the latter research policy. Therefore we reasoned that apart from practical and methodological considerations, guest researchers should be left as free as possible to pursue research they themselves considered most promising. The permanent staff members meanwhile concentrated their research efforts on a few bigger, long-term projects.

In parapsychology three main areas of research can be distinguished. These are research in spontaneous paranormal experiences, research with gifted subjects such as mediums or healers, and laboratory research mainly with unselected subjects. In our Laboratory, research has been done in all these areas, with probably an emphasis on laboratory research in which guest researchers especially were involved. It is outside the scope of this paper to discuss extensively all the research projects carried out in the Laboratory. As stated above, we are much indebted for many ingenious studies to our guest researchers. However, in order to give an impression of our research philosophy and of the resulting kind of investigations, which perhaps might be considered typical for our Laboratory, I will discuss a few of our bigger research projects. Because of their long-term nature, these projects were run or supervised by permanent staff members.

Spontaneous paranormal experiences mainly involve general extrasensory perception, that is telepathy, clairvoyance or precognition. The frequency of occurrence of spontaneous psychokinetic phenomena is much lower. Hence our research has focussed on SPE's (spontaneous paranormal experiences) of the first type. In the study of SPE's two views of its importance and function dominate. The older one, supported by the authors of 'Phantasms of the Living' is that collections of SPE's in itself might provide proof of the existence of ESP. The other, supported among others by L.E. Rhine, considers the study of SPE's as useful mainly for the generation of hypotheses. Our research was based on a different rationale. One argument raised against research in parapsychology is that before ESP can be studied it first should be proven that it exists. Hence as long as we have no convincing evidence for ESP (for instance the repeatable demonstration) there is no argument to justify research in this field. But we felt that if it could be proven that SPE's are not merely a collection of random coincidences but that they are characterized by common properties, then this in itself would provide sufficient legitimation to consider paranormal experiences a proper subject for scientific research. Then, depending on the nature of the characteristics observed, hypotheses and models either of a parapsychological or psychological nature can be tested which eventually should lead to the acceptance of one of these as the most likely explanation. The proof of existence of ESP and a repeatable demonstration, if such is possible at all, might well be the end-product of our research instead of a condition which should be fulfilled before we are allowed to carry out research. Repeatable phenomena in other sciences, as for instance physics, are often only repeatable because previous research yielded the knowledge about the proper conditions under which to elicit the phenomena.

From quantitative analyses of different collections of SPE's it was concluded that these experiences can not be considered as a collection of random coincidences but that they are characterized by various properties. It could also be demonstrated that most of these properties can not be explained by familiar sources of response bias. For instance, it has often been assumed that compared to males more female percipients in SPE's are observed because females would be more inclined to report paranormal experiences. No evidence was found to support that assumption. Another example is that it has been assumed that the fact that more SPE's take place between near relatives, can be explained because of the greater likelihood that an SPE will be discovered between near relatives who may have a greater degree of

communication than do other types of relationships. This assumption was also found to be incorrect.

We next investigated possibilities that SPE´s are based on a psychological attribution process. This hypothesis assumes that under certain conditions people will attribute a paranormal character to a coincidental agreement between a subjective experience, for instance a dream, and an external event. Hence, the observed properties are a consequence of these conditions, and not of the experiences themselves. The ´cultural´ hypothesis is an example of such an attribution hypothesis. It assumes that in a certain culture people have a common stereotype about what a SPE will look like and, that consequently, a paranormal character is attributed to coincidences which fit that stereotype. A consequence of this hypothesis is that cultures with different notions or stereotypes about SPE´s should yield collections of SPE´s with different properties. Collections of SPE´s from different cultures and epochs were compared to study whether these collections showed differences in characteristics which could be interpreted as supportive of this hypothesis. It appeared, however, that the characteristics of these collections were surprisingly similar and those differences which were observed could be readily accounted for by other factors. Hence the data do not support well the validity of the cultural hypothesis.

From the nature of the characteristics of SPE´s another attribution hypothesis was derived: the ´worry´ hypothesis. On this hypothesis it is assumed that people often worry, especially about persons close to them such as near relatives. However, the likelihood that serious negative events do happen to them, for instance that a child on the way to school dies in a traffic accident, is low. It might be that when such an accident does happen people, based on such always present vague worries, get the feeling "I knew that something dreadful would happen" and consequently attribute a paranormal character to these feelings which would otherwise have been considered to be simply normal worries. Two studies were carried out in which various aspects of people´s tendency to worry were investigated. Specifically, data were obtained on what people do worry about, about whom they worry and how strongly they worry about these topics and persons. These distributions were then compared to the same type of data obtained in the SPE studies on the assumption, that under the ´worry´ hypothesis the distributions for the worry and SPE studies should be more or less similar. Thus to stick to the example given above: If people often worry about their children having accidents then one can expect a high

frequency of ESP cases in which a parent is the percipient and a child involved in an accident the target person. It turned out that the data of these studies did not yield evidence supportive of this hypothesis.

A related area is the question of the relationship between SPE´s and belief in ESP and, in general, the question about what the reasons are for people to believe or not to believe in ESP. Various studies were carried out in this area but so far with little success. That is to say, no variable has been found which appears to have a strong relationship with degree of belief in ESP. The strongest correlations we observed were between r=.40 and r=.50 and concern the relationship between belief in ESP and religious attitude. In addition in two studies evidence was found which indicate that when parapsychological topics or paranormal experiences were discussed in the environment in which people grew up, this seems to have an effect both on reporting SPE´s and on belief in ESP.

If the data obtained in all studies on SPE´s are interpreted in terms of a psi model then these data suggest that psi should be considered an all-or-none process in which one essential item of information is obtained rather than a process in which detailed ´pictures´ are received from which the item of information is derived. It looks like as if the percipient ´receives´ one elementary piece of information, for instance X was killed in an accident, and that details based on existing knowledge about the person and the situation are added during the process by which the information becomes conscious. As regards psychological explanations we can conclude that the various hypotheses investigated so far are not well supported by the data. However, that does not imply that the possibilities in this direction are exhausted. Some, still unreported, studies have been undertaken in which it was attempted to discriminate directly between an attribution model and a psi model. The results, however, were unclear. Some findings appeared in favour of an attribution model, other in favour of a paranormal explanation. One possible explanation for this confusion might well be that our data base, the cases of SPE´s, are made up of both genuine ESP experiences (spontaneous paranormal experiences based on an ESP process, whatever that may mean) and quasi spontaneous paranormal experiences (QESPE´s: coincidences to which the experient attributes for psychological reasons a paranormal character). Further research might yield indicators which enables us to discriminate (roughly) between the two types of cases.

There is another approach to the basic question: 'Why do people attribute a paranormal character to certain experiences?' Attributing a paranormal character implies that the person attributes a special meaning to the experience involved. Hence one could also formulate the problem: Is meaning attributed to these experiences because certain psychological conditions exist (a 'psychological' explanation) or is it something in the process or experience itself which generates the attribution of significance (a 'parapsychological' explanation)? Therefore studying the more general problem of the attribution of meaning to experiences and events might also yield a fruitful approach to the problem of explaining spontaneous paranormal experiences.

Another long-term project involved the study of gifted subjects. 'Physical' mediums are hardly found nowadays but there are still quite a number of psychics or paragnosts active. Paragnosts are people who are convinced that they are able to obtain paranormal impressions more or less on 'command' and who use this gift to advise clients who turn to them for help. Previous research with paragnosts, by W.H.C. Tenhaeff and H. Bender among others, have mainly involved the more famous paragnosts such as Croiset, and focussed in general on the more successful and spectacular cases. Such cases have always drawn extensive public interest. Perhaps therefore many people believe that everyone who calls himself or herself a paragnost must be equally gifted and so, in The Netherlands, paragnosts are still often consulted for a wide range of problems.

A systematic study with paragnosts, involving different conditions and over 200 sittings, was carried out with three aims: 1) to describe what happens when a client consults a paragnost; 2) to evaluate how valid the paragnosts' statements are; and 3) to investigate in a systematic way the effects of various conditions on the paragnosts' statements. In this study all sittings took place at the homes of the paragnosts, never in the laboratory, and apart from restrictions imposed by the conditions, the paragnosts were left entirely free to behave as they are used to when interacting with clients. Contrary to most previous studies with psychics, the evaluation was based on all statements made by the paragnosts and not only on the more interesting ones.

In order to be able to evaluate the huge amount of statements obtained in this study, over 10,000, a new method to evaluate verbal material was developed. In general in traditional methods for the evaluation of verbal material all statements are judged for

correctness or incorrectness. By asking different possible target persons to judge all statements on applicability, estimates of probability of each statement for applicability can be obtained. Then based on these probabilities a p-value can be obtained for the set of correct and incorrect statements as judged by the target person for which the statements were intended.

Major drawbacks of these methods are that they require all target persons to judge all statements and that they do not allow for feedback by the client to the paragnost after each statement. In our method statements are first screened by independent judges on degree of probability and only those statements which have at least a minimum amount of what we call 'potential paranormal value' are retained for the final analysis. Hence statements as for instance 'She is sometimes nervous' are probably rejected but a statement like 'He recently traveled to Iceland' might be retained. By this method, which we have called the window model, the number of statements to be judged by the target persons become strongly reduced. Only the statements with sufficient potential paranormal value are retained for the analyses and have to be checked for correctness later on. Another major advantage of this method is that it also allows an evaluation of statements acquired under feedback conditions. This is so important because when clients consult a paragnost they nearly always are eager to provide feedback, hence the feedback condition is the condition under which paragnosts normally work.

The study was carried out with 15 paragnosts, among them the more famous ones of Holland, and two control groups of non-psychics. Conditions under which statements were acquired were systematically varied and involved among others different degrees of feedback, of seriousness of the events about which the paragnost was consulted, the nature of the object presented to the paragnost, and the social context in which the sitting took place. In order to describe what normally takes place in a sitting between a paragnost and a client, analyses of statements made in sessions were carried out involving among others length of the sessions in number of statements, nature of the topics discussed, statements made about past, present and future situations, and reactions of paragnosts to negative feedback, that is, feedback in which the sitter denies the correctness of a statement made by the paragnost.

Other analyses concerned the (for many most important) question, how are valid the statements made by paragnosts. We consider it impossible

to obtain true p-values for the correctness of most verbal statements. Hence the usual way to establish a possible paranormal character that of demonstrating that statements made by paragnosts are significantly more often correct than expected by chance, seem to us not applicable to these types of studies. Consequently we rephrased the question and asked: Do paragnosts more often make correct statements of potential paranormal value than do matched groups of non-psychics? (Matched either as regards gender and age or as regards experience in dealing with problems of other people). If that is the case, we can not only infer that paragnosts must have some special ability to do so but also, that it is meaningful to make use of this ability.

From the many interesting results of this study only a few can be discussed here. It turned out that the number of statements (roughly comparable to the number of paranormal impressions) varies strongly with the amount of externally provided information, e.g. the nature of the object presented and the amount of feedback. The less external information the lower the number of statements. Paragnosts turned out to be very stable over the 5 years that the data was gathered in that each paragnost was characterized by a personal style which hardly changed over this period. The most common reaction after negative feedback is that the paragnost provides a new interpretation for the statement which is in better agreement with the reaction of the client.

According to the independent judges, on average only about 10% of the statements made by the paragnost should be rated as sufficiently specific considering their nature and in view of the feedback given previously in the session. Of these statements on average again only about 10% turned out to be correct as well. Hence of all statements made about target persons, only approximately 1% is sufficiently specific and correct as well, a figure which does not suggest strong ESP abilities. More important, this figure does not differ from the results obtained with matched groups of non-psychics. Hence we concluded that there is little evidence that paragnosts can have psi impressions at will. That implies that we can not speak about a psi ability or gift. That does not imply that paragnosts can not have occasionally spontaneous paranormal impressions, as all people might have, but we do not believe that paragnosts have an ability different from other people in that they are able to provoke such experiences.

What often seems to take place between a paragnost and a client is a kind of quasi communication. The paragnost makes statements probably

mainly based on life experience and especially on experience in dealing with similar problem cases, whereas the client tries to give these statements a meaning within the context of her or his own life. So although the two processes start out as largely independent of each other, in a sense the two parties implicitly cooperate to find as much agreement as is possible. The feedback serves as a connecting variable and leads, in the course of the session, to more agreement. Therefore at the end of the sitting, not surprisingly, both persons are sometimes genuinely impressed by the results. In addition, paragnosts do built up considerable experience in certain problem areas, for instance locating missing persons, which, perhaps occasionally combined with spontaneous paranormal impressions, can lead to unexpected results. So paragnosts and clients might have convincing reasons to believe that something unusual is going on. Therefore paragnosts should certainly not be considered as all frauds who knowingly deceive clients. Also in view of their experience and knowledge in certain problem areas, one might well profit from consulting a paragnost. But it is wise to attach no more meaning to their opinions than one would to the advice of some other person with a special expertise.

A last major project briefly discussed here concerns the field of paranormal healing. In paranormal healing, a form of alternative healing, the healers believe that the effect of their treatment is due to a paranormal influence of the healer on the patient. Outwardly the healing method is characterized by the healer making 'passes', moving the hands alongside parts of the body of the patient without touching it. Another form of paranormal healing is healing-at-a-distance in which the healer concentrates at a previously arranged time on a photograph of the patient. There are quite a number of paranormal healers active in The Netherlands. About 600 of them are member of one of the several professional organizations of paranormal healers. In addition an unknown number of healers are active who are not connected with any of these organizations.

In the case of paranormal healing, the most important questions seem to be: Does it work and if so, why? As regards the question of effectiveness of the method two types of criteria should be distinguished: subjective criteria (Does the patient feel subjectively improved) and objective criteria (Can improvement be demonstrated according to objective medical measurements, for instance improved peak flow meter values in the case of asthma or lower blood pressure values for hypertension patients). In the case of objective criteria a

difficult problem constitutes the choice of base-line values. If improvement in patients is observed, to what should the degree of improvement be compared? Comparisons with scores of no-treatment groups is seldom possible, one can't ask ill people not to seek treatment because we would like them to serve as controls for our studies. To assume that without treatment the scores would have remain approximately unchanged is also debatable, patients unfortunately often deteriorate but spontaneous remissions do also happen. Perhaps the most practical solution, and one that serves directly the interest of society, is to compare the objective effects of paranormal healing with the effects of other types of treatment, as for instance conventional medication. We believe that in this area a step by step approach is the most sensible.

As regards the question of why the method works, assuming that it works at all, one has to take into account that each method of healing involves a complex system of variables having an effect on the patient. Hence it seems difficult to construct a condition in which the effect of only the method itself, in our case the laying-on-of-hands, can be studied and the influence of other variables like placebo-effects, expectations of the patients, etc., is totally eliminated.

Following a step by step approach it seems sensible first to study whether patients feel subjectively improved, and if so, whether this subjective improvement is accompanied by objectively measured improvements, before studies are carried out directed at the question of how paranormal healing might work. Hence a first investigation was aimed at the question: Do patients benefit subjectively from the treatment by healers and if so, how lasting is the effect and do the effects differ over types of complaints? In three studies questionnaire and interview data were collected from over 4000 patients of 65 paranormal healers. Based on the analyses of these data a number of interesting findings were observed. It appears that about 2/3 of the patients are females. All age groups are represented with an average age for the patients of 47 years. The average duration the patient suffered from the complaints before turning to a healer is approximately 7 years. That figure implies that most patients have a long history of probably less successful treatment in the accepted medical circuit before they try a paranormal healer. This fact should be kept in mind when considering the reported subjective success rates. Patients come to see healers for all sorts of ailments, excepting ones which are so serious that the patient is unable to

visit a healer, such as those ailments which require hospitalization. About 50% of the patients come to see a healer because they suffer from pain. As regards types of ailments 70% concern problems related to body functioning (e.g. chronical headache, backpains); the remaining are of a psychological nature (e.g. depressions, phobias).

The number of treatment sessions given by healers is on average 18 but varies widely. Some patients even have a lifelong treatment relationship with a healer. About 85% of the patients report that they feel subjectively improved by the treatment (which varies from a little improvement to strong improvement) whereas only a few percent report further detoriation. The degree of (subjective) improvement appears to remain stable over a period of half a year after the treatment was terminated. Most interestingly is the finding that no strong differences in degree of subjective improvement was observed for the different types of ailments. Hence our expectation that stronger improvement would be observed for complaints of a possible psycho-somatic nature was not confirmed.

Because many patients appear to experience subjective improvement due to the treatment, the next question is whether this improvement is also found when applying objective measures. However, answering this question requires systematic experimentation. Because objective measures are related to specific types of ailments, such experiments can only involve patients suffering from that ailment. In other words, answering this question requires basically as many experiments as their are ailments for which patients are treated by healers, a clearly impossible task. However, we were able to carry out two of such studies, one involving patients suffering from asthma, the other involving patients suffering from hypertension.

The two studies had basically the same design. In the first condition, the ´optimal´ treatment group, patients received the treatment which is typical for paranormal healers: the laying-on-of-hands. All forms of communication between healer and patient were permitted in this condition. In fact, apart from the sessions being held in the laboratory, healers treated their patients just as they would in their regular practice. Two other conditions were introduced to gain insight into the possible influence of psychological and paranormal factors on healing, in case an objective effect of the treatment in the optimal condition was observed. In the second condition, the treatment ´at-a-distance´ condition, patients were treated from behind a one-way screen which ensured complete

absense of sensory communication between healers and patients. The third condition, called the 'control' condition, was in all respects similar to the treatment 'at-a-distance' condition, except that no healers were present and hence that no treatment was given. In the latter two conditions patients did not know whether or not they were treated.

In the asthma study 90 patients participated, 30 in each condition, for a duration of two months. Eight treatment sessions were held. Patients were selected from a large number of volunteers in such a way, that three comparable groups of subjects for the three conditions could be formed. Various objective measures for asthma were taken after treatment sessions and daily during the duration of the experiment. The asthma study was of a double-blind nature.

From the results it appeared that according to several objective measures patients of the optimal condition improved during the experiment. The effect, however, cannot be solely attributed to the treatment itself. It turned out that as far as the dependent variables are concerned no significant differences could be observed in degree of improvement between the three conditions, although there was a trend that for some measures greater improvements were found for the optimal condition. There is, however, a difference in subjective evaluation. In the optimal group patients felt subjectively much more improved than the patients did in the other two conditions. Hence it was concluded that according to objective and subjective measures the treatment had a positive effect on the patients, but that these effects are due to psychological factors rather than to the treatment itself. A possible psychological factor contributing to the results could have been that by participating in this study patients created a new opportunity to improve their condition. In this experiment no indication was found for the operation of a paranormal factor in healing.

The hypertension study followed the same design as the asthma study but was carried out with 120 patients, 40 in each of the three conditions. In the optimal and healing-at-a-distance conditions patients received once a week treatment for a duration of 15 weeks. Patients were selected with blood pressure values ranging from still healthy to serious high levels to avoid a 'regression to the mean' effect. In this study a triple-blind procedure was followed. Neither the patients (except for the optimal condition) nor the experimenters knew which subjects were assigned to what conditions, and in addition

the statistical analyses on the three data sets were independently carried out by statisticians not knowing the relationship between data sets and conditions.

The results of this study appeared comparable to the results of the asthma study but perhaps due to the increased number of treatments (15 compared to 8 in the previous study) some effects became more pronounced. All three groups of patients improved strongly. However, in this study a significant difference was observed between degrees of improvement in the three conditions in favour of the healing-at-a-distance condition. This is the only indication we found for a possible paranormal influence of the healers on the patients. However, the results should only be viewed as suggestive as a direct comparison of degree of improvement between the healing-at-a-distance condition and the control condition did not yield a significant difference. In this study again strong evidence was found for the effect of psychological variables on the condition of the patients. As in the previous study patients in the optimal condition felt subjectively much more improved than the patients in the other two conditions who had had no personal contact with healers. It was also of interest to note that in all conditions the degree of improvement was equal for patients using medication as for the patients not using medication.

From these studies we concluded that according to both subjective and objective criteria paranormal healing must be considered effective. However, the effects are not very strong. According to the medical specialists who participated in the hypertension study the (objectively measured) strength of the effect of the treatment was comparable to the effect of mild medication, but with no undesirable side-effects. The effect of the treatment subjectively experienced in terms of improved feelings of well-being appear stronger than the objectively measured effects. From the point of view of the patients this finding is important because most patients come to see healers because they do not feel well or suffer from pain. The effects of the the treatment seem mainly due to psychological factors. No indications have been found that under the conditions of our research, in which medical supervision was permanently available, paranormal healing might have had a negative effect on patients.

As stated above, these studies are only a few of the many carried out in the Laboratory.

As a consequence of the closing of the Parapsychology Laboratory, the prospects for further research in parapsychology at the University of Utrecht are not promising. The demise of the Laboratory is more unfortunate because it was the only regular university department for parapsychology in The Netherlands. In view of the steady or perhaps even increasing popularity of parapsychological topics in the media, often presented in a corrupted and sensationalized form aptly dubbed by Martin Johnson ´paraporno´, it is important that at least one independent institution should remain to which society can turn for objective and scientifically valid information. Such information can not always be expected from organizations in society not connected to the university, who, in general, are inclined to take positions and to advocate either positive or skeptical beliefs about the paranormal. It is my conviction that parapsychology, as the objective scientific study of paranormal experiences and phenomena, is a legitimate branch of science and should have its place within academic and other scientific institutions.

NOTE

The various publications on research with spontaneous paranormal experiences can be found in issues the European Journal of Parapsychology of 1979, 1981, 1982, and 1984, in the Journal of Parapsychology of 1983, in the Zeitschrift fuer Parapsychologie of 1983, and in the Proceedings of the 1984 International Conference of the Parapsychology Foundation. Of the research with paragnosts publications appeared in the Journal of Parapsychology of 1983, and in various issues of the European Journal of Parapsychology. All these paragnost studies are presented together in the Ph.D. thesis of H.G. Boerenkamp published in 1988 and called ´A Study of Paranormal Impressions of Psychics´. The statistical data on paranormal healing and the experimental study with asthma patients are published by Attevelt in his Ph.D. thesis of 1988 called ´Research into Paranormal Healing´. These studies were supported financially by the N.F.P.N., the largest professional Dutch organization of paranormal healers, and by the Johan Borgman Foundation. The results of the experimental study on paranormal healing with hypertension patients, supported by the Ministry of Health, appeared among others in a report from 1987 by Attevelt et. al. (in Dutch) to the Ministry of Health called ´Effectiviteit Paranormale Geneeswijze´, and in publications by

Beutler et. al. in the Journal of Hypertension of 1987 and in the British Medical Journal of May 1988.

EUR.J.PARAPSYCHOLOGY., V.7, 1988-1989

A FREE-RESPONSE STUDY IN A REAL-LIFE SETTING. A REPLICATION

Sybo A. Schouten
Peter Haneveer
University of Utrecht

In general most people experience only a few or no spontaneous paranormal experiences in their life. Various studies by Boerenkamp (1988) indicate that people, who considers themselves gifted as regards psi abilities and claim to be able to use these abilities at will, seem not able to substantiate such claims. Hence if we assume an ESP process to be possible it appears that in daily life natural conditions tend to inhibit the occurrence of experiences based on ESP. That implies that the probability of finding evidence of ESP under experimental conditions must be low and that in many studies no evidence of psi will be observed. Most investigators in this field will know by experience this to be a fact.

As long as we do not know which conditions favor psi experiences or what the nature of psi is, it is difficult to develop an optimal research method for studying psi phenomena. In a previous study (Schouten and Merkestein, 1985) a research method was explored which was based on the findings from analyses of spontaneous cases that topics which are of importance to the

---

Note: This paper fulfils the publication policy of this journal

percipient seem to have a higher likelihood of becoming the subject of a psi experience. The research method applied involved a free-response study in which subjects provided their own target sets consisting of issues important to them at that time. One of these issues was randomly selected and acted as target for the day of the trial. The mentation of the subject was not obtained in the laboratory but based on the occurrence of unusual and striking inner experiences during the day of the trial. Hence the method tries to imitate as much as possible the situation in which spontaneous paranormal experiences might occur. The main exception was that for obvious reasons the target person could not be a relative or friend of the subject, as is nearly always the case in spontaneous paranormal experiences. Instead the target person was a stranger recruited by one of the experimenters. Given the data we have on spontaneous experiences the latter constitutes a weak point of the method and might be the reason why the study mentioned above did not yield positive results.

Another aim of the study was to improve judging procedures. The judging procedure is an important variable in free response studies. Judging procedures can be either of a global or atomistic approach. The global approach has the advantage that in the case the subjects do the judging, the assessment can be based on all aspects of the mentation including aspects which were not mentioned to the experimenter. With the atomistic approach the separate items of the recorded mentation or inner experiences are compared to each picture and the total score for each picture is based on the sum of the scores of the separate items for that picture. The advantage of this method is that it appears more objective and reliable and less influenceable by judging criteria irrelevant to the task (such as, for instance, differences in 'attractiveness' of targets).

Which approach is to be preferred depends, among other things, on the nature of the assumed psi process. If the conclusion from analyses of spontaneous cases (Schouten, 1979a, 1979b, 1981, 1982) is correct that in ESP experiences in general only one 'main' message is 'transmitted', then the psi process can be assumed to be essentially of an all-or-none nature. In that case the global approach seems preferable. However, if the Tyrrell model holds and paranormal experiences involve the transmission of detailed information of which most often only a part reaches consciousness, then the atomistic approach might be better suited. Since the targets in the study described above consisted mainly of concepts, for instance music, a global judging procedure was adopted.

However, it is likely that with the global approach subjects employ different judging strategies, often unknown to the experimenter, and that such judging strategies might also vary for different conditions of the experiment. If that is true then the interpretation of differences in ESP results between conditions in terms of an effect of the independent variable might be disputable. Unfortunately little is known about how subjects match mentation and targets and how specific conditions might effect the judging. Therefore the previous study explored ways to improve the objectivity of the judging procedure, by reducing the set of mentation items to a smaller set of items but with a higher likelihood of obtaining hits.

One attempt in the previous study proved successful. When an independent judge based his judging on mentation involving only items of memory he obtained 7 hits in the 10 trials in which mentation involved personal memories. Hence in this case the independent judge achieved significant scoring whereas judging by the subjects based on all mentation items did not.

This finding does not necessarily imply that a relationship exists between memory and ESP. The analysis was one of several made to explore the data and hence the finding has at best a suggestive value. Secondly, the judging procedure was of improved objectivity only in so far that it was based on a specific type of mentation. This limits strongly the number of items of mentation which might be used for the matching and therefore reduces the problem associated with global judging, namely, how to select between all the possible connections between many items of mentation and the different aspects of the possible targets. In fact, in nearly all trials of the previous study the set of mentation items became reduced to one item related to personal memories, thereby greatly reducing the task of the external judge. However, the matching itself was still global. The external judge applied his own subjective criteria for the matching. Therefore it would be useful to know how reliable the matching of the judge has been. Another possibility already mentioned in the report on that study is, that the relationship between memories and ESP scoring only holds when the targets are topics taken from the subject's life and of high relevance to the subject.

The present study primarily represents an attempt to replicate the finding of a possible relationship between memory and ESP. Hence we tried to replicate the Schouten and Merkestein study as carefully as possible. The main characteristics of that study were: a free response design, one trial a day, no special session

but unusual experiences during the normal activities of the day were used for matching, and each subject had a different target set which consisted of important topics in the life of the subject.

It was hypothesized that the scoring would be higher when the judging was carried out by an external judge who followed a formal judging procedure and used only inner experiences reflecting memories, than when the judging was done by the subjects themselves using a global approach, matching all their recorded and unrecorded inner experiences.

A second aim of the study was to investigate whether the relationship between memory and ESP, if observed in the data of the present study, depended upon whether the targets were topics taken from the subject´s life and of importance to the subject. To this end a second condition was introduced in which the targets were of a similar nature of the targets employed in the first condition but consisted of topics of importance to some other person, not the percipient. Targets for the second condition were those used in the Schouten and Merkestein study, because these targets were very important to the subjects in that study. It was checked whether these targets, which were supposed to be neutral to the subjects of the present study, were really neutral to them. This way we ensured that the targets of the two conditions were comparable as regards their potential importance to people.

It was decided to test the differences between the two conditions based on the ratings assigned by a external judge using only mentations from the category ´personal memories´.

A third aim of the study was to further improve the objectivity of the judging procedure by adopting a more detailed procedure for the establishment of the degree of correspondence between mentation items and potential targets. The approach consists in establishing the correspondence between an experience and a potential target on three aspects: literal content, associative content, and agreement of details. Rules were formulated about how ratings had to be assigned to reflect the degree of correspondence for these three aspects. In order to obtain an estimate of the reliability of this judging procedure additional judges were employed who carried out the same matchings according to the same rules.

In addition to the experimental data belief in ESP and various personality variables of the subjects were measured for

explorative purposes.

THE JUDGING PROCEDURE

The judging procedure applied by the external judges involves two aspects.
1) The experiences or mentation of the subjects are classified in distinct categories by the subjects.
2) The content of the category 'personal memories' has to be rated on its degree of correspondence with each item of the target pool. In addition one external judge rated the content of the other categories too on degree of correspondence with each item of the target pool.

The classification of experiences

Based on data obtained in two unreported non-ESP studies, carried out by Schouten using psychology students as subjects, the following categories for classification of mentation have been applied.
1) Memories. Mentation related to inner experiences or events which have taken place in the life of the subject before the day of the trial.
2) Mentation related to events or activities on the day of the session or related to the future.
3) Emotions and feelings.
4) Daydreams and fantasies.
5) Spontaneous, unexpected and unclassifiable mentation.

When subjects are asked to classify each of their own mentation items in one of these categories this system appeared to function well in the two previous studies with psychology students. It apparently allows relatively quick and easy classification by the subject of his or her different thoughts or images. On the form used in the present experiment the categories were indicated by questions. For instance, the category personal memories was phrased as the question: "Did the experience mainly involve a personal recollection?"

The rating procedure

For judging by external judges targets and mentation are judged on three aspects.
1) Content. The literal content of a target and of an experience as given by the subject.
2) Associations. These are the associations given by the subject for each of the items of the target pool and mentations prior to the running of the study. The use of associations is partly derived from Sondow's (Sondow, 1979) conclusion that "all vague associations to a target may outweigh a single striking correspondence to a non-target picture" (ibid, p.125).
3) Details. Aspects of the target or experience which are not covered by its main theme. For instance, if the main theme is 'house', then doors and windows are not counted as details, but a car drawn alongside the house is. Details were established by the experimenter responsible for handling the subjects.

The judging procedure for the external judges consisted of comparing each potential target for the trial with each mentation item on these three aspects, and to rate the degree of correspondence on each aspect on a scale ranging from 0 to 30. Then for each drawing or potential target picture the sum of the ratings (for each experience there were three ratings, one for each aspect) over all experiences was taken, and these sums (actually these sums divided by the number of experiences) decided the ranking order of the drawings. Not surprisingly the drawing with the highest sum was taken as the 'guess' for that trial.

Subjects also rated the agreement between targets and mentation for each trial but in the traditional global way. They were instructed to compare all significant mentation of the day of the trial with each target and to assign ratings on a scale from 0 to 30 for degree of agreement.

METHOD AND PROCEDURE

A trial lasted a day. During the day an agent concentrated occasionally on a drawing randomly selected from one of two sets of five drawings made by the subject for the two conditions. The subject was requested to pay attention during the normal activities of the day to inner experiences which might be of relevance to the task. At the end of the day an experimenter

interviewed the subject about the mentation items described by the subject, asked the subject for his or her associations related to these mentations, and established the subject's ratings for degree of agreement between the mentation and each of the possible targets.

The tasks of the two experimenters (P.H and S.A.S.) who participated in the study were strictly separated. P.H. acted as Es (experimenter subject) and was responsible for all dealings with the subjects. In addition Es acted as the independent judge. Es did not know which condition or target was involved in a trial. S.A.S. acted as Ea (experimenter agent) and was responsible for randomisation of targets and for supervising the agent. Ea knew the condition of a trial for a given date but not the target. Es did not know the identity of the agent and Ea did not know the identity of the subjects. Ea and the agent lived near Utrecht whereas Es and the subjects lived in or near Rotterdam, approximately 70 km's from Utrecht. During the experiment Es and Ea had very little contact with each other, and if so only about technical matters concerning the experiment or matters not related to the experiment were discussed.

Ea selected the two further independent judges in order to obtain reliability scores for the judging procedure. These external judges had previously carried out a Ganzfeld study themselves and so were familiar with this type of judging. Es was not informed about the identity of these two judges. When the two additional judges carried out their tasks neither they nor Ea knew how Es had assigned his ratings.

Each subject participated in two trials one for each condition. For the PT condition (personal targets) the subject made a drawing for each of five topics he or she rated as important in his or her personal life. For the NT condition (neutral targets) the subject made drawings of five topics randomly selected from the list of topics of personal importance given by the subjects in the previous Schouten and Merkestein study. Hence the neutral topics in the present study can be supposed to be matched in degree of seriousness or importance to people with the topics of the PT condition. Ea prepared for each subject a list of 10 neutral topics randomly drawn from the list of topics used in the previous experiment. Es used the first five topics of the prepared list of neutral topics for that subject but checked with each subject that each neutral topic realy was neutral to the subject. If not, or if it resembled one of the personal topics given by the subject, that topic was substituted for the next topic on the list of ten

neutral topics for that subject. Hence each subject made 10 drawings, 5 of personal topics and 5 of neutral topics.

Es made 4 copies of each set of 10 drawings and then mailed the originals to Ea in two separate envelopes, one for the PT drawings and one for the NT drawings. Es assigned each subject a number and on the envelopes he noted only the subject's number, not a name. On one of the two envelopes he also printed the dates of the trials agreed upon by subject and Es. Two sets of copies were used for the ratings by the subject on the two days of the trial. In the two sets randomly ordered PT and NT drawings followed each other successively but in reverse order. However, Es deviated from the random order of PT and NT drawings in the case two adjacent drawings resembled each other. A third set of copies was used by Es to obtain after the second trial a rankorder of all ten drawings by the subject based on degree of importance of the topics of the drawings to him or her. The fourth set of copies were used by the external judges for their judging. The external judges were not allowed to write on these drawings and used forms to list their ratings.

Ea prepared a list of subject numbers and order of conditions for each subject. For the first subject he randomly decided the first trial to be the PT condition. Consequently the second trial for this subject would be the NT condition. Then for each following subject number he reversed the relationship between first and second trial and condition. This way the PT and NT conditions became evenly distributed over first and second trials. As stated above Es was not informed about this procedure nor about its result until after all data were collected and the ratings assigned.

For each subject number, and first and second trial, a Bierman RNG board provided a random number between 1 and 5. Ea prepared the target envelopes. For each of the two trials the envelope with the original drawings for the corresponding condition was used. The set of drawings was taken out of the envelope by Ea with the faces down and based on the random number for that trial Ea selected the target, counting from the top down. Thus in the case the number was five, Ea took the fifth drawing from the top, i.e. the bottom drawing. This drawing was put into another envelope, still upside down to ensure that Ea did not know the content of the selected target. On the envelope he wrote the subject number, trial date and the letter P or N to indicate the condition.

The agent was a young female of 18 years of age, the same person

as employed in the previous study. The agent was not informed about the specific aims of the study or about the meaning of the letters P or N on the target envelopes. She was asked to concentrate occasionally on the target during the day of the trial. She was requested to keep a notebook in which she recorded date, letter, and a description of the content of the drawing. The agent kept all target envelopes and the notebook until all judges had completed their rating tasks.

Thirty subjects, 19 females and 11 males, age ranging from 18 to 64 years with a mean of 33, volunteered to participate. They responded to invitations put up at bulletin boards in the Rotterdam Public Library and in buildings of the University of Rotterdam. Subjects differed strongly in level of education. Sixteen of them stated that they had had one or more ESP experiences in their life. The group included a psychic and a paranormal healer, both females. None of the subjects were prior to the experiment acquainted with Es and none had prior experimental experience.

Before the trials were held Es had a meeting with the subject in the subject's home. In this meeting Es explained the nature and procedure of the experiment (but not its aims) and asked for the five most important topics in the life of the subject. Then the drawings were made, except for the first three subjects first for the personal topics and then for the neutral topics. Subject's indicated on each drawing with a word or with a short description the main topic of the drawing. After this subjects filled in the 'Belief in ESP' and personality questionaires. These questionnaires were not evaluated before the sessions were completed to avoid such knowledge influencing Es's attitude towards the subjects. Finally the dates for the two trials were settled. A minimum of three days and a maximum of three weeks was kept between the two dates. The night before a trial date Es called the subject to remind him or her of the trial scheduled for the next day. Subjects were not informed about the hypotheses tested nor that on the two trial days targets would be used chosen from different sets of drawings. They knew, however, that not the same target would be used for the two trials.

On the evening of the trial day, Es visited the subject and recorded the unusual or striking experiences of the day on tape and, in addition, in condensed form on paper. For each experience Es elicited associations which were also recorded. In addition the subject classified each experience in one or more categories of the classification system. Finally Es took a set of copies of the

drawings and spread them out in front of the subject, from left to right. The subjects then rated all drawings for degree of correspondence with their experiences of the day. These ratings were written down on the copies. Hence in each trial the subject was asked to rate all ten drawings to avoid Es and the subject learning the nature of the condition involved in that trial. For the final evaluation only the five ratings associated with the drawings of the proper condition were used. After the second trial was completed Es asked for associations of the subject related to the ten drawings. In addition, Es requested the subject to rank all drawings in order of importance to him or her. This provided a check on the validity of the difference between the two operationalizations of the independent variable: the personal and neutral targets. If that difference was still as it was when the subject provided the five drawings of topics important to him or her one can expect the subject to assign these drawings the first five positions.

After collecting the data from all subjects, Es transcribed the tapes. Next Es judged, for each trial, all experiences against the ten targets in each set on three aspects: literal content, associations, and details. Because of the amount of work involved the two external judges did similar ratings only for the experiences which the subject had listed in the category 'personal memories'.

For evaluations based on the ratings given by the subject the 5 drawings of the set, associated with the condition of the trial, were ranked according to the ratings assigned by the subject. In the case of the external judges a more complex procedure was applied for establishing ranking order. When Es acted as external judge, each experience was compared with each possible drawing on the three aspects mentioned above. For each experience the highest value of the resulting three ratings was taken. Then for each drawing these highest values were added and divided by the number of experiences. The resulting averages were used to establish the rankorder of the five drawings of the target set for that trial. If the drawing with the highest rank matched the target a hit was counted. Analyses have been based on the number of hits. The two additional external judges followed the same rating procedure but they used only experiences reflecting personal memories.

After all data were collected from subjects and external judges Ea retrieved the target envelopes and notebook from the agent and together the two experimenters established the targets for each trial.

Personality data were obtained by using two tests. One measured belief of ESP and consisted of two scales of 8 and 10 items respectively. These two scales were developed using a Guttman scaling procedure. Since both scales are meant to measure belief of ESP the correlation between the subject´s scores on the two tests can be used as a check on the reliability of the scales. The other test is the N.P.V., a Dutch personality test measuring traits like among others rigidity, dominance, and feelings of inadequacy.

RESULTS

Subjects reported in the 60 sessions in total 171 striking or noteworthy experiences they had on the day the trial was held, an average of about 3 experiences per trial. These experiences were categorized as follows: memories 57; related to events of the day 69; emotions or feelings 81; daydreams or fantasies 22; and spontaneous and unclassifiable 41. The sum is higher than 171 because a number of experiences were classified into more than one category. Females reported on average a slightly higher number of experiences but not to a significant degree. In the first trial subjects reported in total 96 experiences, in the second trial 75 experiences. The difference does not reach a statistical significant level (n=171, p=1/2, E=85.5, D=10.5 z=1.53, p=n.s. two-tailed).

It appeared that contrary to expectations, subjects had occasional trouble with respect to applying the category ´personal memory´. For instance, the experience "I was thinking today about the experiment" was classified as a memory, whereas an experience like "I had to think strongly about one of the drawings I made for the experiment" was not rated as a memory. In such instances Es reclassified the experience himself. Originally subjects had rated 49 experiences as memories, due to reclassification this number increased to 57. Es had the feeling that this problem could have been avoided by using a different phrasing for the question which asked subjects whether the experience could be classified as a memory.

Based on the ratings assigned by the subjects, 12 hits were obtained, a result in agreement with chance expectation (n=60, p=1/5, E=12, D=0). In 34 trials the subjects had reported one or more experiences related to personal memories. Based on these experiences Es assigned ratings to the drawings of the trials

involved following the rating procedures detailed above. He obtained 3 hits, a deviation of D=-4 (n=34, p=1/5, E=7, D=-4). The difference between the number of hits obtained by the subjects based on their ratings and by the external judge based on mentation involving memories only is not significant and opposite to the predicted direction. The first aim of the study is thus not reached.

Not surprisingly in view of the negative result of the main analysis an equally non-significant result was found when comparing Es scores obtained in the P and N conditions. In the P condition the 15 trials in which personal memories were recorded yielded 2 hits, in the N condition the 19 trials involving memory experiences yielded 1 hit.

The third aim of the study was to improve the objectivity of the judging procedure by adopting a more detailed procedure for assigning ratings. As stated above Es applying these rules obtained 3 hits in the 34 trials involving memory experiences. The two additional judges applying the same rating procedures obtained 5 and 6 hits respectively, a relatively large difference in view of the small number of trials involved. However, in all instances in which the two additional judges had obtained a hit and Es had not, Es had classified that drawing in the second position. The values of the rankcorrelations computed between all rating assigned by Es and assigned by the two additional judges are presented in Table 1.
The correlations range between .59 and .79 and are thus all highly significant. Hence it appears that the reliability of the scoring system is higher than the differences in number of hits obtained by Es (3) and the two other judges (respectively 5 and 6) suggest.

FURTHER ANALYSES

The non significant difference between the P and N conditions is confirmed when the evaluation is based on the ratings assigned by Es for all trials using all mentation items. In the 30 trials of the P condition he obtained 6 hits, in the N condition 4 hits. A similar non significant result is found when this comparison is based on the ratings given by the subjects. The subjects obtained 7 hits in the P condition and 5 hits in the N condition. From these data it also follows that Es using a more objective rating procedure obtained in total 10 hits in the 60 trials of the study whereas the subjects using a global and more subjective rating

TABLE 1
Rank correlations between ratings assigned by external judges

	R	t	df	p
PT Condition				
Es and judge 1	.78	10.76	73	<<.01
Es and judge 2	.79	10.95	73	<<.01
NT Condition				
Es and judge 1	.64	7.95	93	<<.01
Es and judge 2	.59	7.01	93	<<.01

procedure obtained 12 hits.

During the experiment the Es noticed that the experiences belonging to the category spontaneous and unclassifiable might be especially related to the targets of the trials. Hence an analysis was carried out based only on these experiences. This analysis involved 29 trials in which 7 hits were obtained. Thus no indications were found that special importance should be attached to these experiences.

No difference in scoring was observed between males and females. Although subjects with previous spontaneous psi experiences obtained higher scores, 9 hits in 32 trials as compared to 3 hits out of 28 trials for subjects without previous psi experiences, the difference does not reach a significant level (zd=1.36, p=n.s., two-tailed).

After the second trial was completed each subject was asked to rate all 10 drawings on degree of importance to them. Thus the relationship between degree of importance of topics to the subject and ratings assigned by the subjects to the drawings when carrying out the judging could be studied. To this end for each subject the set of drawings was split into the five most and the five less important drawings and into the drawings which were given the 5 highest and the 5 lowest ratings. Then for first and second day of

the trial separately two-by-two tables were formed, based on importance of drawing and ratings assigned to them. The data are presented in table 2.

TABLE 2
Relationship between importance of drawing to subjects and ratings assigned

	First trial Ratings		Second trial Ratings	
	5 highest	5 lowest	5 highest	5 lowest
Drawings				
5 most important	105	45	108	43
5 least important	58	92	63	86
	Chi-square=28.43 df=1 p<<.01		Chi-square=24.98 df=1 p<<.01	

The table show highly significant distributions indicating a strong positive relationship between importance of target to the subject and assigning high rating values. This relationship is probably caused by a tendency of the subjects to assign higher ratings to targets which are more meaningful to them. But this correlation could also be based on a third variable, related both to the assignment of ratings and to importance of targets. Most likely, it might be that subjects have more often spontaneous mentation related to topics important to them, and consequently observe more often correspondences between their mentation and such meaningful topics resulting in higher ratings. However, from inspection of the content of the subject´s mentation our impression is that the latter is not the case.

In order to study the relationship between personality variables and ESP scoring the 30 subjects were split up into a group of 10 subjects who had obtained one or two hits, and a group of 20

subjects who did not obtain a hit in any of the two trials. Then scores on the various personality variables and belief of ESP scales were compared for the two groups applying t tests. Of the 8 personality variables studied, only one yielded a marginally significant difference. The 'hitters' appeared to have an average score on this trait in contrast to the 'missers' who scored below average, but no meaning was attached to this finding due to the number of analyses made and because of the nature of the relationship. In addition no difference was observed between the two groups of subjects as regards their scores on belief of ESP.

The Spearman rank correlation between the two belief in ESP scales was .57. This is still a significant value but in view of the higher correlations obtained in previous studies it might indicate that the scales may be losing their stability. However, it should be noted that the two groups of subjects generally had rather high belief scores (not surprising in view of the fact that they all volunteered to participate in the study without compensation) and as a result the rank correlation tends to become lower.

DISCUSSION

From the results it is clear that this study yielded no indications of ESP. Therefore the conclusion must be that the experimental approach applied in this study and in the previous one seems not successful. This approach is characterized by using daily-life experiences and individual target sets for each subject, made up of topics important to the subject. Different reasons might be considered and discussed for this failure, from the non-existence of psi to the involvement of an agent not related to the subject, but such speculations seem rather meaningless. As long as our research is in an explorative phase only approaches which seem successful without further ado are worthwhile to pursue. The aim of this replication was mainly to confirm the post-hoc finding in the previous experiment that items related to personal memories yielded strong significant scoring. This finding could not be confirmed. Hence two successive attempts to apply this (non-laboratory) experimental approach have failed.

Another disappointing finding was that the reliability of applying a more objective judging procedure was not as high as was hoped for. In 34 trials using only mentation related to memories Es obtained half the number of hits as did one of the two

additional external judges; a not very satisfactory outcome. This is aggrevated by the finding that the ratings assigned by the subjects can not be considered very valid either. The fact that strong correlations are observed between importance of topic to the subject and rating values is perhaps indicative of how complex the judging procedure is. No doubt other non-ESP factors influence the judging too. For instance, potential targets who are rich in content will perhaps receive higher rating values than targets with a more simple content. Hence agreement between mentation and content of potential target might be only one of many factors which influence the value of the ratings assigned. Especially in research dealing with weak effects as is the case in our field such undesirable influences can easily wash out any real effects. One consequence of this observation is that it might be advisable to ask only experienced subjects to take part in free-response studies. It might well be that the consistent results, published by some authors in free-response studies, are more due to the fact that they often work with small groups of experienced subjects than to some mysterious experimenter effect.

ABSTRACT

In a previous study a new research approach was explored, based on daily life experiences. Subjects were asked to note unusual mental experiences during the day the trial lasted. The target sets were composed of topics important to the individual subject. In that study a strong correlation was observed between ESP scoring and mentation items reflecting personal memories. The present study was an attempt to confirm this finding. In addition we studied the possibility, that the observed relationship is only valid when targets are applied which are meaningful to the individual subjects. A third aim of the study was to improve the objectivity of the judging procedure in free-response studies.

None of the aims of the study were reached. No indication for ESP was observed. It turned out that the judging procedures applied by the external judges did not yield an agreement on the number of hits. From the data it appeared that the importance of the potential targets to the subjects had a strong influence on the rating values they assigned to these targets. This implies that, although the ratings should reflect the degree of agreement between their unusual experiences on the day of the trial and the potential targets, other factors possibly also influence the ratings. It is concluded that it might be advisable to use only

experienced subjects in free-response studies.

ACKNOWLEDGMENT

The authors are indebted to Dr. Deborah Delanoy for her critical appraisal of this manuscript.

LITERATURE

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

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

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

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

Boerenkamp, H.G. ´A study of paranormal impressions of psychics. Part IV. The second group of experimental series´, E.J.P., 1986, 6, 107-128. (a)

Boerenkamp, H.G. ´A study of paranormal impressions of psychics. Part V. The group of control series with non-psychics´, E.J.P., 1986, 6, 107-128. (b)

Boerenkamp, H.G. ´A study of paranormal impressions of psychics´. Ph.D. Thesis, University of Utrecht, 1988.

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

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

Burdick, D.S. and Kelly, E.F. ´Statistical methods in parapsychological research´. In Wolman B.B. (ed) ´Handbook of Parapsychology´, New York, Van Nostrand Reinhold, 1977.

Schouten, S.A. ´Analysing spontaneous cases´, Research Letter Parapsychology Laboratory University of Utrecht, 1979, 9, 55-63. (a)

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

Schouten, S.A. ´Analysis of spontaneous cases. A replication based on the Sannwald collection´, E.J.P., 1981, 4, 9-49.

Schouten, S.A. ´Analysis of spontaneous casse. A replication based on the Rhine collection´, E.J.P., 1982, 4, 113-159.

Schouten, S.A. and Merkestein, J. ´A free-response study in a real-life setting´. E.J.P., 1985, 6, 19-32.

Sondow, N. ´Effects of associations and feedback on psi in the Ganzfeld: Is there more than meets the judge´s eyes?. J.A.S.P.R., 1979, 73, 123-150.

EUR.J.PARAPSYCHOL., V.7, 1988-1989

AN EXAMINATION OF SUBJECT AND AGENT MENTATION IN THE GANZFELD

Deborah L. Delanoy
University of Edinburgh

The ganzfeld technique encourages the subject to experience abundant mental imagery/mentation. One problem which arises from this involves the difficulty of distinguishing relevant (target-related) from irrelevant (not target-related) imagery. This ganzfeld study was conducted in hope of shedding some light upon this problem by examining the subjects' mentations for components which may help distinguish potentially target-related responses from other mentation items. Another related question involves the role of the agent in ganzfeld work. Typically, an agent is involved in sending activities for a relatively lengthy period of time, but is provided with little instruction as to how to send the target most effectively. This study examined different categories of agent activity/mentation to see if any agent sending activities appeared to be more successful than others in relation to the subject making apparently target-related mentations.

One aspect of subject mentation which was studied involved trying to find factors in the mentations which could assist judges in identifying target/response correspondences during the judging procedure. This was investigated because responses made by subjects when describing their imagery in free-response ESP experimentation are

often characterised as presenting distorted or in some way altered information about the target (Warcollier, 1938 & 1948/1963; Sinclair, 1930/1962; Terry & Honorton, 1976; Palmer et al., 1977; and Sondow, 1979). These distortions or transformations can create difficulty when assessing possible correspondences between targets and responses during the judging procedure in that the judge may fail to recognize potential correspondences due to such transformations. But perhaps it would be possible to identify responses which were more likely than others to convey target-related information. If certain characteristics were found more commonly in apparently target-related responses than in responses which apparently had no correspondence to the target, then judges could be instructed to pay particular attention to responses having these characteristics. For example if responses categorised as being particularly bizarre in content were discovered to relate to the target significantly more often than were responses lacking this characterization, then judges could be instructed in future experimentation to examine any bizarre responses more carefully than other responses and to put greater judging weight on any apparent correspondence which was discovered.

Mentation characteristics have, of course, been examined in many previous studies. However these examinations have often been confined to subjects making overall judgements about characterisations of the whole of their mentation report, by means of a post-session questionnaire. The type of information gained from such methods is too global to allow for any specific conclusions to be drawn concerning the efficacy of any particular mentation characteristic in identifying target-related responses.

There are a few experiments which have asked their subjects to examine each individual mentation item or response for certain characteristics. Sargent, Bartlet, & Moss (1982) asked their subjects to check through their mentation reports and make note of any items which seemed particularly clear or strong, and of surprising or unusual items. The unusual items showed higher scoring than the non-unusual responses, particularly in naive subjects, but the result was not significant. With regard to clarity, naive subjects scored significantly better on unclear imagery, while experienced subjects scored better on clear imagery. Milton (1984), using a method of examining mentations similar to that of Sargent, Bartlet, & Moss, found a significant relationship between surprising imagery and psi-hitting. In a later study, Milton (1985) discovered that fleeting imagery related significantly to psi-missing. These studies have

started to examine mentation imagery in a way which may provide judges of free-response data with useful insights as to how to best perform their task.

This study examined fifteen different types of mentations (e.g., bizarre, fleeting, and auditory imagery) to determine if any response type related significantly to psi-hitting. If certain mentation types, relating significantly to successful ESP scoring, can be identified, judges could then be instructed to give greater emphasis to such responses during the judging procedure. This could provide a valuable aid to judges, by helping them to identify and differentiate psi- from non-psi-mediated responses.

Another area of inquiry which could provide help to subjects in their judging involves the importance of weak and strong correspondences. In the past, few experimenters have presented any details of their judging instructions. Two exceptions to this are Palmer et al. (1979) and Sargent (1980). Both of these studies' judging instructions stressed the need to differentiate between responses, according to the apparent strength of the observed correspondence. Whilst it may seem to be common sense to give greater weight to apparently better correspondences, the author is unaware of any work having been done which justifies this conclusion. Indeed, the fact that the ESP content of responses often appears to be transformed and/or distorted, potentially to an unrecognisable degree, could be seen as arguing against this perspective. If, due to transformations, the ESP content of a response is 'masked', potentially 'strong' correspondences may not be recognised as such. Thus, the procedure of automatically giving greater judging weight to apparently strong correspondences, may be counter-productive. In order to further examine this possibility, this study examined weak and strong correspondences, to see if one was more related to psi-hitting than the other.

The role of the agent was also examined in this study. To date, the data concerning the benefit of even having an agent in ESP studies has been conflicting (Palmer, 1978). Nonetheless, some findings have suggested that the agent's activity during the sending period may influence the subjects' receptivity of apparently target-related impressions. Warcollier (1938), Van de Castle (1970), and Delanoy (1982, 1986) have noted that subjects sometimes appeared to make more target-related responses when the agent was not actively concentrating on the target. Work by Harley & Good (1981), Sargent, Milton, Payne,

& Bennet (1982), and Milton (1986) has suggested that the presence of an agent imposes a rational structure on the subject's responses which is lacking in clairvoyant conditions. If particular aspects of agent activity ('sending modes') can be related to successful ESP reception, then the training of agents could become of paramount importance. To investigate such a possibility this study examined 25 different kinds of mental activities or mentations experienced by the agent during the sending period in relation to temporarily-matched subject mentations. The objective was to try to determine whether certain agent sending modes were related, significantly more frequently than others, to the subject making apparently target-related responses.

In order to examine whether certain types of agent sending activity might correspond to the subject producing target-related responses, the agent in this study made a mentation report of all her on-going thoughts, experiences, emotions, and sensations during the sending period. These mentations were then categorised according to 25 different agent activity characteristics. The agent's and subjects' mentation reports were then temporally matched. From this one could determine whether certain agent sending modes were related, significantly more often than others, to the mediation of target-related responses, on the part of the subject, than were others.

While there is little indication in the literature that one should expect to find exact temporally-linked correspondences between a subject's and the agent's mentations, it was noted in a previous study (Delanoy, 1986) (in which the agent was able to hear the mentations being made by the subject) that on occasion the subjects' mentations appeared to mirror the contemporaneous thoughts of the agent. As the design of this study allowed for a temporally-matched examination of subject/agent mentation correspondences, it was decided to conduct an informal analysis exploring the occurrence of time-linked correspondences. The method for accomplishing this comparison of mentations was to be relatively crude, and no firm conclusions were intended to be drawn from it. However, it was felt that it might provide suggestive results which could then benefit from future systematic research.

Various ganzfeld studies have reported findings which have been judged by both the subjects and by an independent judge(s). Several of these have found that the scoring based on the subjects' judging obtained a greater degree of psi-hitting, or a lesser degree of

psi-missing than did that of the independent judges (Sargent, Bartlet, & Moss, 1982; Sondow, 1979; Child and Levi, 1980; and Palmer et al., 1977). This could indicate that the subject has a greater source of information available to him than does the independent judge (e.g., subjective knowledge of his experience in the ganzfeld, and of his personal associations with his responses and the target pictures), and hence is better able to recognize response/target correspondences. However, one study (Palmer et al., 1979) obtained results which suggested that the independent judges were able to better detect mentation-target correspondences than were the subjects. The judges in the latter study were two research assistants who, one assumes, were familiar with parapsychological research, and may thus have had greater judging experience than some of the persons employed as independent judges in other studies. It may be that the level of success of judges depends upon their skill and experience in judging free-response data. However, as few studies provide details as to the judges' training in, or prior experience of, free-response judging, no conclusion can yet be made regarding this point.

This study was judged both by the subjects, who received training in how to conduct the judging, and by an independent judge, who had considerable experience in free-response judging. The scoring outcomes of the two were compared, but no prediction was made regarding whether the results based on the independent judge or the subjects' judging would disclose the greater degree of ESP-scoring.

METHOD

Participants

Twenty subjects, ten males and ten females, ranging in age from 19 to 48, took part in this experiment which was conducted in the spring and early summer of 1985. Ten of the subjects were studying for an undergraduate degree. Excepting one, the rest of the subjects already had at least one university degree. The majority of the subjects learned about the experiment from attending parapsychology seminars which are regularly presented by the Parapsychology Laboratory at the University of Edinburgh. Other subjects were friends of the experimenter, or friends of other subjects, who had expressed an interest in taking part in an ESP experiment. Only two of the subjects

had previous experience with parapsychological research, with both having prior experience of the ganzfeld.

Subjects were selected if they expressed an interest in participating in the study, had an open-minded attitude to the possible existence of psi, and if they were judged by the experimenter to be generally well-adjusted socially and to have no evident vested interest in either having or not having psi ability. Each subject participated in one training session, designed to familiarise them with the experimental procedure, and two ganzfeld sessions. Prior to recruitment, all potential subjects were told about the nature of the study, and the general ganzfeld procedure was explained to them.

The agent for all sessions was the experimenter, who has considerable experience with the ganzfeld, having acted as the agent and experimenter in over 100 previous ganzfeld sessions and as a subject in over 15 sessions, While she has yet to participate in a ganzfeld study which achieved an overall level of significant psi scoring, she does believe the ganzfeld to be a psi-conducive procedure and expected this study to obtain a significant, positive outcome. The experimenter's interactions with the subjects were friendly and informal, and frequently involved conversations about topics unrelated to the study. The experimenter tried to make the subjects feel relaxed and comfortable in the experimental environment.

The data from this experiment was independently judged by Julie Milton (Milton, 1986). Milton has considerable experience with the ganzfeld and has participated in many previous studies in a variety of roles, including those of experimenter, subject, and independent judge.

Setting

The experiment was conducted in three rooms in the Psychology Department of the University of Edinburgh. The department is centrally located, and is easily accessed with the area being served by numerous buses and a car park located near by. The building is placed opposite a garden on a square in the centre of the University complex, and is relatively unaffected by traffic noises. The department is housed in a large Georgian building which originally contained several separate flats, and more recently housed a girls' school. Given the age of the

building, it is well maintained and generally impressive in appearance.

The Parapsychology Laboratory, located on the top (second) floor of the department, was a large room with four very big windows and skylights, which gave it a bright, sunny atmosphere. The room was normally used as an office by the experimenter and as a meeting place for parapsychology seminars. It contained normal office equipment including about eight desks and tables, storage cabinets, and bookshelves. This room was used to greet the subjects before starting a session, and was later used by the subjects when they were performing the judging procedure. During the experimental session, the subjects would not normally meet any members of the parapsychology unit other than the experimenter.

A very small room, located off one end of the lab, served as the ganzfeld stimulus room. This room contained a reclining chair and a small table for experimental equipment, there being little room for any further furnishings. When the door connecting the stimulus room to the lab was closed, the small room was quite dark and cosy, described by one subject as being 'womb-like'. The only access to the stimulus room was through the lab.

The agent's sending room was an office located on a different corridor, on a different level of the building. Although the room was a relatively short distance from the lab, it was sensorily isolated from it. The sending room normally provided office space for two people. There was a screen installed dividing the room so that the area used for sending was not visible from the other desk in the room, and vice versa. The person using the other desk was a post-graduate in the department. He was not usually in the room during the sending period; however, on the few occasions when he was present, he would maintain silence. It should be noted that the other person was not involved in the study in any way, and, to the best of my knowledge, did not know any of the subjects taking part in the study.

Targets and Target Selection

The experimenter compiled twenty target sets, each containing four pictures, for use in the study. The pictures were composed primarily of postcard-sized art prints and photographs. They were chosen so that

the pictures in any set were as different from each other as possible in terms of content, theme, dominant shapes/forms, and colour. No pictures were chosen whose theme or content was judged by the experimenter to contain elements which could be construed as being violent, frightening, sexual and/or religious. Duplicate target sets were compiled for the subject and agent, with each picture being enclosed in a separate envelope for the agent's set, so that the designated target could be removed without the other pictures being viewed. The target sets were compiled several weeks before the start of the experimental sessions.

The target designation for each session was randomly determined by means of random number tables by a person not otherwise connected with the study using Rand Corporation random number tables (1955). According to a predetermined set of rules, two tosses of a ten-sided die established the entry point into the table. The target designation was then established by the consecutively listed numbers so that the numbers one and five designated target number one, two and six designated target number two, and so on, with the numbers zero and nine being excluded. The target set to be used in each session was chosen using Rand Corporation random number tables by a second person who was otherwise not involved in the experiment. Selection of target sets was made by a predetermined quasi-random method, similar to that described above, but which allowed for at least fifteen sessions to take place before any one target set was used for a second time. This was to prevent the agent from becoming overly familiar with the contents of any set, by seeing the same set twice in close succession. No subject was presented with the same target set more than once.

All target and target set designations were placed in individually sealed envelopes, labelled by trial number, by the individuals performing the randomization. Separate lists of the target and target set designations were kept by the randomizers to prevent any tampering with the original random ordering. All target sets, and the sealed target and target set designations (given to the experimenter by the randomizers), were kept in a locked box in the parapsychology lab to which the experimenter had the only key.

Apparatus

Four reel to reel tape recorders were used. One, located in the

ganzfeld stimulus room, relayed the white noise via headphones to the subject. Another, located in the room where the judging took place, recorded the subject's mentations. The third tape recorder was located in the agent's sending room and recorded the agent's mentations. The fourth recorder, located in the judging room, and connected to both the subject's and the agent's mentation recorders, played a recording of a talking clock. The talking clock reported the time at five second intervals, progressing from zero to forty minutes. The recorders used to record the subjects', and agent's, mentations were both two track. Whilst one track was recording mentations, the second track recorded the talking clock, as relayed by the fourth recorder. This enabled the subject's and agent's mentations to be matched to each other at any given point, within a maximum discrepancy of five seconds (for details of the procedure used to time-match the subject and agent mentations, see analysis 4 in the Results section).

A switch in the stimulus room, located within easy reach of the reclining chair, could sound a buzzer in the sending room. This buzzer was used to summon the experimenter if any difficulty arose during the session, and to signal the experimenter when the subject had completed the judging procedure. Unknown to the subject there was another buzzer, which would sound in the sending room if the drawer containing the subject's target set was opened. The purpose of this was twofold: first, it was a precaution against the subject removing himself from the ganzfeld prior to the end of the stimulus period to illicitly look at the target set. Secondly, it informed the experimenter of when the subject was starting the judging procedure.

During the ganzfeld stimulus, the subject lay in a comfortably padded reclining chair. Halved ping pong balls were affixed over their eyes by means of surgical tape, which could be removed painlessly. Any cracks between the subject's face and the halved hemispheres were filled by cotton wool. A flexipoise lamp, containing a 25 watt, 200-250 volt red bulb, was shone onto the subject's face from a distance of approximately 1 to 2 1/2 feet, depending upon the subject's preference. At the closer distance, a perceptible, but not intense, warmth could be felt coming from the bulb. White noise was conveyed to the subject, by means of headphones, for 37 minutes, the volume and tone being adjusted to the subject's preference. After the 37 minutes of white noise, there was a one minute pause in the recording during which time nothing was heard by the subject. Following the pause, a message played which informed the subject that the stimulus period was over and instructed them to remove themselves

from the ganzfeld and to proceed to the judging procedure in the lab.

PROCEDURE

Training Session

Prior to the actual testing sessions, each subject participated in a group training session. These sessions varied in size from three to ten subjects, and were all presented by the experimenter. The training sessions, which were held in the parapsychology lab, were informal in nature with refreshments being served. They generally lasted between 1 1/2 and 2 hours.

The primary purpose of the training session was to familiarise the subjects with the ganzfeld testing and judging procedure. While they received no direct experience of ganzfeld stimulation, the technique was described in detail. In a previous ganzfeld study (Delanoy, 1982 & 1986), subjects had encountered various problems involving both the receiving of impressions during the stimulus period, and the recognizing of various types of response/target correspondences during the judging procedure. These problems were discussed and examples of their occurrence were presented (see Appendix 1 for an outline of this discussion and the points covered). To assist with the judging procedure, three mentation records, from past sessions in which the experimenter had been the subject, were reviewed in detail along with their appropriate target set. The three mentation reports, all of which had received direct hits, were selected to highlight the different ways in which target/response correspondences may manifest themselves.

The subjects were also instructed how to rate their correspondences, using a mentation item-by-item rating scale of 0 to 5 points (where 0 represents no correspondence and 5 an excellent correspondence) for each picture in the target set. Thus, the subjects would rate the degree of correspondence between each item of mentation to each of the four pictures in the target set. The subjects were given a brief summary of the main points covered in the training session to review prior to their test sessions (see Appendix 1).

The training session was also used to familiarise the subjects with

the various imagery categories which they would be using to catalogue their mentations. Fifteen different categories of mentations were examined. These were grouped together in such a way as to make them easier for the subject to deal with. The categories, divided into their general groups, were:

Type of Image:
1. the image interrupted an ongoing chain of thought
2. the image was the result of one image transforming into another
3. the image developed into a recognisable one from an unclear one
4. appeared spontaneously

Duration:
5. the image was fleeting
6. the image was persistent
7. the image was recurrent

Clarity:
8. the image was undeveloped
9. the image was detailed
10. the image had intense colour

Content:
11. the image was bizarre
12. the image related to a personal memory or experience

Miscellaneous:
13. there was an auditory component
14. an impression of a sensation occurred
15. the subject experienced an actual physical reaction to an image

The subjects were instructed to categorize each item of mentation according to the presence or absence of the above fifteen characteristics. For each item of mentation, they were to tick as many, as few, or none of the fifteen categories, as they felt appropriate. The subjects were provided with a list of these mentation categories to review before their test sessions (see Appendix 2).

Ganzfeld Sessions

The design of this experiment differed from most ganzfeld

procedures, in that there was no subject experimenter. Instead, the subject's mentations were recorded, by means of a microphone suspended over the reclining chair, onto the subject's mentation tape recorder. If the need arose, the subject could contact the experimenter at any time by means of the buzzer located in the stimulus room, which when rung would sound in the room occupied by the experimenter during the stimulus and judging period. The buzzer was located on a table next to the reclining chair, within easy reach of the subject.

Upon arrival at the lab the subject was greeted by the experimenter and offered refreshments. The experimenter would then converse casually with the subject, during which time the procedure for the session would be reviewed. When the subject indicated that he was ready to begin the session, the subject and experimenter would then proceed to the ganzfeld room, where the subject was seated in a reclining chair. Halved ping pong balls, surrounded by cotton wool to eliminate any gaps, were affixed by means of surgical tape over the subject's eyes. The red light was shone onto the subject's face and the headphones relaying the white noise were placed over the subject's ears. When the subject was ready, the experimenter would start the white noise tape and leave the room, returning to the lab.

Immediately upon leaving the stimulus room (returning to the lab) the experimenter would turn on the subject's mentational tape recorder and the recorder which played the talking clock, both located on the table in the lab where the subject would perform the judging. Up to this point the agent was unaware of the identity of either the target set or the target picture for the session. She now unlocked the box, in the lab, in which the target materials were kept, retrieved and opened the envelope which contained the designation as to which target set was to be used for the session (the randomizer had noted which envelope was to be used for which trial on the outside of each of the sealed envelopes). She would then remove the designated subject and agent target sets (as specified in the target set designation envelope she had just opened) and the sealed envelope designating the target number for the session, and would then relock the box. The envelope containing the subject's target set would be placed in the drawer which was rigged with the alarm buzzer. The experimenter would then leave the lab, taking with her the agent's target pack, and the as-yet unopened envelope containing the target designation. The door of the lab was then locked, to prevent anyone from entering it whilst the subject was in the ganzfeld.

The agent then proceeded to the agent's sending room. There she would turn on the tape recorder which would be used to record her mentations. The volume would be adjusted to enable her to hear the talking clock which was recording on both her and the subject's mentation tapes. As the volume on each of the two tracks (one recording the talking clock and one recording the mentations) of both the agent's and subjects' tapes were controlled separately, the volume of the talking clock could be turned up or down, depending upon whether or not one wished to hear the talking clock concurrently with the mentation report. The volume of the talking clock on the subject's mentation tape recorder was turned off, so that he would not hear the clock when he was reviewing his mentations during the judging procedure.

The agent then busied herself with other things, until the talking clock announced 14 minutes. (The time announced by the talking clock corresponded to how long the subject had already been in the ganzfeld.) The envelope containing the target designation was then opened, and the envelope containing the appropriate target was removed from the target set. When the clock announced 15 minutes its sound was turned off, and the agent would remove the target from its envelope and commence sending.

The agent would send the target for 15 minutes. During this period, she would say out loud all of her thoughts, imagery, feelings, and activities. At the end of the sending period, she would replace the target in its envelope, and relax until the subject's stimulus period had ended.

The subject received ganzfeld stimulation for 37 minutes. At the end of the period there was a one minute pause in the white noise recording, during which time nothing was heard by the subject. Following the pause, a message played which informed the subject that the stimulus period was over. The message instructed the subject to remove himself from the ganzfeld and proceed to the judging procedure. On the judging table in the lab, he would find instructions to aid him during the judging (these instructions are reproduced in Appendix 1).

After helping himself to more refreshments, the subject removed the target set from the drawer and was instructed to study the pictures as taught in the training session. He would then rewind the tape upon which his mentations were recorded and replay his mentations, writing each item of mentation onto the judging sheet. He had been instructed

to stop the tape recorder after entering each mentation item and finish the judging for that particular item, before progressing to the next.

For each mentation item, the subject was to place a tick in any of the imagery categories which applied to it. He was free to tick as many or as few (or none) of the categories as he felt appropriate for each item. He was then to rate the item's correspondence to each of the four pictures in the target set, using a 0-5 point scale. After having judged all of his mentation report in the above manner, he totalled the points allocated to each target, and rank-ordered the target pictures accordingly. Most subjects took between one to two and a half hours to complete the judging. When finished, the subject summoned the experimenter by means of the buzzer in the stimulus room.

Once the subject had finished his ganzfeld stimulus period, the agent would then rewind the tape containing her own mentations, and proceed to write down each mentation item, and the time at which it was made (as conveyed by the talking clock), on to a judging form similar to that used by the subject. The agent would also catalogue each mentation item, in accordance with 25 different agent mentation categories (see Appendix 3). These 25 categories where chosen to represent the variety of sending activities/modes which the agent might use during the sending period.

Having completed the mentation categorisation, the agent remained in the sending room until summoned by the subject. On two occasions subjects contacted the agent by sounding the buzzer before she had completed the judging of her mentation reports. These two sessions were not included in any analyses involving the agent's mentations, due to possible contamination, as the agent would have been aware of the subject's target ranking, prior to having completed her judging. When the buzzer rang notifying the agent that the subject had completed his judging she would return to the lab, where she would give feedback as to the identity of the actual target, and generally discuss the session with the subject.

Independent Judging

The independent judge, Julie Milton, made transcripts of all the sessions from the subject's mentation tape, noting the time at which

each mentation item had been made, as indicated by the talking clock. She received no feedback as to the actual target for each session until she had completed the judging of all of the sessions. The independent judge rated each item of mentation to each of the five pictures using the same rating scale as the subjects.

Three experimental sessions were not judged by Milton. For two of these she was the subject, and the third was not judged as she received information regarding the target before the judging was completed. For greater detail regarding the independent judging procedure see Milton (1986).

METHOD OF ANALYSIS AND RESULTS (note 1)

All analyses were preplanned. As this was an exploratory study, no predictions were made and all of the given probabilities are two-tailed, unless otherwise stated.

1. a) A sum of ranks (Solfvin et al., 1978), based on the subjects' target rankings, was the primary overall psi measure and should represent the study outcome. The subjects' overall ESP scoring did not differ from chance expectancy as measured by a sum of ranks (MCE sum of ranks = 100, obtained sum of ranks = 102, $z = -0.21$, corrected for continuity).

b) A sum of ranks based on the independent judge's data also was non-significant (MCE sum of ranks = 92.5, obtained sum of ranks = 81, $z = 1.62$, $p = 0.053$, one-tailed). A comparison of the scoring obtained by the subjects' and the independent judge (Solfvin et al., 1978) showed no significant difference between the two ($z = 0.11$). The distribution of the subjects' and the independent judge's target ranks is shown in Table 1.

2. The results were also examined to determine whether weak or strong correspondences best identified the target.

Method of Analysis:

This was analysed by means of sum of ranks, the ranking of the

TABLE 1
Target rank distribution of the subjects and independent judge

Characteristics using a Wilcoxon Test

Subjects	Target Rank				Independent Judge				
Target Rank	1	2	3	4		1	2	3	4
Session 1	5	5	5	5	Session 1	6	7	2	3
Session 2	5	3	7	5	Session 2	7	3	6	3
Overall	10	8	12	10	Overall	13	10	8	6

pictures being determined by the rating points assigned by the subject to each picture for each item of mentation. Weak correspondences were considered to be those receiving ratings of two or less, and strong correspondences were those receiving ratings of three to five. The pictures were then twice re-ranked, one ranking based upon the weak correspondences and one upon the strong, and two sum of ranks analyses were conducted upon the re-ranked data.

Results:

Neither weak nor strong correspondences obtained significant results (MCE sum of ranks = 100; weak correspondences: sum of ranks = 105, $z = -0.64$; strong correspondences: sum of ranks = 98, $z = 0.21$); nor was the difference between weak and strong correspondences significant ($z = -0.30$). Thus, in this study neither weak nor strong correspondences conveyed a significant degree of target-related information, nor did one convey a significantly greater amount of target-related imagery than did the other.

3. The data were also analysed to determine if certain response types, as defined by the fifteen various characteristics, conveyed target-related information more frequently than other types of responses.

Method of Analysis (note 2):

The data were examined by the following method: the proportion of all points (the correspondence rating points given for each mentation item by the subjects) allocated to the target on the basis of each imagery category, were compared to the proportion allocated to the target on the basis of the remaining imagery for each trial, using the Wilcoxon matched-pairs signed-ranks test (Siegel, 1965).

Results:

Overall, 25.6 per cent of all item-by-item rating points were assigned to the target (MCE = 25 per cent). Fifteen different mentation types (or response characteristics) were examined, the results are given in Table 2. Of the
fifteen categories, one did not provide enough data to allow computation (this category was when the imagery prompted a physical reaction in the subject). The category of undeveloped imagery (when an image was unrecognisable, its features usually being described in terms of form/geometric shapes; and/or when the subject had a thought or idea which never developed into a mental picture) was the only category of the fourteen remaining categories which was shown to convey a significantly greater proportion of target-related information than the others ($N = 32$, $T = 158.5$, $z = 1.973$, $p = .046$).

4. The data were also examined to determine whether certain types of agent activity, as defined by the agent's 25 mentation categories, could be identified as particularly corresponding to the subject making target-related responses, using the same analysis as those used in analysis 3.

Method of Analysis (note 2):

To conduct this analysis, the agent's mentations were time-matched to those of the subject so that any response made by the subject was matched within a five second interval, to a mentation of the agent. If the subject made a response at a time where there was no corresponding mentation from the agent, the subject's response was categorised under the agent's mentation category of 'blankly' looking at the target (this category referred to those instances when the agent was thinking of nothing in particular, and just blankly looking at the target; during such periods the agent would not be making mentations). If the

TABLE 2
Results from the subjects' judging comparing mentation characteristics using a Wilcoxon Test

Mentation Type	N	T	p
1. Imagery interrupted on-going thoughts	23	134.5	n.s.
2. Image transformed from another image	30	180.5	n.s.
3. Image developed from unclear imagery	25	159	n.s.
4. Imagery appeared spontaneously	36	331	n.s.
5. Imagery was fleeting	37	351	n.s.
6. Imagery was persistent	34	287	n.s.
7. Imagery was recurrent	34	208.5	n.s.
8. Imagery was undeveloped, vague	32	158.5	.046
9. Imagery was detailed, clear	34	229.5	n.s.
10. Imagery was intensely coloured	27	185.5	n.s.
11. Imagery was bizarre	21	83.5	n.s.
12. Imagery related to a personal memory	33	277	n.s.
13. Auditory imagery was experienced	15	51	n.s.
14. Imagery suggested a physical sensation	20	74	n.s.
15. Imagery prompted a physical reaction	6	10	---*

* There was insufficient data to compute this analysis.

agent made a mentation at a time when there was no corresponding subject mentation, the agent's mentation item was discarded. Mentations were only matched to each other if both fell within the same five second interval, as reported by the talking clock. In cases where an item of either the subject's or agent's mentation lasted more than five seconds, it was matched to any/all of the mentations of the other which occurred during its duration. The correspondence rating points allotted by the subject to the target picture could thus be assigned to the appropriate time-matched agent mentation. The proportion of all rating points allocated to the target on the basis of each agent mentation category was then compared, by means of a Wilcoxon test, to the proportion of points allocated to the target on the basis of all the remaining imagery for each trial.

Results:

Overall, 26.3 per cent of item-by-item correspondence rating points were assigned to the target, based on the agent's mentation report (MCE = 25 per cent). Of the 25 agent mentation categories, six did not have enough data to allow analysis. Of the remaining eighteen categories, four were found which corresponded to the subject making target-related responses to a significantly greater degree than the other characteristics. The four categories were: a) when the agent was concentrating on actively sending to the subject ($n = 30$, $T = 118$, $p < 0.02$); b) when the agent was experiencing mental imagery ($n = 33$, $T = 118$, $p < 0.005$); c. when the colour of an object was particularly noticed by the agent ($n = 28$, $T = 115$, $p < 0.05$); and, d. when the agent experienced a vague and/or unclear mental image or thought ($n = 7$, $T = 0$, $p < 0.02$). The results of the analyses of all the agent categories are presented in Table 3.

5. A non-statistical analysis was conducted to examine whether the agent's and the subjects' mentation reports showed evidence of concurrent time-matched correspondences.

Method of Analysis:

The comparison between the agent's and subjects' mentations was informally carried out, by the experimenter simply scanning the reports to note any observed time-linked correspondences, using the same time-matching approach described in analysis 4. Additionally, the experimenter also tried to look for correspondences with the subjects' comments which immediately proceeded and followed her mentation. It should be noted that there were literally thousands of mentation items to be considered, and the experimenter only looked for literal correspondences (e.g., both the subject and agent mentation items referred to a dog). Given the nature of this examination, any trends which appeared to arise could only be viewed as providing suggestive information, as no firm conclusions could be drawn from such informal analyses.

Results:

This comparison yielded only two instances where the mentations of the subject appeared to correspond, within a five second interval, to those of the experimenter. Given the number of mentation items which

TABLE 3
Results comparing agent mentation categories
using the Wilcoxon Test

Mentation Characteristic	N	T	p*
1. Blankly looking at target	30	232	n.s.
2. Actively sending target	30	118	0.02
3. Visually looking at target	33	215	n.s.
4. Experiencing mental imagery	32	113	0.005
5. Thinking about target (vs. imagery	29	143	n.s.
6. Concentrating on a detail of target	32	707.5	n.s.
7. Viewing target as a whole	33	188	n.s.
8. Making general associations	30	147	n.s.
9. Making personal associations	13	31	n.s.
10. Mentation was not target-related	27	181	n.s.
11. Concentrating on shape(s)	27	172	n.s.
12. Concentrating on colour	28	115	0.05
13. Agent experiencing emotion	16	60	n.s.
14. Conveying a sense of motion	4	---**	---
15. Mentation had auditory component	17	75	n.s.
16. Mentation conveyed a sensation	3	---**	---
17. Agent physically experienced item	19	85	n.s.
18. Mentation experienced fleetingly	4	---**	---
19. Mentation persisted in agent's mind	8	10	n.s.
20. A recurrent mentation item	24	143	n.s.
21. Very clear mental image or thought	7	11	n.s.
22. Vague, unclear mental image or thought	7	0	0.02
23. Spontaneously occurring mentation	0	---**	---
24. Mentation interrupted on-going thoughts	3	---**	---
25. A bizarre mental image or thought	1	---**	---

* all p values are two-tailed
** insufficient data to compute

were compared, it is not surprising that a few correspondences were observed. It should be added that they were not exact correspondences, nor were they concerned with unusual or bizarre topics. No notable

correspondences were found between the agent's mentations and those of the subject which immediately preceded or followed the agent's.

DISCUSSION

The subject's scoring in this experiment was remarkable only to the degree in which they mimicked exact chance scoring (MCE sum of ranks = 100, subjects' sum of ranks = 102). Yet the independent judge's results approached a one-tailed significance ($p < 0.053$, one-tailed). This may indicate that the subjects, despite their having attended a training session designed to improve their judging ability, were still not able to assess the correspondences between their mentations and the target pictures as well as the independent judge (or, for the sake of completeness, the independent judge may have been displaying more psychic functioning than the subjects). As there was no control group who did not experience the training session, it is not possible to determine what effect, if any, the session had on improving the subjects' judging. However, these findings may suggest that the instruction given as to how to best perform the judging was not sufficient to enable naive subjects to judge as accurately as the independent judge. Regarding the importance of personal experiential knowledge of one's mentations in relation to the judging procedure, these results suggest, as did those of Palmer et al., (1979), that, in some cases at least, skill and experience may be more helpful in recognising target/mentation correspondences than subjective knowledge of one's experiences during the stimulus period.

As the subjects were unable to detect, to a significant degree, any psi-related imagery which they may have experienced, the examination of the data for weak and strong correspondences may have been futile. Yet the findings here may also suggest that weak and strong correspondences are equally important in conveying psi-mediated information. The independent judge, whose judging outcome approached significance, conducted a study of her own, based on the subject data collected in this study (Milton, 1986). In that study, Milton performed the same analysis, based on her ratings of the strength or weakness of the correspondences, and obtained similar results to those in this study (e.g., there was no significant difference between the two sum of ranks outcomes, where one was based upon weak and one upon strong correspondences).

The one significant finding of the analyses of the subject´s mentation characteristics, relating undeveloped imagery to target-related responses, is difficult to assess given the lack of significant psi-scoring and the possible influences of multiple analysis (this was the only significant finding out of the 14 tests run for this analysis). However, this finding does offer support to those of Sargent, Bartlett, and Moss (1982), who found that naive subjects scored significantly better when their imagery was unclear, rather than clear.

It is also possible that the types of mentation which might convey target-related impressions could vary from subject to subject. If this were the case, it is unlikely that any mentation characteristic could be identified using an across subjects analysis, as was done in this study. Thus, by only considering the various mentation analyses (weak and strong mentations, and subject and agent characteristics) across subjects, as done in this study, it is possible that some target-related mentational characteristics, specific to any given subject, may not have emerged.

The results of the analyses of the agent´s mentations found significant relationships between the subject making target-related responses and when the agent was : a) experiencing mental imagery (her eyes were closed); b) experiencing vague, unclear mental imagery, c) was concentrating on a colour found in the target; and d) actively sending to the subject. The anecdotal information suggesting that particularly striking target-related imagery would sometimes occur when the agent´s attention was temporarily distracted away from the target picture received no direct support, as this ´strategy´ was categorised in this study as occurring when the agent´s mentation had nothing to do with the target, and was not found to relate significantly to the subject making an immediate target-related response.

The significant agent mentation findings suggest that those specific agent ´activities´ or strategies may be beneficial to the subject receiving immediate target-related impressions. However, these results may be exaggerated due to multiple analyses. Also, the agent did not randomly vary her sending strategies throughout the sessions. Nor were any analyses made to see whether certain sending strategies may have significantly related to the subjects´ making target-related responses at some time during the session other than within a five second interval of the agent´s activity. Furthermore, the findings from this

study were based on the sending of only one agent. Given the possible effect that different agents may have with different subjects (White, 1976), these findings may not be generalisable to other agents. Thus the results from the agent mentation analyses may be seen as suggesting possible agent strategies beneficial to the subject receiving target-related impressions, but further research will be needed before any firm conclusions could be drawn.

The analysis looking for time-linked correspondences between the subjects' and the agent's mentation reports yielded no results which could even be considered to be suggestive of any such effect. However, it should perhaps be noted that there were several problems encountered in conducting this analysis. First, there were literally thousands of mentation items to be judged. The method of judging simply consisted of the experimenter reading one mentation of the subject, comparing it to the appropriate time-matched mentation of the agent, looking for obvious correspondences, and so on through all the mentation reports. After the first few reports had been examined, and it had become obvious that few if any time-matched correspondences had occurred, the experimenter quickly tired of the time-consuming judging process. The size of the judging task was such that no types of correspondences, other than literal correspondences (e.g.,, agent's mentation was 'dog', subject's mentation was 'dog') were considered. Furthermore, as there were no 'controls' against which the correspondences were being judged, no real claims could be made for any findings which arose.

A possible methodological improvement to the judging of this data would have been to have had a randomly selected sample of the agent's mentations from each session transcribed by a person not otherwise connected with the study. Such a transcript could then be independently judged for correspondences with the appropriate time-matched mentation of the subject, which would be placed with a control sample of other mentation items, and then also independently judged. This would have decreased the number of correspondences under consideration, allowing for a more thorough consideration of correspondence types. It would also have provided a MCE which would have allowed for various statistical analyses to be carried out on the results. However, there were no findings in this analysis to suggest that a more formal analysis would have received a different outcome.

One aspect of this study which may have contributed to the chance level of scoring was that the judging procedure of the subjects was

very time-consuming, and required a good deal of concentration. Most subjects took between one and two and a half hours to complete their judging. Furthermore, as they were unsupervised during the task, if they did tire, there was no supportive person to try to restore their morale.

Two pilot sessions using this judging procedure had been conducted, and in both sessions the subjects finished their judging within an hour and a half's time. Both subjects also found the mentation-characteristic form easy and interesting to complete. However, as the two subjects in the pilot sessions were the experimenter and the independent judge, it would not be surprising if they had found the judging procedure of greater interest than the other (main study) subjects did. Furthermore, as both of the pilot session 'subjects' were experienced judges, they may have found the judging task in general to be much quicker and less demanding than the other, largely naive, subjects.

There is no way of knowing how the length and complexity of the judging procedure may have affected the study's outcome. However, if a similar study were to be run again, a subject experimenter would be used. This would eliminate the need for the subject to go through the time-consuming task of replaying his mentation tape, and entering each mentation item onto the judging form. Aside from substantially decreasing the duration of the judging, a subject experimenter may have been useful in clarifying how to classify the mentation responses according to their characteristics, and would have helped to keep the subjects' motivation for the task at a higher pitch.

Another possibly confounding factor is that the subjects' imagery may have been somewhat self-conscious and/or forced, if they had felt any implied responsibility or obligation to produce imagery which would correspond to the fifteen mentation categories. This may have also left them with a tendency to monitor their imagery during the stimulus period for such correspondences. Such activities would obviously restrict the spontaneity of their imagery, a factor which is often associated with the apparent success of the ganzfeld technique, and thus may have been counter-productive. It should be noted that neither the experimenter nor the independent judge noticed any obvious differences between the mentations obtained in this study and that obtained in previous ganzfeld studies which they had conducted. However, no formal analysis was conducted comparing the mentations obtained in this study to those of others. Thus, no conclusions can be

drawn regarding the possible influence of the training session and judging procedure upon the subjects' mentations.

ACKNOWLEDGMENT

This paper is an abbreviated version of a study which is reported in full in Delanoy, 1986 (a Ph.D. thesis, University of Edinburgh). Further details of the full study can be obtained from the thesis or the author. The author would like to thank Dr. John Beloff who supervised this study and who, along with Dr. Murdo MacDonald, preformed the randomization procedures. Grateful thanks are also extended to Dr. Julie Milton for her independent judging, and her constant advice and support. Finally, Prof. Robert L. Morris is thanked for his helpful comments on the write-up of this study.

NOTES

1. Due to space limitations, not all the analyses conducted for this study are reported herein. The author has only reported those analyses which she felt were most central to the study. Details regarding the analyses which have not been reported here (involving the frequency of occurrence of the mentation categories, pre- and post-session questionnaires, and confidence and target-liking ratings), can be obtained from the author, and have been reported in Delanoy (1986).

2. The statistical analyses used in analyses 3 and 4 differ from those originally planned. Initially a proportions test was used, however, the use of this test was criticised as being inappropriate since the analysis units (individual mentation items) were not independent from each other. Thus, the use of a proportions test was abandoned, and the data were examined by the method reported in this paper. For finding an appropriate method with which to analyse this data, grateful thanks are extended to Ephraim Schechter, Donald McCarthy, George Hansen, Jessica Utts, and Julie Milton.

ABSTRACT

This study was designed to investigate the relationship of specific

´types´ of subject and agent mentation to psi-scoring using the ganzfeld technique. A related area of enquiry was whether weak or strong correspondences between mentation imagery and the target picture best conveyed psi-related impressions. The data was judged by an experienced independent judge and the subjects, who received training in how to preform the judging.

Neither the independent judge nor the subjects´ judging obtained a significant level of psi-scoring, and there was not a significant difference between the two outcomes. Examining the subjects´ mentations, undeveloped imagery was found to convey target-related imagery more frequently than the other types of responses ($p = .046$, two-tailed). It was also found that four categories of agent activity (or agent mentations) corresponded to the subject making target-related responses to a significantly greater degree than the other agent activities being examined. These four categories were: a) actively sending the target ($p = .02$, two-tailed); b) experiencing mental imagery ($p = .005$, two-tailed); c) concentrating on the colour of the target ($p = .05$, two-tailed); and, d) experiencing vague, unclear mental imagery or thoughts ($p = .02$). Neither weak or strong correspondences obtained significant results, nor was the difference between the two significant.

REFERENCES

Child, I.L. & Levi, A. ´The use of judges´ ratings to test hypotheses about psi processes´, Journal of the American Society for Psychical Research, 1980, Vol.74, no.2, pp.171-181

Delanoy, D.L. ´The training of psi in the ganzfeld´. In W.G. Roll, R.L. Morris & R.A. White (Eds.), Research in Parapsychology 1981, Scarecrow Press, Metuchen, New Jersey, 1982, pp.157-159

Delanoy, D.L. ´The training of extrasensory perception in the ganzfeld´. Ph.D. Thesis, University of Edinburgh, Edinburgh, Scotland, 1986

Harley, T. & Good, D. Unpublished paper, 1981

Honorton, C. & Harper, S. ´Psi-mediated imagery and ideation in an experimental procedure for regulating perceptual input´, Journal of the American Society for Psychical Research, 1974, Vol.68, no.2,

pp.136-168

Milton, J. 'The effect of the presence of an agent on ESP performance and of the isolation of the target from its controls on displacement in a ganzfeld clairvoyance experiment'. In R.A. White & R.S. Broughton (Eds.), Research in Parapsychology 1983, Scarecrow Press, Metuchen, New Jersey, 1984, pp.85-86

Milton, J. 'The effect of agent strategies on the percipient's experience in the ganzfeld'. In R.A. White & J. Solfvin (Eds.), Research in Parapsychology 1984, Scarecrow Press, Metuchen, New Jersey, 1985, pp.1-4

Milton, J. 'Displacement effects, role of the agent, and mentation categories in relation to ESP performance'. Ph.D. Thesis, University of Edinburgh, Edinburgh, Scotland, 1986

Palmer, J. 'Extrasensory perception: research findings'. In S. Krippner (Ed.), Advances in Parapsychological Research 2. Extrasensory Perception. Plenum Press, New York, 1978, pp.59-243

Palmer, J., Bogart, D.N., Jones, S.M. & Tart, C.T. 'Scoring patterns in an ESP ganzfeld experiment', Journal of the American Society for Psychical Research, 1977, Vol.71, no.2, pp.121-145

Palmer, J., Khamashta, K. & Israelson, K. 'Scoring patterns in an ESP ganzfeld experiment'. Journal of the American Society for Psychical Research, 1977, Vol.73, no.4, pp.333-348

Rand Corporation 'A Million Random Digits with 100,000 Normal Deviates'. The Free Press, New York, N.Y., 1955

Rogo, D.S. 'Free response ganzfeld experiments with a selected subject'. In J.D. Morris, W.G. Roll & R.L. Morris (Eds.), Research in Parapsychology 1975, Scarecrow Press, Metuchen, New Jersey, 1976, pp.176-179

Sargent, C.L. 'Exploring Psi in the Ganzfeld'. Parapsychology Foundation, Inc., New York, 1980

Sargent, C.L., Bartlet, H.J. & Moss, S.P. ´Response structure and temporal incline in ganzfeld free-response GESP testing´. In W.G. Roll, R.L. Morris & R.A. White (Eds.), Research in Parapsychology 1981. Scarecrow Press, Metuchen, New Jersey, 1982, pp.79-81

Sargent, C.L., Milton, J., Payne, J. & Bennet, S. Unpublished study, 1982

Siegel, S. ´Nonparametric Statistics for the behavioral sciences´ (International Student Edition). McGraw-Hill, Inc., London, 1955

Sinclair, U. ´Mental radio´. Charles C. Thomas, Springfield, Illinois, 1930/1962

Solfvin, G.L., Kelly, E.F. & Burdick, D.S. ´Some new methods of analysis for preferential-ranking data´. Journal of the American Society for Psychical Research, 1978, Vol.73, no.2, pp.93-110

Sondow, N. ´Effects of associations and feedback on psi in the ganzfeld: is there more than meets the judge´s eye?´ Journal of the American Society for Psychical Research, 1979, Vol.73, no.2, pp.122-150

Terry, J.C. & Honorton, C. ´Psi information retrieval in the ganzfeld: two confirmatory studies´. Journal of the American Society for Psychical Research, 1976, Vol.70, no.2, pp.207-217

Van De Castle, R.L. ´Psi abilities in primitive groups´. Proceedings of the Parapsychological Association, 1970, Vol.7, no.1, pp.97-122

Warcollier, R. ´Experimental telepathy´. Boston Society for Psychical Research, Inc., Boston, Massachusetts, 1938

Warcollier, R. ´Mind to mind´. Collier Books, New York, 1948/1963

White, R.A. ´The influence of persons other than the experimenter on the subject´s scores in psi experiments´, Journal of the American Society for Psychical Research, 1976, Vol.70, no.2, pp.133-166

Deborah L. Delanoy
Department of Psychology
University of Edinburgh
7 George Square
Edinburgh EH8 9JZ
Scotland, U.K.

APPENDIX I

Instructions Given to the Subjects in the Training Session (copies of these instructions were also provided for review during the ganzfeld sessions)

THINGS TO REMEMBER DURING THE STIMULUS PERIOD

1. Relax!
2. Keep your eyes open.
3. Say out loud everything you experience, as it occurs.
4. Do not try to obtain imagery, just relax and let it come to you.
5. Forget about the experiment. Do not worry about the agent or whether you are receiving any ESP related imagery.
6. If some imagery does not take any recognisable form, do not try to 'turn it into' some recognisable object. Just describe its form as best you can.
7. Try not to follow long chains of thought.

THINGS TO REMEMBER DURING THE JUDGING

1. Take your time! Do not be in a hurry to finish.
2. Study each of the four target pictures very carefully. Try to examine each picture from these two perspectives: 1st: Look at it analytically, consider the picture as a whole and in terms of its various components. Observe carefully everything that is contained in the picture, down to its smallest details. Make note of the various shapes contained in the picture and examine these from different angles (e.g., look at the various shapes in the picture when it is viewed upside down); 2nd: Look at the picture from a more personal and holistic perspective. Try and receive general impressions from the picture. Also think of any associations, personal or otherwise, which you might have towards the picture.
3. Make a note of each item of mentation on the judging form. Complete the imagery categorisation for each mentation.
4. Rate each picture according to its degree of correspondence to each mentation. Use a 1 - 5 scale where:
    1 = a small degree of possible correspondence
    2 = a stronger correspondence than for a 1 rating, but still not a particularly convincing match
    3 = there is a definite correspondence, but it is

lacking some respect(s)
4 = a good correspondence, there are marked similarities between the target and the mentation item
5 = an excellent correspondence with little, if any, dissimilarity

5. The various types of correspondence which you should be looking for are as follows:
-- Direct correspondences - What you experienced is contained in the picture (to some degree); e.g., you saw a flower and there are flowers in the picture
-- Associative correspondences - The mentation has an associative relation to the picture. This association can be either personal, e.g., you saw your father and there is a father-like figure or something which you associate with your father (a pipe perhaps) in the picture; or, of a more general nature, e.g., your felt thirsty and the picture is of a desert scene.
-- Similarities of features and/or shape - e.g., you have an image of a floating half circle and there is an umbrella in the picture
-- Similarities of meaning and/or purpose - e.g., you have an image of a train which could relate to another form of transportation or to travelling.
-- Symbolic correspondences - Examples of this could be having an image of a thunderbolt when Zeus was in a picture, or seeing a lion when the picture has a courageous theme or component.
6. When performing the final ranking (1st choice, 2nd choice, etc.) of the target pictures, you should rank them in order of the number of correspondence points each picture received. Thus, the picture which is chosen as the target for that session should be the one which had the highest score after you added up all the ranking points given to each picture. The only time the above procedure should not be followed is when you had some outstandingly excellent correspondences to a picture, which nevertheless ended up with fewer points than another picture, which contained more correspondences but of a much poorer quality. However, as a general rule, do not allow your judging to be swayed by a few mentation items. You may not tie any ranks. If two pictures have the same number of correspondence points, rank the picture which had the better quality of correspondences above the other.
7. When doing the judging try to disregard any personal preference in terms of liking or disliking which you may have towards the pictures.
8. As a general rule of thumb, let logic, not intuition, be your guide whilst judging.

APPENDIX II

The 15 Imagery Categories (Characteristics) for the Subjects' Mentations (The following list was given to the subjects.)

IMAGERY CATEGORIES

Type of Image:
1. Interrupts a Chain of Thought: an image which interrupts a series of related images or a scene of related action
2. Result of a Transformation: when one image turns into another; e.g. "I see a beach ball...the beach ball just became a clown's face"
3. Developed from an Undeveloped Image: when a recognisable image develops from an unrecognisable image; e.g., "I see something rather like a floating half circle, oh, now I can see that it is a dome of a church"
4. Spontaneous: an image about which you have no idea why it occurred and which is unrelated to other imagery .

DURATION:

5. Fleeting: a brief image which quickly appears and disappears
6. Persistent: an image which stays in the mind awhile
7. Recurrent: an image which appears several times throughout the session

CLARITY:

8. Undeveloped: an image which is unrecognisable; its features are describable in terms of pure form (including geometric forms); e.g., "I see several straight lines intersecting each other, rather like a Y with several tops" Note; it may be useful to make sketches of such imagery; this category may also refer to an idea or thought which never develops into an image, e.g., "I have a thought of a frog, but I'm not seeing one"
9. Detailed: an image which is very clearly defined in terms of lines and shapes; this image may be either hallucinatory or dream-like, or be as clear as you would expect it to be using normal vision
10. Intense colour: an image which has quite bright and vivid colouration

CONTENT:

11. Bizarre: an image which contains an unusual combination of elements; e.g., "I see a green and purple striped strawberry"
12. Personal memory or experience: an image which related to a personal memory or a personal experience

MISCELLANEOUS:

13. Auditory: an image which had an auditory content; e.g., "I hear many voices in conversation as one might at a party" or "I just heard the word ´frog´"
14. Physical Sensation: a sense of experiencing some form of physical sensation which may or may not be related to an image; e.g., "I feel as if I´m floating"
15. Physical Reaction: when one does experience an actual physical response to an image; e.g., "that image actually caused my body to jump in the chair"

APPENDIX III

The List of the 25 Agent Mentation Categories (characteristics of the agent's thoughts and/or mental images)

1. Blankly: the agent was looking at the target, but had no particular thoughts passing through her mind, and therefore, was making no mentations
2. Sending: the agent was actively trying to send the target to the subject
3. Vision: the agent was looking at the described item
4. Imagery: the agent was experiencing mental imagery (her eyes were closed)
5. Thinking: the agent was aware of thinking about the mentation item
6. Detail: the agent was concentrating on a detail of the target picture
7. Overall: the agent was concentrating on the whole of the target picture
8. General Association: the item content had a general, consensual associative connection to the target
9. Personal Association: the item content had an associative connection to the target which was personal to the agent
10. Non-Target: the mentation had nothing to do with the target indicating that the agent's attention was temporarily diverted away from the target
11. Shape: the shape of an object contained in the target was particularly noticed
12. Colour: the colour of an object was particularly noticed
13. Emotion: the item conveyed an emotional response of the agent
14. Motion: the item conveyed a sense of motion
15. Auditory: the item had an auditory component
16. Sensation: the item conveyed a sensation
17. Experiential: the agent physically experienced the item
18. Fleeting: the agent experienced the item very briefly
19. Persistent: the item remained in the agent's mind for a relatively long period
20. Recurrent: the item occurred to the agent more than once
21. Detailed: a mental image or thought was quite clear and/or detailed
22. Vague: a mental image or thought was unclear and/or vague
23. Spontaneous: a mental image or thought occurred spontaneously, with no relation to the agent's on-going stream of consciousness

24. Interruption: the item interrupted the agent´s on-going steam of consciousness
25. Bizarre: the mentation item contained a bizarre, unusual element

EUR.J.PARAPSYCHOL., V.7, 1988-1989

WHY THE GANZFELD IS CONDUCIVE TO ESP: A STUDY OF OBSERVATIONAL THEORY AND THE PERCIPIENT-ORDER EFFECT

Joop M. Houtkooper
Loftur R. Gissurarson Erlendur Haraldsson

INTRODUCTION

A sizeable number of ESP experiments employing the ganzfeld condition have shown significant above chance scoring. Claims have even been made that the ganzfeld was the condition ensuring parapsychological experiments with reproducible results (Honorton, 1978; Sargent, 1980). We would not make such a strong claim ourselves, but recent reviews of ganzfeld experiments indicate that the procedure has been rather successful (Schouten, 1981; Honorton, 1985), the critical review by Hyman (1985) notwithstanding. Moreover the rationale of the ganzfeld condition is prima facie convincing and theoretically sound, in spite of the continuing debate about the success and quality of previous experimental work (Hyman, 1985; Honorton, 1985; Blackmore, 1987; Harley and Matthews, 1987; Sargent, 1987 and Parker and Wiklund, 1987).

Further research may provide more insight in the theory of the ganzfeld condition and we will show that more than one explanation of its success may be constructed. Testing such explanations and therewith the theoretical framework from which they emerge might be important for the direction of further research.

This is one of the aims of the present study. In working out the details of the experimental procedure we benefitted from Hyman´s criticism of the ganzfeld experiments, and the care that has been taken in describing the procedure can be considered a response to his critical review (Hyman, 1985; Haraldsson and Gissurarson, 1985).

We will consider here two models of ESP in the ganzfeld: first, the usual "transmission model" and, secondly, we will present here an interpretation of ganzfeld-ESP from the viewpoint of observational theory, i.e. the "observational model".

The transmission model supposes that information is received by the percipient during the period of sensory deprivation that constitutes the ganzfeld condition. This information may, for instance, manifest itself as visual or auditory imagery during the ganzfeld condition. Since the auditory and visual stimulation of the percipient is very constant and therefore without any information content, it is assumed that any information about the actual target, impinging on the conscious or subconscious of the percipient, will be likely to become so prominent that the percipient will make a correct guess about the target. The information the percipient receives, is usually assumed to be "sent" by an agent, who is instructed to make a conscious effort to do so. The "sending" almost always takes place contemporaneously, although efforts to produce precognition, employing sending at a later time, e.g. the morning after, have been made. We may consider this ganzfeld procedure to have a good rationale. It hinges on the assumption that information is received as a weak signal, most easily depicted as "mental radio", though more elaborate mechanisms might be called upon. Nevertheless, this assumption suggests to us to call the above-described model of ganzfeld ESP the "transmission model".

To explain the observational model of ganzfeld ESP, we will first introduce a few concepts from observational theory (Millar, 1978, Houtkooper, 1983). In observational theory the different paranormal phenomena are regarded as having a common denominator. The supposed common mechanism is the influence an observer of a random event has on the probabilities of the different possible outcomes of that event. The magnitude of this influence is assumed to depend upon properties of the observer, such as his motivation. Different mathematical formulas may be given to describe the effects on the probabilities of the outcomes of the observed random event (Schmidt, 1975; Walker, 1975), especially if more than one observer is taken into consideration (Hartwell, 1977; Schmidt, 1976, 1982; Houtkooper, 1983).

These issues of what form the theory should take are important by themselves, but they hardly concern us here. At present, dealing with ganzfeld-ESP, we have to be careful to note what exactly is the random event and who the observer, or observers.

First, we regard the random event to consist of the correspondence between the target sent by the agent and the guess made by the percipient. (We will not consider here the case of the percipient producing a protocol and judges producing the correspondence between protocol and the different target alternatives.) The actual target is assumed to be randomly chosen from a well-defined finite set of targets (in the present experiment, four different pictures) and the percipient makes his guess by chosing from this set. Neither the guess, nor the actual target alone contain information about the correspondence between the two (an exception has to be made for known non-random guessing behavior together with known actual target). Only the combination of target and guess contains the information which decribes the outcome of the random event. Its possible outcomes are two: correspondence and non-correspondence, or a hit and a miss. These have well-defined a priori probabilities, given that the choice of the actual target is generated randomly.

Secondly, the observer of the random event is the person who observes the correspondence between the guess and the actual target. In principle the influence is not limited to only the first observer, although the supposed influence of later observers entails methodological and theoretical problems, which fall outside the scope of the present discussion.

The observational model of ganzfeld ESP therefore considers the observer to be the person checking the outcome of the trial, that is, he (or she) might be one of the experimenters, or the agent, or the percipient, depending on the exact experimental protocol.

The explanation of the success of the ganzfeld procedure thus hinges on the checking of the correspondence between target and guess. For instance, if the percipient himself, after making his guess, immediately checks the correspondence with the target, we may assume that he still is in a relaxed and receptive state from the experience of the ganzfeld condition. This state may well be psi-conducive. Moreover, he may be all the more relaxed if he thinks that his real task is over, and even more so if he had been told that he only had to report what he received from the agent, who had to do the hard work of

concentrating to get something transmitted. (By analogy: in ordinary radio broadcasting, the sender is a big, power-consuming installation, whereas receivers can be tiny things.) Therefore, in ganzfeld experiments the conditions may well be favorable for a 'release of effort' effect (Stanford, 1977) to be produced by the agent.

Besides the agent, the other persons involved in the experiment have to put much effort in an experimental session and this might enhance their interest in getting some result. As it is essential to observational theory that the observer is motivated to obtain a particular outcome, it is the psychological condition of the checker, or checkers, which favors the occurrence of better than chance scoring. Thus, there is a good rationale for above chance scoring in ganzfeld ESP, also from the viewpoint of observational theory. The difference with the classical viewpoint is that the checker, instead of the percipient and the agent, is crucial to the obtained result.

To test the relative merits of the observational model and the transmission model, our experimental procedure differs from other ganzfeld experiments. After the percipient has made his guess, a formalized observation procedure takes place, in which two different conditions are employed: In the first, the percipient is the first person to observe the correspondence between his guess and the target, and the experimenter (E2) who stayed with the percipient is the second to observe. In the second condition, the agent first observes the correspondence between the target he has sent and the guess of the percipient, while the experimenter (E1) who stayed with the agent is the second observer.

In both conditions, the other persons involved in the experiment are informed about the outcome after the second observer has become knowledgeable about it. The two conditions entail observations by persons in different psychological conditions, of which presumably one is sufficiently more psi-conducive than the other, so that we expect a difference in scoring between the two conditions. As the reason for the success of the ganzfeld procedure has been put on the condition the percipient is subjected to, a higher level of scoring may be expected when the percipient is the first observer of the outcome than when the agent is.

It might be argued that this way of testing the observational hypothesis is complicated by a possible effect of the experimenters (E1 and E2), as second observers of the outcome. This effect may or

may not be in the same direction as the primary effect of the agent or the percipient. Although this source of possible contamination may cause a cancellation (or an enhancement) of the effect of the first observers by the second observers, it should not affect the conclusion if an effect is found in the present experiment: if the effect is due to the second observers, it also confirms the observational model. This can be grasped from the fact that the observational condition is imposed after the trial is finished in the classical sense. In the last section of this paper we will discuss the significance of the present experiment for the theory.

Successful experiments showing the existence of "experimenter effects " or "analyzer effects" have been reported in the literature, as reviewed by White (1976). We believe the lack of limits and the lack of control of experimenter effects has led to the point where any effect obtained can be ascribed to the preferences of the experimenter. This is not a fruitful situation, where progress by experimentation has become a doubtful enterprise. The present experiment is designed to clarify the role of the observer, which may lead to ways to avoid experimenter effects.

In research with plethysmographic ESP (Haraldsson, 1972, 1980; Houtkooper and Haraldsson, 1985), it has been found that in the case in which two persons alternate the roles of agent and percipient in two successive sessions, the second percipient has the tendency to produce psi-hitting whereas the first percipient produces null results or psi-missing. To test this percipient-order effect in ganzfeld ESP, we adopted a similar procedure, employing pairs of subjects who switch roles. The percipient-order effect in ganzfeld ESP has been studied by Haraldsson and Gissurarson (1985). In two experiments with 32 and 38 ganzfeld sessions the percipient order effect was in the expected direction in one experiment ($z=1.57$; $p=.06$) but zero in the other.

EXPERIMENTAL DESIGN

In the present experiment we decided to test three hypotheses:

H1: The overall effect: there will be a greater number of hits than the mean chance expectancy. This is the hypothesis that the ganzfeld condition is indeed a psi-conducive and we expect it to work that way in our experiment. It is to be tested by the binomial test, one-

tailed.

H2: The observational effect: More hits are expected in the condition where the percipient first observes the outcome, than if the agent first observes it. To be tested by the chi-square test, one-tailed.

H3: The percipient-order effect: More hits are expected for the second session than for the first session of the two in which a pair of subjects alternate in the roles of percipient and agent. To be tested by the chi-square test, one-tailed.

METHODS

The experimental procedure was adapted from the earlier version (Haraldsson and Gissurarson, 1985) to include the observational condition.

Subjects:

Twenty selected unpaid persons, all of whom had been tested before in various experiments, participated in pairs in the experiment on two occasions. Each occasion consisted of two ganzfeld sessions. In one session the subject had a role of percipient and in the other of agent. This arrangement provided a pool of a total of 40 ganzfeld sessions, as each subject participated in four sessions. Only three subjects had participated before in ganzfeld experiments.

Preliminary procedure in the ganzfeld:

The subjects participated pairwise in every session. First one subject acted as the percipient and the other as the agent. A second session followed immediately after the first in which the subjects changed roles, such that the first percipient became agent and the first agent took over the role of the second percipient. This procedure was used to test the occurrence of the "percipient-order effect". When a pair of subjects arrived at the lab they were welcomed by two experimenters. All necessary details of the ganzfeld technique

were explained to them and a coin was tossed by one of the experimenters to determine which of the subjects acted first as the percipient. One of the subjects was asked if he preferred to act first as percipient or agent. If the emblem turned up in the coin-toss, then that subject obtained the role which he stated a preference; if the krona turned up he started in the other role.

From then on the first experimenter (E1, Haukur Hjaltason, a medical student) took care of the agent and brought him/her to the agent´s room. E1 then left, after closing the door of the agent´s room, to an adjacent room to select a target picture for the session.

The second experimenter (E2, LRG) handled the percipient and showed him to his room. One room was situated between percipient´s and agent´s rooms (see floorplan in Figure 1). The percipient was laid comfortably in a reclining chair ("Lazy Boy"), where E2 placed on him a headphone and two yellow ping-pong ball halves over his eyes. A 60 watts white lamp was lighted 18 inches in front of his face. (Percipients have usually reported this distance to produce a totally unpatterned visual field without shadows or uncomfortable heat). When E2 had given instructions to the percipient, he turned on the tape recorder which the headphone was attached to, left the percipient´s room and went into the middle room where he stayed till the end of the ganzfeld. He closed both doors between the two rooms. Through his headphone the percipient first heard a piece of music composed by Mozart for a few minutes and then white noise till the end of the session. A microphone in the percipient´s room was attached to a second tape recorder which was placed in E2´s middle room. Through a headphone E2 listened to all statements made by the percipient who had been told to "think out loud". Whatever he reported was recorded on tape and also written down simultaneously by E2. A one-way mirror between E2´s and the percipient´s room allowed E2 to keep a watch on the percipient.

Target pool and randomization procedure:

100 Target pictures had been selected by EH as a targetpool (see for details Haraldsson and Gissurarson, 1985). All targets were glued on thick 18 cm x 26 cm cards. The target pool consisted of 25 different, fixed target sets, each set containing four pictures that differ widely in content. The pictures in each set had been hand-shuffled and

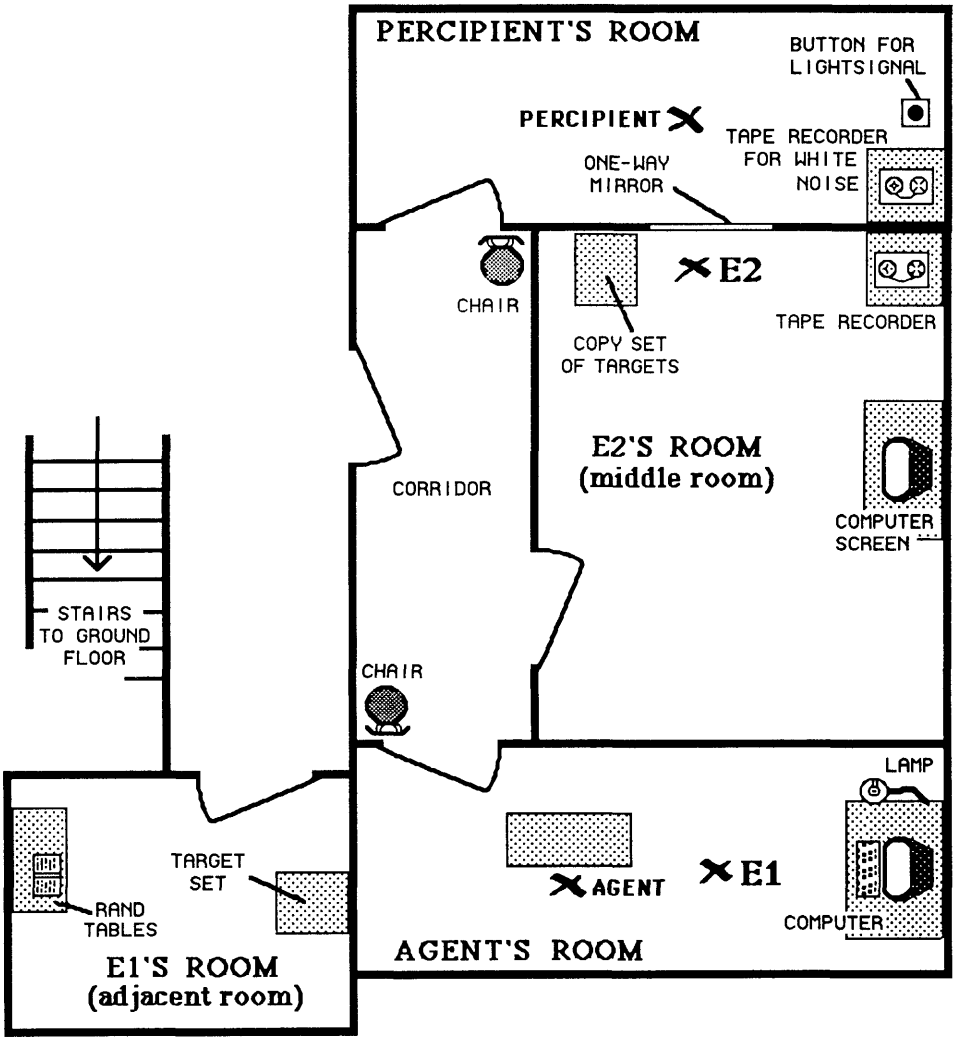

FIGURE 1
The experimental rooms in the basement of the
Faculty of Social Sciences at the University of Iceland

then numbered from 1 to 4, with the number written on the back of each picture and on the opaque envelope which contained each picture. The four targets (envelopes) of each set had been placed in a larger envelope and numbered in the same fashion from 1 to 25. For each target set envelope there was an exact copy set envelope numbered from 1 to 25, each including four duplicate pictures in four opaque envelopes.

The first set was used for the agent and the copy set for the percipient's evaluation which took place immediately after the ganzfeld session. The target picture sets were kept in one drawer in a filing cabinet in E1's adjacent room which was locked. The copy picture sets were kept in a drawer in a locked filing cabinet in E2's room. The key to the cabinet containing the copy sets remained with LRG (E2) all the time and the key to E1's room with the filing cabinet containing the target sets, remained with E1. E1 never got the opportunity of touching the copy set. He could thus unconsciously not have marked, folded or left any sort of trace on the correct envelopes. The copy set remained "clean" all the time.

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

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

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

E1 picked the selected "target" set envelope out of the filing

cabinet in his adjacent room. He then took the envelope with the target picture out of the set envelope, got an empty, new envelope (for the observational part), and went to the agent´s room. The rest of the set for the agent was left in the large set envelope in E1´s room. Before entering the agent´s room, E1 locked his room, closed the door to the corridor that leads to his room and knocked two times on the door of the percipient´s room indicating that he was ready. (Although farfetched, it could be argued that this act by E1 provides opportunity for unintentional cueing of the percipient´s experimenter, E2, as to the number of the correct picture. For such cueing to develop unintentionally, a very much longer series of trials probably has to be conducted.)

When E2 had taken care of the percipient and upon hearing the two knocks from E1, the experiment was ready to start. E2 signalled to E1 that the ganzfeld period had started by pressing a button in the percipient´s room, which turned on a light in the agent´s room.

Instructions to the percipient:

With percipient seated in the reclining chair: "You are to keep your eyes open as much as possible throughout the session. You will first hear music by Mozart that will help you to relax, followed by one minute silence. Then you will hear a sound called white-noise throughout the session, after I have placed the headphones over your ears. I will leave and monitor you from the next room. I will turn out the main lights before I leave. When the ganzfeld period is over, in about half an hour, I will turn them on again".

While putting the equipment on percipient: "Sit as comfortable in your chair as you can. Try to relax as well as you can while the music is on and try to stay relaxed all the time. When you hear the white noise, begin to observe your mental processes and start "thinking out loud". Continue to share all your images, thoughts and feelings, which pass through your mind, with me throughout the session. Try not to concentrate on anything special. Do not bother about, or cling to the sender and his picture. He will do all the concentration work. At some point during the session he will send you a message or pictorial images, forms, shapes, colors or orientations. Do not try to guess, anticipate or conjure up the images or message. It might appear spontaneously, automatically, and effortlessly in consciousness at the

appropriate time. The only thing you have to do is to relax and remain awake all the time. Notice and observe your impressions and thoughts without strain and do not cling to any of them. Just observe them as they go by and speak fluently and continually about everything you think of; ideas, impressions, emotions, sensations and feelings, sudden impulses and pictures."

In the rating period: "The picture will very seldom fit exactly. You are supposed to evaluate what picture is most likely to correspond to your thoughts, feelings and ideas during the ganzfeld period. Perhaps there will only be few items or details in the pictures that correspond to your mentation report".

Instructions to the agent:

When E1 finished the randomization procedure (and filling in the form for the session) and E2 had signalled the start of the ganzfeld period, E1 gave the following instructions to the agent: "In this envelope is the target picture that you are to try to "send". In a little while I will open it and hand you the picture."

When the sending period started: "Look carefully and attentively at the picture, imagine that you are seeing what the picture is actually of, and at the same time think of your companion sitting in the other room. Imagine that he is seeing the picture, that its contents are in his mind and that he is surrounded by it. Try to make the picture a living reality in your mind and hope that the percipient in the other room comes to think of the contents of the picture". The sending period started after 15 minutes of ganzfeld (four minutes of Mozart, one minute of silence and ten minutes of white noise). Then E1 took the target out of the envelope and handed it to the agent who kept it for 20 minutes. The experimental session therefore lasted 35 minutes.

Procedure at the end of the ganzfeld:

At the end of the ganzfeld period E2 went back to the percipient´s room (bringing along a brand-new envelope for the observational part) and signalled for a second time to E1 that the ganzfeld was over. E2 then released the percipient of the headphone and the ping-pong ball

halves. E2 asked the percipient if he wanted to add something to his mentation report, draw sketches of, or associate around his impressions. After that E2 signalled the light for the third time.

Upon seeing the third signal E1 wrote the number of the target set, which had been randomly selected, into a computer. Two screens were connected to the computer, one situated in the agent's room and another in E2's room. E2 left the percipient's room and entered his middle room. He read the number from the screen in his room and opened the locked filing cabinet with the copy sets. He then picked up the correct duplicate set and returned immediately to the percipient's room. E2 opened the envelopes and handed the percipient the 4 pictures of the set and asked him to rank them according to their similarity to his mentation report. While the percipient was thinking about the pictures E2 read for him the mentation report. The percipient was also asked to explain why he ranked the pictures the way he did, from the most similar to his report and memory to the least similar. He was finally asked to score them on a scale from 0 to 100. These scores were gathered for eventual post-hoc analysis of how confident the percipient was of the similarity between the target and his images or mentation report.

Testing observational theory:

After the ranking period - when E2 had joined the percipient and he had made his guess of the target picture and ranked all four pictures on the scale from 0 to 100 - E2 pressed the button in the percipient's room for the fourth time, signalling to E1 that the guess had been recorded. Then E1 took the target picture, put it into an unmarked brand-new envelope which he sealed, and went out of the agent's room placing the envelope on a chair in front of the percipient's room and knocked on its door. About a minute later E2 opened the door, noticed carefully whether the other door was closed, and told the percipient to pick up the target in the sealed envelope. E2 told him not to open the envelope until he (E2) had gone out of the room. Then E2 went to the corridor closing the door of the percipient's room behind him. E2 waited in the corridor about one minute while the percipient took a look at the two pictures (the target and the guess) to see whether they corresponded or not. E2 returned into the percipient's room and took a look at the results too, E2 thus being the second person to observe the results. After that E2 and the percipient joined the

others in the agent's room, the agent and E1 together being third to observe the results.

When the pair, having had the percipients observe the results first, arrived for their second time the observing procedure was changed, such that the agent became the first observer. After the ganzfeld procedure, E2 took the picture being the percipient's guess of the target and sealed it in an envelope. E2 placed it on a chair in front of the agent's room and knocked on the door, signalling to E1 that the picture had been put in front of the door. E1 opened the door after a minute, and if the door to the percipient's room was securely closed he told the agent to pick up the envelope. E1 then left the agent's room, closing the door behind him, and waited in the corridor for a minute to allow the agent to compare the guess and the target. E1 returned into the agent's room, being the second to observe the results. After that E1 and the agent joined the percipient and E2 in the percipient's room, E2 and the percipient together being the third party to observe the results.

The first pair of subjects to arrive in the experiment had the role where the percipient looked first at both pictures before E2. This arrangement was repeated in session two which immediately followed the first one (to control for testing the percipient-order effect; both subjects when having the different roles got the same treatment). On next occasion, when the same pair of subjects arrived again to take two sessions in a row, the agent was the first to look at both guess and target, followed by E1 looking at the same pictures. This was repeated unchanged for the two sessions.

The second pair of subjects to arrive in the experiment had the role of the agent looking first at both pictures and then E1, this arrangement being unchanged for the two sessions. When the second pair arrived next time the percipient observed both pictures before E2, doing so for both sessions. The third pair had the same role that the first pair had and so forth. The participants were kept unaware of which course was taken until at the end of the ganzfeld procedure.

This manipulation of the roles of subject pairs has been done to achieve an experimental design in which the experimental variables: observer and percipient-order and the session-order for each subject are balanced. This setup allows for a comparison between results of percipient receiving the feedback first against the agent receiving the feedback first, with fair certainty that if a "decline effect"

plays a role, it will do so equally for both groups. In the same way, it allows the percipient-order effect to be tested.

Summary of the roles of E1 and E2

Following are the two experimenters' separate roles, when the percipient is the first to view both target and guess pictures together. When the agent is the first observer, steps 7 and 8 for the first experimenter and steps 12 to 14 for the second experimenter have to be changed accordingly.

First experimenter, E1 (H.H.):

1) E1 took the agent to his room.

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

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

4) E1 fetched the target set envelope out of the locked filing cabinet. He took the target picture (in envelope) out of the set envelope and went with it (leaving the rest of the pictures inside the set envelope in his room but taking an extra brand-new envelope for the observational part with him) to the agent's room after locking the doors to his room and the corridor, and knocking twice on the percipient's door.

5) E1 gave the agent instructions and chatted with him if time allowed. After 15 minutes of the experimental session (i.e. after the first light) E1 handed the agent the target picture who started "sending", doing so for the next 20 minutes.

6) E1 waited until E2 had gone out of his room at the end of the session and into the percipient's room (the second light signal). After the third light signal E1 put the number of the target set in the computer, making it appear on a screen in E2's room.

7) E1 waited for the fourth signal light, which indicated that he

was to put the target picture in an envelope and place it on the chair in front of the percipient's room.

8) E1 and the agent waited for E2 and the percipient to arrive after they had finished observing the results.

9) E2 placed the pictures on a table in the order that the percipient had ranked them. E1 asked the percipient to describe to him and to the agent why he had ranked each picture the way he had. After the percipient had done that, some questions and discussions took place.

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

11) E1 and E2 took the target and copy set envelopes back to the filing cabinets.

Second experimenter, E2 (LRG):

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

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

3) When hearing the two knocks on the door, E2 turned on the tape recorder for the headphones as well as the 60 watts lamp. Then E2 signalled to E1 that the experiment had started (by pressing the button which turned on a light in the agent's room). E2 left the percipient's room.

4) E2 entered his room (the middle room between the agent's and the percipient's rooms), closing both doors on his way. Fifteen minutes after he had turned on the tape recorder producing the sound for the ganzfeld condition, he turned on the second tape recorder in his room which was connected to the percipient's microphone, so that everything the percipient said was recorded.

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

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

7) E2 left his room after 35 minutes and went to the percipient and released him of the equipment. E2 signalled that the ganzfeld period was over.

8) E2 asked the percipient to add to his mentation report any associations, feelings or impressions he may have had, this being written down by E2 and also tape-recorded. The percipient was, furthermore, asked if he wanted to add drawings of his impressions to the mentation report.

9) E2 signalled for the third time. He went back to his middle room, looked at the computer screen, read the number of the target set and opened the locked filing cabinet with a key (which had remained in his possession all the time). He picked out the exact copy set envelope and returned to the percipient´s room.

10) E2 read the mentation report to the percipient.

11) E2 asked the percipient to evaluate all 4 pictures and rank each one of them, such that the picture corresponding most with the mentation, being given the lowest number (1). The percipient was also asked to tell why he had ranked the pictures in the way he did. He was asked to grade each picture in the set on a scale from 0 to 100.

12) E2 signalled the light for the fourth time to E1 (who was in the agent´s room). E2 waited about a minute and then he let the percipient fetch the envelope outside the door of the percipient´s room.

13) He left the percipient alone with the sealed envelope, closed the door to the percipient´s room, and waited in the corridor while the percipient looked at the results.

14) E2 returned to the percipient´s room to take a look at the results. Then E2 and the percipient left the room and joined E1 and the agent.

RESULTS

The results for the numbers of hits in the different conditions in the experiment are given in Table 1. The results for the hypotheses are:

TABLE 1.
Hits obtained under different conditions of first observer and percipient order

First observer: Percip.order:	percipient	agent	total
1st percipient	3	3	6
2nd percipient	2	2	4
Total	5	5	10

H1: The total number of hits, 10, is exactly equal to the mean chance expectation for 40 trials with P ≈ 1/4. No overall effect appears in the experiment.

H2: For both conditions, the outcome being first observed by the percipient and the agent, the number of hits is equal to MCE. The chi-square test gives: Chi-square = 0, indicating no difference between observer conditions.

H3: The first percipients in the pair of trials obtain 6 hits out of 20 trials, whereas the second percipients obtain 4 hits out of 20 trials. The chi-square test gives: Chi-square = .13, a figure which is not significant.

Therefore, none of the tests of the three hypotheses causes the

null-hypothesis, that there are no extra-chance effects, to be rejected.

DISCUSSION OF EXPERIMENTAL RESULTS

The absence in the present experiment of any effect indicating a paranormal influence, does not allow a conclusion to be drawn about the hypothesized differential effects. With regard to the overall effect, we must conclude that in this case the ganzfeld has not produced the expected success. The percipient-order effect which has stood up in previous experiments, turns up in the present experiment in the direction opposite to expectation, although to a nonsignificant degree, the first percipients obtaining two more hits than the second percipients.

TABLE 2
Average rank of the target picture obtained under different conditions of first observer and percipient order

First observer: Percip.order:	percipient	agent	total
1st percipient	2.2	2.1	2.15
2nd percipient	2.6	3.0	2.8
Total	2.4	2.55	2.475

It might be argued that using direct (P = 1/4) hits is an insensitive method for obtaining results. To explore this, we calculated the average rank assigned by the percipient to the target picture. The percipient assigned ranks 1, 2, 3 and 4 to the four

pictures, 1 being assigned to the designated target and no ties being allowed. As can be seen in Table 2, no different conclusion is suggested than on the basis of the evaluation of direct hits.

Finally, the present results do not indicate a preference for either the "transmission model" or the "observational model".

DISCUSSION OF THEORY

The design of the present experiment may serve to show that the procedure to test the two models against each other, is quite simple. More importantly, it does not affect the conditions of the ganzfeld experiment from a classical point of view, so there is no disadvantage of the observation procedure except for the small effort involved.

What can be learned by it, had there been some effect? In the case of an overall effect without an effect due to the observer of the outcome, this would have been unlikely for observational theory, as we have argued. On the other hand, a differential effect associated with who observes the outcome, is difficult to reconcile with a classical theoretical framework. In both cases, no doubt counter-explanations would be put forward. Just the strain such counter-explanations form for the theoretical frameworks they try to defend, eventually causes these frameworks to be abandoned.

Had we concluded at this point, we would have offered an interesting experimental test, but, disappointingly, in the present experiment without a decisive result. However, further advances in parapsychology will be brought about by two factors, of which one cannot go without the other, at least in the long run. One is obtaining interesting experimental results that vindicate the existence of particular unexpected phenomena. The other is the development of at least one theoretical framework that gives an explanation of, why paranormal phenomena are at all possible, or, under what conditions human or animal subjects produce such phenomena, or, what limitations can be set for the phenomena.

We return here to an issue of theoretical interest that was raised in the introduction: An observational model of a GESP experiment, of which the ganzfeld ESP experiment is an example, entails the specification of what is to be considered the "random event" to which

the observer applies his influence to affect the probabilities of the possible outcomes. In theoretical work of one of us (Houtkooper, 1983, chapter 7) it has appeared necessary to distinguish variables that possess (quantum mechanical) "randomness", from variables that do not, or that are not essential to the outcome of interest of the particular experiment. This distinction is required in order to provide a consistent concept of "possible outcomes".

In the observational model of the ganzfeld experiment we can distinguish three random variables, one describing the designated target, the second describing the guess of the percipient and the third, describing the correspondence between target and guess, whether it is a hit or a miss. It is evident that to an omniscient being the third variable is superfluous, as it is completely known if you know the first two. However, from our viewpoint outside the realm of omniscient beings, what we do not know, we may imagine to exist, but its existence still is indefinite. This may seem to be a subtle point, but it is crucial to the model employed in the present experiment.

Let us have a look at the situation after the agent has looked carefully at the target picture (and "sent" it) and after the percipient has made his guess and communicated it to the experimenter in the same room with him. To anybody with a normal sense of what is real, both guess and target are real and so is therefore the fact whether a hit has been obtained or not.

What is assumed by the observational model is in plain contradiction with this: The target may be real to both the agent and the experimenter present with him, but the guess is not. The guess may be real to the percipient and the other experimenter present with him, but the target is not. Therefore, the variable describing whether a hit or a miss has been obtained, is in a state of "indefiniteness", from either viewpoint. (And we do not entertain the viewpoint of the hypothetical omniscient observer!) This indefiniteness of the random variable (as distinct from lack of information about it) allows the observer of the outcome to have an influence on its eventual value, that is, on the outcome being a hit or a miss. It is exactly this influence that is tested in the present experiment.

We do not regard the picture we have just drawn, as one that is appealing on first sight. However we do consider the above rendering of the state of affairs in our experiment to possess a fair consistency with previous work and to be the only consistent way to

unify PK and different forms of ESP, of the GESP (telepathy) type in this case. To unify different paranormal phenomena has been the claim of Schmidt´s (1975) model, but pitiful little effort has been spent on taking advantage of his unification theorem, although it should rank highly with the wielders of Occam´s razor. In this paper, we hope to rekindle interest in the unification of paranormal phenomena and the fruitful theoretical development that it might engender.

For those, who hesitate to take our speculation seriously, we would like to draw attention to what two eminent theoretical physicists, Aharonov and Albert (1980) say: "... it is not possible to define the quantum state of a system [because] no consistent description of how the state changes as the result of a measurement can be developed." The concept of reality rendered by theoretical physics is showing some puzzling aspects at variance with everyday experience. Very few would doubt the statement that reality is at a certain instant of time in a certain state. What Aharonov and Albert say is essentially: it is not.

We hope that further ganzfeld experiments along lines similar to ours may help to shed light upon these fundamental issues, which are leading us to believe that the nature of paranormal phenomena is still stranger than what we can think of at present.

ABSTRACT

The ganzfeld condition appears to be successful in facilitating ESP, despite criticisms. To explain this we put forward a novel model, based on observational theory, in which the observer of the correspondence between target and call is assumed to play a crucial role.

A ganzfeld experiment was designed to test this explanation against the classical explanation based on a transmission model of ESP. The earlier found percipient order effect was also tested. The results gave no indication of ESP, so that none of the hypotheses was corroborated. The hypothesized model of ganzfeld ESP is discussed with regard to its significance for observational theory.

ACKNOWLEDGEMENTS

Thanks are due to Haukur Hjaltason, who acted as co-experimenter throughout the experiment, except for one pair of trials, for which we acknowledge the help of Johanna M. Thorlacius in taking his role.

REFERENCES

Aharonov, Y., Albert, D.Z. ´States and observables in relativistic quantum field theories´, Physical Review D, 1980, 21, 3316-3324.

Blackmore, S. ´A report of a visit to Carl Sargent´s laboratory´, J. S.P.R., 1987, 54, 186-198.

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.

Haraldsson, E., Gissurarson, L.R. ´Perceptual defensiveness, ganzfeld and the percipient-order effect: Two experiments´, E.J.P., 1985, 6, 1-17.

Harley, T., Matthews, G. ´Cheating, psi, and the appliance of science: a reply to Blackmore´, J. S.P.R., 1987, 54, 199-207.

Hartwell, J.W. ´A bound for the observational theories of psi´, E.J.P., 1977, 2, 1, 19-28.

Honorton, C. ´Psi and internal attention states: Information retrieval in the ganzfeld´. In: B. Shapin, L. Coly (Eds.) Psi and states of awareness. New York: Parapsychology Foundation Inc., 1978.

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

Houtkooper, J.M. ´Observational theory: A research programme for paranormal phenomena´. Lisse, The Netherlands: Swets & Zeitlinger, 1983.

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

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

Millar, B. ´The observational theories: A primer´, E.J.P., 1978, 2, 3, 304-332.

Parker, A., Wiklund, N. ´The ganzfeld experiments: towards an assessment´, J. S.P.R., 1987, 54, 261-265.

Sargent, C.L. ´Exploring psi in the ganzfeld´. Parapsychological monographs, 17. New York: Parapsychology Foundation Inc., 1980.

Sargent, C.L. ´Sceptical fairy tales from Bristol´, J. S.P.R., 1987, 54, 208-218.

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

Schmidt, H. ´PK effect on pre-recorded targets´, J. A.S.P.R., 1976, 70, 267-291.

Schmidt, H. ´Collapse of the state vector and psychokinetic effect´, Foundations of Physics, 1982, 12, 565-581.

Schouten, S.A. ´An overview of details of published ganzfeld studies´. Research Letter, 1981, 11, Parapsychology Laboratory, University of Utrecht.

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

Walker, E.H. ´Foundations of paraphysical and parapsychological phenomena´. In: L. Oteri (Ed.) Quantum physics and parapsychology. New York: Parapsychology Foundation Inc., 1975 (pp.1-44)

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.

Joop M. Houtkooper
Houtkooper Software Engineering
De Maalkom 7
1191 LP Ouderkerk aan den Amstel
The Netherlands

Loftur R. Gissurarson
The Koestler Chair of Parapsychology
Department of Psychology
University of Edinburgh
7 George Square
EH8 9JZ Edinburgh
Scotland

Erlendur Haraldsson
Department of Psychology
University of Iceland
101 Reykjavik
Iceland

EUR.J.PARAPSYCHOL., V.7, 1988-1989

A POSSIBLE 'DIRECTIVE' ROLE OF THE AGENT IN THE GANZFELD

Julie Milton
Psychology Department
University of Edinburgh

INTRODUCTION

The Ganzfeld may be a particularly appropriate setting to examine the role of the agent in ESP. Honorton and Harper (1974) expected the Ganzfeld to be psi-conducive because, among other reasons, exposure to the Ganzfeld might help to establish an emotional link between percipient and agent, increasing the percipient's desire for communication: Bertini, Lewis and Witkin (1972) in their non-parapsychological study of the Ganzfeld, had noted that some subjects developed during the session a preoccupation with the experimenter, suggestive of a process akin to transference.

However, very few Ganzfeld studies have compared ESP performance with and without an agent. Both Raburn and Manning (1977) and Sargent, Milton, Payne and Bennet (1982) found scoring without an agent to be close to chance and significantly less ($p<.05$ in both cases) than scoring in a GESP condition. A number of experimenters have suggested that there may be a relationship between the presence of an agent and the occurrence of within-trial displacement in a free-response setting.

Both Stanford and Neylon (1975) and Rogo (1979) have suggested that the presence of an agent may help to prevent the occurrence of displacement. Stanford and Neylon, noting that some displacement seemed to have occurred in their Ganzfeld clairvoyance study, suggested that in a GESP procedure the presence of an agent, with whom the percipient might develop a transference-like relationship due to the Ganzfeld stimulation, might serve to focus the percipient's attention on the target. Rogo, on the basis of some apparent cases of displacement in his own Ganzfeld studies, has also suggested that the agent serves a directive purpose, and that percipients may be more likely to displace without an agent. Displacement was suspected to have occurred in the clairvoyance, but not the GESP condition in a Ganzfeld study by Sargent, Milton, Payne and Bennet (1982), which would tend to support the suggestions of Stanford and Neylon and of Rogo. However, an interesting aspect of Rogo's account of the work in which displacement seemed to occur is that on some trials, displacement occurred to some experimental target from another trial about which the agent claimed to have been thinking at the time. Any directive function which the agent might fulfill could therefore involve the agent's attention as a guide, rather than his presence as such; thus, an agent may help to prevent displacement if he or she only attends to the target, but may contribute to the occurrence of displacement if he or she also attends to the controls. It should be noted that in all of these studies no statistical analyses were conducted for displacement, and such evidence as there was for displacement solely consists of the authors' judgment.

In the present study, the main hypothesis was that the presence of an agent attending to the package containing the target should result in higher scoring than when no agent was present. It was not felt desirable to have the agent view the target in the former situation because this would add the possibly confounding effects of the agent acting as a channel for information about the target, rather than merely acting in a directive capacity. A secondary hypothesis was that, if all of the pictures in the judging set were with the agent during the trial, displacement would be more likely to occur than if the target only was with the agent, because the control pictures would also be attended to by the agent during the trial.

In order to test these hypotheses, a within-subjects design was used with three conditions; although an agent was present in two of the conditions, he or she never saw the target, which was always enclosed in a sealed opaque envelope. In the first condition, an agent remained

with the target in an isolated room during the Ganzfeld session; in the second condition, there was no agent, and the target remained in an isolated room during the session; and in the third condition, the agent remained in an isolated room with all of the pictures in the judging set, both target and control pictures. The percipients and the experimenter were always blind as to which condition was operating. It was predicted that overall scoring would be higher in condition (1) than in condition (2), and that displacement might occur in condition (3), but not in conditions (1) or (2).

The possible influence of the agent's psychological state on performance was examined by having agents complete at the end of each session for which they were present a questionnaire measuring their mood, motivation, the degree to which they had found the session interesting, and the percentage of the session for which they had been able to concentrate on the package containing the target. Each questionnaire measure was correlated with the rank assigned to the target on each trial, but no predictions were made concerning the outcome of this analysis.

To allow for the possibility that scoring might manifest itself on a control picture as well as on the intended target, the correspondence rating given to the picture most similar to the mentation on each trial was also correlated with the agent state measures, in order to see if the strength of association to the target which may have been the object of the percipients' responses correlated with the agent's state. Again, no predictions were made as to the outcome of these analyses.

Correlations between various measures of the percipient's attitude towards each picture in the judging set, and the correspondence ratings assigned to each picture were examined, in order to investigate whether the percipient's attitude towards the pictures in the set might determine to which picture he or she would displace, if displacement occurred.

Finally, some predictions were made concerning scores on Ganzfeld mentation which the percipients reported as being unrelated to the previous image; not recognisable as being related to the percipient's memories; and surprising. Following a suggestion from Stanford's (1978, 1979) conformance behaviour model, it was predicted that the percipients would score better on unrelated than related images; following the work of Sargent, Bartlett and Moss (1982), Sargent, Moss

and Bartlett (1982) and Sargent, Milton, Payne and Bennet (1982) (see Milton, in press, for a review), it was predicted that scoring would be better on surprising than unsurprising imagery, and on novel than memory-related imagery. All of the predictions and planned analyses are given in full detail in the results section.

For all the analyses, the ratings and rankings assigned by two independent judges to the correspondence between the subjects' mentation reports and judging sets for each trial were used, rather than those of the subjects.

METHOD

Design

A within-subjects design was used with three experimental conditions (one condition in each of three trials) and a pretest trial to familiarise subjects with the experimental procedures. A double-blind operated so that the agent did not learn of the outcome of any of the percipient's trials until the end of his or her part in the experiment, and so that neither experimenter nor percipient knew which condition was underway in any trial. The order in which the percipients did the three conditions was counterbalanced to avoid order effects, and the percipients were pseudo-randomly assigned to one of the six possible orders by a person (note 2) otherwise independent of the experiment, using a method specified by the author, closely similar to that used for target and target set selection described below.

Subjects

Twelve percipients (8 male, 4 female, aged 19-35) took part. All were friends or acquaintances of the author, and mostly students. Eleven of the percipients chose to act as agents when the person who had acted as their agent took part as a percipient. No-one acted as agent for more than one percipient. Percipients who did not wish to bring a friend to act as agent were offered a choice of agents, and were introduced to them in a social setting by the author before the

experiment. All subjects were at least open to the existence of ESP. Only two had had previous Ganzfeld experience. All subjects were blind to the purposes of the study, although percipients were told of the different clairvoyance procedures to be used on each trial. The phenomenon of displacement was not mentioned to the subjects.

Targets and Target Selection

Twenty-two target sets, each containing four pictures, were used in the study. The pictures were black and white and colour postcards of paintings, cartoons, illustrations and photographs, and the sets were composed so that the pictures in them were as contrasting as possible in terms of style and content. No pictures with a negative emotional tone were included, both for reasons of ethics and the practical reason that the use of unpleasant targets might result in avoidance; in addition, pictures with erotic content were not included, since it was felt that percipients might be embarrassed about reporting erotic imagery if target-related imagery occurred.

Duplicate target sets were made for the agents to use. Each picture in the agent's set was contained in a sealed, opaque envelope, while all four of the pictures in the percipient's set were contained in a single envelope. Each set was assigned a set number, from 1 to 22. Each picture in the agent's set was identified by a letter A, B, C or D; the corresponding pictures in the percipient's set were identified by a different, randomly-applied code unknown to either percipient or agent, so that there would be no possibility of the agent (who had to be unsupervised because there were no other personnel available) cheating by, for example, thumping the laboratory wall three times for picture C.

The target set and target to be used for each trial were selected according to a procedure specified by the author by a person (note 2) otherwise unconnected with the experiment, who used numerical codes to translate the outcome of a series of coin-flips into an entry-point into a random-number table (Rand Corporation, 1955), and other codes to translate the ensuing sequence of random numbers into numbers and letters identifying target sets and targets such that the selection of each set or picture was equiprobable (e.g. for the selection of target pictures within the set, the digits 1 and 2 indicated picture A, 3 and 4 picture B, 5 and 6 picture C, 7 and 8 picture D, with the digits 0

and 9 being ignored). A constraint was applied to this otherwise random procedure such that no percipient ever saw the same set of pictures twice. Having made these selections, the randomiser enclosed a slip of paper identifying the target set and another identifying the target each into a separate opaque envelope bearing the name and trial number (1, 2, or 3) for each percipient, and sealed the envelopes. All target sets and designations were kept locked away by the experimenter until each trial was ready to begin.

Setting and Apparatus

The experiment was conducted using three rooms in the Psychology Department at Edinburgh University. The Parapsychology Laboratory, a large office, was used as the Ganzfeld stimulation room, and contained a comfortable reclining chair on which the percipient lay during the Ganzfeld stimulation. On a table next to the chair was a flexi-poise lamp fitted with a 60 watt red bulb. A tape-recorder, which relayed white noise from a tape through comfortable headphones to the percipient, was also on the table, as was a sensitive microphone which relayed the subject's verbalisations to a second tape-recorder, contained in a cubicle in the corner of the laboratory, and separated from the laboratory by a door. The subject's verbal report was taped on this recorder, as well as being concurrently relayed over headphones to the experimenter (the author), who was also in the cubicle. The author's office, which was separated from the laboratory by another room, contained the target set, target and condition designations in their sealed envelopes. A third room, three floors down from the laboratory, served as the agent's room during the sending period.

Procedure

At the beginning of each trial, the experimenter met the percipient and agent in the laboratory and offered them refreshments. On the first experimental trial, the experimenter explained to both subjects that all three sessions involved a clairvoyance procedure, with minor variations, and stressed that the percipient and agent should not discuss any aspect of any trial, until all were over. Both practice and experimental trials followed the same basic procedure, except for

full feedback being given to both parties after the practice trial which was conducted to familiarise both subjects with the situation.

When the percipient and agent seemed to be at ease and ready to begin, the experimenter escorted the agent to her office and showed him or her the sealed envelopes containing the target set, target, and condition designation for the trial, and the pile of sealed envelopes in a box containing the target pool of pictures, each identifiable by a number and letter on the envelope. Starting two stopwatches simultaneously, the experimenter handed one to the agent and asked him or her to open the envelope labelled 'CONDITION' and to follow the instructions it contained. She then left the office before the agent opened the envelope. For condition (1), in which the agent was to attend to the envelope containing the target picture, the agent was instructed how to obtain the target envelope, and to go to the agent's room downstairs. He or she was to wait in the agent's room until the stopwatch read 15 minutes, and then to place the envelope on the table and to sit and concentrate, as far as possible, upon the envelope for half an hour. After this time, he or she was to complete a questionnaire provided and to leave the building. In condition (2), in which the target was to remain in the room without the agent, the agent was instructed to place the envelope on the table in the agent's room and to leave the building at once, before the stopwatch read 15 minutes. In condition (3), the agent was instructed to take all of the pictures in the target set to the agent's room, and to place all but the target envelope behind the chair out of sight and to concentrate upon the target envelope as in condition (1). After half an hour, the agent was to complete the questionnaire, return the control pictures to the experimenter's office (so that she could not tell by their absence that that trial belonged to condition (3)) and leave the building. In each condition, the agent was also instructed to place the instructions, questionnaire, etc. in an envelope and to seal it, to leave the stopwatch running, and to leave these things and the envelope containing the target picture on the table in the agent's room to make all conditions seem alike to the experimenter.

The absence of anyone to act as agent-experimenter raises the question of whether it would be possible to cheat. The only practicable method would have been for percipient and agent to devise some sort of content-related code before the trial, such as two knocks (on the walls or heating pipes, for example) indicating the presence of people in the target, and so on. However, in order to be heard by the percipient during the trial, any raps or knocks would have had to

have been loud enough for the percipient to hear them over the white noise, in which case the experimenter would have been unlikely not to have noticed them also; the experimenter would have been able to hear (as pre-tested) if the percipient had lifted either headphone earpiece during the session, or if the percipient had attempted to remove the ping-pong balls, or to leave the reclining chair. Given that the percipients were mostly friends of the experimenter, with no ostensible investment in cheating, it is arguable that with these subjects, cheating was unlikely.

Meanwhile, after leaving the agent in her office, the experimenter returned to the laboratory and locked the door. The percipient was seated in a reclining chair, with the visual Ganzfeld provided by halved ping-pong balls. The percipient adjusted the position of the red light to his or her preference (between 1 and 2 feet from the face), and the experimenter switched on the white noise tape, which had been set at a comfortable volume by the percipient.

Having switched on the tape, the experimenter immediately retired to her cubicle, started the tape which recorded the percipient's verbal mentation report, donned the headphones so that she could hear the report, and transcribed the subject's mentation as he or she reported it; the percipient remained in the Ganzfeld for thirty minutes, and had been asked to report aloud all of his or her thoughts, imagery, feelings and sensations. The end of the Ganzfeld period was signalled to the subject by the fading out of the white noise on the tape, followed by J. S. Bach's 'Air on a G String'. When the music had finished, the percipient extracted him or herself from the Ganzfeld, and was again offered refreshments.

The experimenter now gave the percipient the mentation transcript, and asked him or her to indicate the divisions between images which occurred separately, in the sense that they were separated by a period of time during which no ideation occurred, or by a change in content. The percipient was also asked to indicate which images were surprising, which novel and which unrelated to the previous image. The percipients were also encouraged to point out which images were noteworthy for some other reason, but none did.

The experimenter then went to her office, and picked up the judging set of pictures (specified by a slip of paper left on the desk in the office by the agent, according to instructions.) Returning to the laboratory, she showed the percipient the four pictures which made up

the set, and asked him or her to complete a short questionnaire concerning the pictures. The percipient was asked to rate his or her liking for, interest in, familiarity with, and feelings of personal significance towards each picture, and to place the four in rank order on the basis of which the percipient would prefer to look at for ten minutes. In case the percipient suspected that the measures might be expected to affect the judging process, the experimenter told him or her that the data were being collected as part of an independent study of the characteristics of the target pool.

The percipient then went through the mentation transcript, item by item, rating each picture for its correspondence to each image on a scale of 0 to 5, and recording the ratings on a form. The percipient was instructed to be alert for correspondences which might be literal, formal, symbolic, emotional, or associative. These instructions were not extensive because only independent judging data were to be analysed, and the percipients' judging was conducted mainly for their own interest. When the judging of the whole transcript was completed, the total number of correspondence rating points assigned to each picture was summed, and the sums used as rough but not binding guides to rank the pictures in order of their correspondence to the mentation (the picture corresponding best to the mentation being ranked first, down to the worst being ranked fourth), and to give each picture a rating on a scale of 0 to 5 for its overall correspondence. Subjects were allowed to tie ratings, but not rankings (this rating procedure has been used successfully in a number of experiments by Sargent (see, e.g., Sargent, 1980).

The experimenter now went downstairs to the agent's room, retrieved the experimental materials, and returned to the laboratory, where she gave the sealed target envelope to the percipient, who opened it to see which picture had been the target. The experimenter then answered any questions the percipient might have concerning the trial, and the session ended.

At the end of the experiment, the transcript for each trial was sent to the two independent judges, as well as the target set used on the trial by the percipients, and a control set never seen by the percipient. The control set was randomly selected for each trial by a person independent of the experiment, after the experiment was over, by means of a pre-specified procedure using random number tables. The judges were asked to rate each mentation item on a 0-10 scale for its correspondence to each picture in the two sets provided for that

trial, judging one set at a time (the order of judging the two sets had been pseudo-randomly counterbalanced across trials), and to use the point sums for each picture as a guide in assigning to each picture a 0-10 rating of overall correspondence, also placing the pictures in rank order of correspondence to the mentation report. Both judges had had previous judging experience, but were reminded to look out for all possible types of target-mentation correspondence such as literal, symbolic, shape, emotional, and thematic. When assigning ratings, they were instructed to consider how close the correspondence was, the complexity of each picture (because there are more chance opportunities of correspondence with a complex than with a simple picture), and the frequency of the mentation item's content in Ganzfeld mentation transcripts, giving higher ratings to unusual ones that were correct.

RESULTS

All of the analyses were planned before the experiment began, except where otherwise stated. Data from one trial in condition (2) was not included in the analysis, because the agent had not followed the instructions to leave the building before the trial began, but instead remained during the trial; the agent made a note on his questionnaire that he had been confused about the instructions, but the experimenter did not examine the questionnaires until the experiment was over, by which time the percipient had left the country, so it was not possible to conduct a re-trial.

According to a post-hoc analysis comparing overall scoring to chance, using the sum-of-ranks analysis of Solfvin et al. scoring was significantly above chance for Judge 2 ($z=1.97$, $p(\text{one-tailed})<0.025$) and non-significantly above chance for Judge 1 ($z=1.06$, $p(\text{one-tailed})<0.10$).

Comparison of scoring between conditions (1) and (2)

The tables below show the distribution of ranks assigned to the target in each of the three conditions, for each independent judge: Contrary to prediction, scoring in condition (1) was not significantly higher than scoring in condition (2), according to the sum-of-ranks

TABLE 1
Target rank distribution for Judges

Judge 1	RANK				SUM OF RANKS	MCE
	1	2	3	4		
CONDITION (1)	5	2	2	3	27	30
CONDITION (2)	2	2	6	1	28	27.5
CONDITION (3)	6	2	1	3	25	30
TOTAL	13	6	9	7	80	87.5

Judge 2	RANK				SUM OF RANKS	MCE
	1	2	3	4		
CONDITION (1)	5	3	2	2	25	30
CONDITION (2)	4	3	3	1	23	27.5
CONDITION (3)	5	2	3	2	26	30
TOTAL	14	8	8	5	74	87.5

test of Solfvin et al. (1978), although the very slight difference between the scores was in the predicted direction for both judges.

Occurrence of displacement in the three conditions

It was predicted that displacement might occur in condition (3), but not in conditions (1) and (2). As a test for displacement within each

condition, the average overall correspondence ratings assigned by each independent judge to the control pictures in the target set were compared to the average ratings assigned to the four pictures in the control set never seen by the percipients, following the analysis of Child and Levi (1980); if displacement had occurred to the control pictures in the target set, their average correspondence rating should have been higher than that of the pictures in the control set.

For both judges in all three conditions, the mean ratings assigned to target set control pictures and to control set pictures were very close; only for Judge 2 in condition (3) were control set picture ratings lower than target set control picture ratings, as would be expected if displacement had occurred, but not significantly so as tested using the related t-test to compare the two average ratings on each trial ($t=0.537$, 11df, $p(\text{one-tailed})>0.10$).

In case the judges had tended to assign low overall ratings to pictures which had not been ranked first (to make an apparent 'hit' more spectacular), hence tending to lower the average target set control picture rating and therefore disguise any displacement, a post-hoc analysis using instead the sum of item-by-item correspondence rating points for each picture was performed. Using these data, point sums were non-significantly higher for target set control pictures than for control set pictures in condition (3) for both judges ($t=0.222$ for Judge 1, $t=1.175$ for Judge 2; 11df and $p(\text{one-tailed})>0.10$ in both cases), and in condition (2) for Judge 2 ($t=0.548$, 10df, $p(\text{one-tailed})>0.10$).

Rogo (1979) noted that at the beginning of the Ganzfeld session, the percipient seemed to describe the target picture and then went on to displace to the control pictures in the target set. Stanford and Neylon (1975) suggested that displacement might occur if the percipient felt that the session was going on too long and began to look ahead to the judging period, when he or she would see all of the target set; this situation might also result in psi-hitting at the beginning of the session, turning into displacement towards the end. After the experiment, it was decided to see if a change from psi-hitting to lower scoring on the target (as would result if, among other things, displacement occurred) was evident during the course of the session, by comparing the percentage of correspondence rating points allocated to the target in the first and second halves of each session, using the Wilcoxon Test. The sessions were divided into halves containing equal numbers of mentation items; if there was an

odd number of items, the item which divided the two halves was omitted for the purpose of analysis. Only in conditions (2) and (3) for Judge 1 was scoring higher in the first than in the second half of the trial, with neither difference being significant. The largest difference was on condition (2) ($N=11$, $T=17$, $0.10>p(\text{one-tailed})>0.05$), but the direction of this difference was reversed for Judge 2. The difference in condition (3) did not approach significance ($N=12$, $T=33$, $p(\text{one-tailed})>0.10$).

Agent questionnaire measures

Spearman correlation coefficients were calculated for the relationship between scores on the four agent questionnaire measures and the rank assigned to the target by each independent judge. The two conditions for which an agent had been present, conditions (1) and (3), were analysed separately because, if displacement had occurred in condition (3), combination of the two conditions could have obscured any effect on the target in condition (1).

The results are summarised in the table below; two agents in each condition failed to complete a questionnaire, and so $N=10$ in both conditions. 100-point line scales were used for all 4 questions, with end points labelled.
Only the negative correlation between concentration and target rank approached significance, for Judge 2 in condition (1), indicating a tendency for high scoring on the target to be associated with a high degree of concentration.

To allow for the possibility that scoring might manifest itself as displaced scoring, it was also planned to calculate Pearson correlations of the agent questionnaire measures with the correspondence rating of the highest-ranked picture on each trial for both judges. The results are shown below:
Two results were significant at the 0.05 level, and both in condition (3). High motivation correlated significantly positively with the highest rating for Judge 1, and bad mood correlated significantly positively with the highest rating for Judge 2.

The characteristics of the data are summarised in table 4.

TABLE 2
Correlations between agent questionnaire measures and target ranks in conditions (1) and (3)

CONDITION (1)	JUDGE 1		JUDGE 2	
	RHO	P(2-T)	RHO	P(2-T)
BAD MOOD	0.25	>0.10	0.05	>0.10
HIGH MOTIVATION	-0.07	>0.10	0.04	>0.10
BOREDOM	-0.34	>0.10	-0.33	>0.10
CONCENTRATION	0.42	>0.10	0.62	0.10>P>0.05

CONDITION (3)	JUDGE 1		JUDGE 2	
	RHO	P(2-T)	RHO	P(2-T)
BAD MOOD	0.01	>0.10	0.08	>0.10
HIGH MOTIVATION	-0.28	>0.10	0.36	>0.10
BOREDOM	-0.08	>0.10	-0.15	>0.10
CONCENTRATION	-0.18	>0.10	-0.26	>0.10

Picture preference measures

Pearson correlations of the first four measures of the percipient´s attitude towards the pictures with each judge´s overall rating of correspondence to each picture were obtained for each independent judge, in order to test whether the percipient´s attitude towards a picture might determine to which, if any, of the pictures he or she

TABLE 3
Correlations between agent questionnaire measures and highest correspondence rating for conditions (1) and (3)

CONDITION (1)	JUDGE 1		JUDGE 2	
	RHO	P(2-T)	RHO	P(2-T)
BAD MOOD	0.14	>0.10	0.11	>0.10
HIGH MOTIVATION	0.42	>0.10	0.12	>0.10
BOREDOM	0.29	>0.10	0.22	>0.10
CONCENTRATION	-0.35	>0.10	-0.47	>0.10

CONDITION (3)	JUDGE 1		JUDGE 2	
	RHO	P(2-T)	RHO	P(2-T)
BAD MOOD	0.29	>0.10	0.59	<0.05
HIGH MOTIVATION	0.61	<0.05	0.43	>0.10
BOREDOM	-0.27	>0.10	-0.37	>0.10
CONCENTRATION	-0.02	>0.10	-0.36	>0.10

would displace. Each picture was rated on a 4 or 5 point scale, with each point labelled. None of the four measures correlated significantly with correspondence ratings; however, the standard deviations of all of the attitude measures were quite low, particularly the measures of how boring and how familiar the pictures were to the percipients.

It had been planned to correlate the rank assigned to each picture

TABLE 4
Descriptive statistics for conditions (1) and (3)

CONDITION (1)

	N	MEAN	S.D.
BAD MOOD	10	28.90	21.93
HIGH MOTIVATION	10	69.10	7.00
BOREDOM	10	39.40	16.83
CONCENTRATION	10	52.10	16.31
JUDGE 1 TARGET RANKS	10	1.90	1.10
JUDGE 2 TARGET RANKS	10	2.00	1.05
JUDGE 1 HIGHEST RATING	10	5.30	1.77
JUDGE 2 HIGHEST RATING	10	3.50	1.58

CONDITION (3)

	N	MEAN	S.D.
BAD MOOD	10	43.60	25.55
HIGH MOTIVATION	10	71.10	15.02
BOREDOM	10	40.40	19.34
CONCENTRATION	10	56.20	26.98
JUDGE 1 TARGET RANKS	10	1.80	1.23
JUDGE 2 TARGET RANKS	10	1.80	0.92
JUDGE 1 HIGHEST RATING	10	6.00	1.63
JUDGE 2 HIGHEST RATING	10	3.80	1.40

on the basis of which the percipients would prefer to look at for ten minutes with the correspondence rating assigned to the picture by the judges, but it was realised at the end of the experiment that such an analysis would be invalid because the assignment of ranks to the pictures within each trial would mean that pictures would not be comparable between trials on such a measure, as would be necessary for such a correlation. Therefore, it was decided to instead calculate a value of Kendall's tau (a measure of concordance) for each trial for the relationship between the percipients' liking ranks and the judges' correspondence ranks assigned to each picture in the set, and to compare the number of positive and negative tau values (equiprobable by chance) using the binomial test. The analysis yielded 16 positive, and 11 negative tau values for the relationship between the percipients' liking ranks and the mean of the two independent judges' correspondence ranks, showing a slight but by no means significant tendency for liking to be related positively to correspondence ($p(\text{two-tailed})=0.4$).

One of the reasons for using the independent judges' data rather than that of the percipients for the analyses reported, was the suspicion that the percipients might allow themselves to be swayed in their judgement of correspondence by their liking for the pictures. A post-hoc analysis was carried out to see if this had, in fact, been the case. Calculating tau values for the relationship between the percipients' liking and correspondence ranks, 23 positive values were obtained, a number significantly greater than the 4 negative tau values obtained ($p(1\text{-tailed})<0.0002$), indicating a tendency for liking and correspondence ranks to be similar.

Mentation categories

The proportion of all item-by-item correspondence rating points allocated to the target on the basis of each mentation category (MCE=25%) was compared to the proportion allocated to the target on the basis of the remaining mentation for each trial, using the Wilcoxon Test (note 3). The results for Judge 1 are summarised in table 5 below (the data of Judge 2 were not usable since he had not clearly identified the mentation items for which the points were assigned).

According to prediction, percipients scored significantly better on

TABLE 5
Percentage of points allocated to target for Judge 1

Novel mentation:	19.6%	Remainder:	28.5%
Surprising mentation:	35.2%	Remainder:	26.3%
Unrelated mentation:	26.7%	Remainder:	27.2%

surprising mentation than on the remainder (N=21, T=57, p(one-tailed)<0.025). However, contrary to prediction, percipients scored worse on novel than on memory-related imagery, the direction of the difference being due to below-chance scoring on novel imagery (5.4% below chance) and above-chance scoring on memory-based imagery (3.5% above chance). The difference was not significant. Also contrary to prediction, scoring was slightly worse on unrelated than on related mentation items, although scoring on both mentation types was above chance.

DISCUSSION

Although overall scoring tended to be above chance (significantly so by a post-hoc analysis of the data of Judge 2, p(one-tailed)<0.025, although the significance of this result could have been due to overanalysis), scoring with an agent was only slightly higher than without an agent, thus giving no real support to the hypothesis that the presence of an agent should yield higher scores. No significant evidence of displacement was found in any of the three conditions, although in the planned displacement analysis, only the results in condition (3), for Judge 2, were in a direction consistent with the occurrence of displacement. In a post hoc analysis which was hoped would be more sensitive to the occurrence of displacement, both judges' data yielded results in a direction consistent with the occurrence of displacement in condition (3), although this was also the case for condition (2) for Judge 2. Again, however, none of the

results approached significance. A post-hoc analysis which compared scoring in the first and second halves of each trial, following suggestions that psi-hitting might turn into displacement later in the trial if it occurred, found non-significant declines in scoring on the target in conditions (2) and (3) for Judge 1. Although agent questionnaire measures failed uniformly to yield significant correlations with the ranks assigned to the target by each judge, two correlations between the highest correspondence rating assigned on each trial and the agent questionnaire measures were significant in condition (3); motivation correlated positively with scoring for Judge 1 ($r=0.61$, $N=10$, $p(\text{two-tailed})<0.05$), while bad mood correlated positively with scoring for Judge 2 ($r=0.59$, $N=10$, $p(\text{two-tailed})<0.05$). Since correlations with the highest rating on each trial had been examined to allow for the possibility of displaced scoring, the occurrence of significant correlations with the highest rating but not with the target rank on each trial, and only in condition (3), might be taken to add to the weak tendency for results directionally consistent with the occurrence of displacement to appear in condition (3), where displacement had been predicted; however, such a tendency is very far from being conclusive, and at best suggests a need for replication with a larger number of trials in each condition.

None of the measures of the degree to which each picture was liked by the percipient, or seemed boring, familiar, or personally significant, correlated significantly with its resemblance to the mentation, apparently indicating that the percipient's attitude to each picture in the set did not determine to which picture he or she might have displaced, although in the absence of any evidence for any displacement in the study, this hypothesis may not have received a fair test. In addition, the variance of the scores on all four of the attitude measures were quite low, and a future examination of this question should perhaps involve deliberate manipulation of the characteristics of the pictures in the target set, so that the range of attitude scores would be large enough to reveal any effect.

A more interesting result, perhaps with practical implications, was the highly significant post-hoc finding that the percipients' liking ranks were very similar to their correspondence rankings ($p(\text{one-tailed})<0.0002$), but not to the correspondence rankings of the two independent judges ($P(\text{two-tailed})=0.4$), seeming to indicate that the percipients had allowed themselves to be swayed in their judging of correspondence by their liking for the individual pictures. This result is perhaps another argument in favour of using independent

judges.

Finally, the examination of scoring on the basis of various mentation categories yielded a significant predicted effect, namely, that of higher scoring on surprising than on unsurprising mentation ($p(\text{one-tailed}) < 0.025$). Contrary to prediction, scoring tended to be worse on novel than on memory-related imagery, although the direction of this difference was due to a below-chance deviation on novel imagery which was larger in magnitude than the slightly above-chance deviation on memory-related imagery. Also contrary to prediction, scoring was slightly worse on unrelated than related imagery with both scores being slightly above chance.

In conclusion, a deliberate attempt to induce displacement was largely unsuccessful, and the few analyses that reached significance at the 0.05 level should be treated cautiously because of the number of both planned and post-hoc analyses carried out.

NOTES

1. This study was conducted between 1982 and 1983, and was jointly funded by the Society for Psychical Research and the Perrot-Warwick Studentship in Psychical Research as part of my doctoral research. Earlier versions of this paper were presented at the 1983 Parapsychological Association Convention and in my doctoral thesis. I am grateful to John Beloff, Deborah Delanoy and Robert Morris for comments on earlier versions, and to Gerry Matthews and Trevor Harley for acting as independent judges.

2. I would like to thank John Beloff for performing this function.

3. The use of the Wilcoxon Test replaces a planned analysis in which the proportion of item-by-item correspondence rating points assigned to the target on the basis of a category of mentation by all subjects was to be compared to the proportion of points allocated to the target on the basis of the rest, using the binomial test. The author is grateful to Drs. Ephraim Schechter and Donald McCarthy for pointing out that, since the item-by-item correspondence points are not assigned independently of each other, the use of the binomial test would be unjustified. Thanks are also due to Drs. Jessica Utts, Ephraim Schechter, Donald McCarthy and George Hansen for suggesting

the use of the Wilcoxon Test instead.

ABSTRACT

Twelve percipients took part in 3 Ganzfeld sessions each in a within-subjects comparison of ESP performance with an agent, without an agent, and with an agent in the presence of control pictures as well as the target. Contrary to prediction, scoring in the first condition was not significantly higher than in the second, and there were only indirect indications that predicted displacement effects had occurred in the third condition. As predicted, scoring was significantly higher on surprising than unsurprising mentation, although two other predictions concerning mentation types were not fulfilled. No measures of the percipients' attitude to the pictures in the target set related to target or displaced scoring, although there was very strong evidence that percipients were heavily swayed in their judgement of correspondence by their preference for individual pictures within the judging set.

REFERENCES

Bertini, M., Lewis, H.B. & Witkin, H.A. 'Some preliminary observations with an experimental procedure for the study of hypnagogic and related phenomena'. In C. T. Tart (Ed.), Altered States of Consciousness. Anchor books, N.Y., 1972, p.95-114

Child, I. & Levi, A. 'The use of judges' ratings to test hypotheses about psi processes'. Journal of the American Society for Psychical Research, 1980, 74, p.171-181

Honorton, C. & Harper, S. 'Psi-mediated imagery and ideation in an experimental procedure for regulating perceptual input'. Journal of the American Society for Psychical Research, 1974, 68, p.136-168

Milton, J. 'A survey of free-response judging practices'. Journal of the American Society for Psychical Research, in press

Raburn, L. & Manning, R. ´Sender relaxation and expectation in telepathy´. In J. D. Morris, W. G. Roll & R. L. Morris (Eds.) Research in Parapsychology 1976, Metuchen, N.J., Scarecrow Press, 1977, p.156-158

Rand Corporation ´One Million Random Digits´. Glencoe, Ill., Free Press, 1955

Rogo, D.S. ´The concept of extrasensory "noise"´. Res. Letter Univ. of Utrecht, 1979, 9, p.40-54

Sargent, C.L.S. ´Exploring psi in the ganzfeld´. New York, N.Y., Parapsychology Foundation, 1980

Sargent, C.L.S., Bartlett, H.J. & Moss, S.P. "´Response structure and temporal incline in ganzfeld free-response GESP testing´. Research in Parapsychology 1981, Metuchen, N.J.: Scarecrow Press, 1982, p.79-81

Sargent, C.L.S., Milton, J., Payne, J. & Bennet, S. Unpublished study, 1982

Sargent, C.L.S., Moss, S.P. & Bartlett, H.J. "Unpublished study, 1982

Solfvin, G.F., Kelly, E.F., & Burdick, D.S. ´Some new methods of analysis for preferential-ranking data´, Journal of the American Society for Psychical Research, 1978, 72, p.93-110

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

Stanford, R.G. ´The influence of auditory ganzfeld characteristics upon free-response ESP performance´. Journal of the American Society for Psychical Research, 1979, 73, p.253-272

Stanford, R.G. & Neylon, A. ´Experiential factors related to free-response clairvoyance performance in a sensory uniformity setting (ganzfeld)´. In J. D. Morris, W. G. Roll, and R. L. Morris (Eds.), Research in Parapsychology 1974, Scarecrow Press, Metuchen, N.J., 1975, p.89-93.

Julie Milton
Psychology Department
Edinburgh University
7 George Square
Edinburgh
Scotland, U.K.

EUR.J.PARAPSYCHOL., V.7, 1988-1989

HANDLING AND GENTLING AS FUNCTIONS OF PARANORMAL HEALING WITH NORMAL AND SICK ANIMALS BY NON-HEALERS

F.W.J.J. Snel and P.C. v.d. Sijde

INTRODUCTION

One of the most used arguments against paranormal healing is that suggestion is the major source influencing the healing process and its results. Suggestion possibly plays that major role, as a function of paranormal healing, when humans are treated. It is generally accepted that animals, especially rodents ('lower order vertebrates'), are not susceptible to suggestion in the same sense as human beings are. Animals, however, are susceptible to behavioural stimuli from the environment.

When we objectively analyse the behaviour of a paranormal healer during healing we observe the making of passes and/or the laying-on of hands. There is a striking resemblance between this behaviour and what is known as handling and gentling with animals. Recent results of experimental studies (Van Bergeijk et al, in press; Beynen and Van Tintelen, in press) show that gentling, when compared with handling,

---

Note: All experiments described in this paper with sick animals were conducted as part of ongoing oncological research.

increased body-weight, while handling when compared to a control condition has a negative effect on body-weight. Gentling might be considered as the 'animal' counterpart of 'human' suggestion.

As we argued, suggestion may play an important role in paranormal healing. In a series of experiments with animals, of which the present paper is the first report, we investigated the validity of this argument. If it can be demonstrated that paranormal healers have a significant healing effect on animals then it follows that the assumed effect of suggestion in this type of healing can not be the only explanation. However, in order to demonstrate a possible healing effect on animals we first have to establish a 'base-line', that is the effect of only handling and gentling on animals. Therefore we first designed a series of experiments using non-healers as handlers and gentlers of both healthy and sick animals. This series of experiments is reported in this paper. Then results of future research with professional (paranormally gifted) healers treating healthy and sick animals can be compared with the data of the present study.

Sick animals have to be included in the studies because it can be assumed that only with sick animals real healing effects can be demonstrated. We obtained sick animals for our studies because we were allowed to participate in ongoing oncological studies, in which animals for research purposes are injected with tumours and made sick.

Handling and gentling with rats and mice have been extensively studied during the 1950's (Bechterev, 1949; Hall and Whiteman, 1951; Griffiths and Stringer, 1952; Weininger, 1953; Weininger et al, 1954; Scott, 1955; McClelland, 1956; Ruegamer and Silverman, 1956; Gertz, 1957; Levine and Otis, 1958; Mogenson and Ehrlich, 1958; Knowles, 1959; Levine, 1959; and King and Eleftherion, 1959).

It appears that various psychological and physiological processes can be altered by handling or gentling animals for a prescribed time each day. Conflicting evidence is reported for changes in emotional stability (Hall and Whiteman, 1951; Weininger, 1953; Scott, 1955; Gertz, 1957; Levine, 1959; Cowley and Widdowson, 1965; Williams and Russell, 1972; Hisjarvi and Junnila, 1988); behaviour (Bechterev, 1949; Griffiths and Stringer, 1952; Scott, 1955; Mogenson and Ehrlich, 1958; King and Eleftherion, 1959); learning capacity (Hall and Whiteman, 1951; Griffiths and Stringer, 1952; Gertz, 1957) and physiology (Weinger, 1953; Weinger et al, 1954; Scott, 1955; Ruegamer and Silverman, 1956; King and Eleftherion, 1959).

Marked effects were also reported on growth, mainly weight-gain (Weininger et al. 1954; Scott, 1955; McClelland, 1956; Mogenson and Ehrlich, 1958; Knowles, 1959) and survival rate/time (Weininger, 1953; Weininger et al, 1954; Scott, 1955; Levine and Otis, 1958; Levine and Cohen, 1959). Different rat strains (mostly Wistar and Sprague-Dawley) were used, generally immediately after weaning (at 21 days after birth). In one experiment (Levine and Otis, 1958) one group of rats was handled from the day after birth till day 21; one group was handled from day 21 till day 42 and the third group was gentled from weaning till day 42 (note 1). This experiment failed to confirm the hypothesis that post-weaning handling results in a greater resistance against severe deprivation in adulthood; the pre-weaning handled group showed a greater resistance and better ability to survive. The ability to survive and the changes in body weight were independent of the amount of handling since there were no differences between the picked-up and gentled rats in either the early or the late handled groups. Finally, in one experiment (Mogenson and Ehrlich, 1958) it was suggested that the effects of gentling on growth would be larger on 'wilder' or more reactive rats than on initially 'tamer' rats. In view of these inconsistent results further experimentation seemed warranted, questioning whether hamsters and rats when given a certain amount of human care would increase in weight (and weight-gain) and/or would be more resistant against disease when compared with the 'ignored' control groups ('ignored' always includes a minimum amount of handling, i.e. cleaning cages).

EXPERIMENT I

In this experiment we studied the effect of handling and gentling on weight-gain. If handling and gentling are functions of paranormal healing we expect a higher mean weight-gain in the experimental (gentling) condition than for the handling condition.

Method

Experimental animals: 23 hamsters (Mesocricetus auratus), 11 males and 12 females. Handler/gentlers: 4 non-healers who voluntarily participated. They were explicitly instructed to try to increase the body-weight of each animal in the experimental condition.

Procedure

The hamsters were randomly divided into four groups. Group 1 consisted of 5 male hamsters in the gentling condition and group 2 of 6 males in the handling condition. Group 3 consisted of 6 female hamsters in the gentling condition and 6 in the handling condition. The hamsters were subsequently assigned at random to the handler/gentlers.

TABLE 1
Assignment of hamsters to handler/gentlers.

	males		females	
h/g	gentling	handling	gentling	handling
1	3	3	-	-
2	2	3	-	-
3	-	-	3	3
4	-	-	3	3

Handling consisted of picking-up the animal every day (weighing once a week) and putting it back into its cage. Two methods were used for gentling: holding the animal in the left hand, placed against the chest so that the animal was 'cuddled' in the palm of the hand and with the right thumb stroking the back of the animal from the head to the base of the tail at a rate of approximately 50 strokes a minute (Weininger, 1953; Weininger et al, 1954; Scott, 1955; McClelland, 1956) or holding the animal and stroking it lightly along the head and back with the free hand (Ruegamer and Silverman, 1956; Levine and Otis, 1958; Mogenson and Ehrlich, 1958). The animals were allowed to move about in the experimenter's hand or arm. Gentling was done for 10 minutes/day during 42 days, starting immediately after weaning (day 22). The gentled hamsters had their left hind leg razored, the handled ones the right hind leg. Experimental and control hamsters of each handler/gentler were housed in the same cage.

Temperature and humidity were controlled, food (rat/hamster 1010 complete food, Hope Farms, Woerden, Holland) and water were available without restrictions. Cages always stood in the same room and on one level (table or shelf) to ensure equalized conditions. Weighing was done with the same scale on the same day and time. Differences between the conditions were tested by means of Student's t-test for male and female animals separately. Differences between individual handler/gentlers were tested by means of the Mann-Whitney test.

Results

Table 2 shows the mean weight-gain of male and female hamsters for the conditions gentling and handling.

TABLE 2
Effect of handling and gentling on mean weight-gain and standard deviation of hamsters

	condition	weight*	weight-gain
male	gentling	72.8 (6.9)	41.5 (5.3)
	handling	83.5 (9.9)	52.0 (11.0)
female	gentling	84.2 (11.9)	54.7 (13.4)
	handling	81.5 (10.1)	50.8 (7.8)

* weight after 42 days (gram)

There is no significant difference between the two conditions for both male hamsters ($t=1.87$; $p<.10$) and female hamsters ($t=.56$; p=n.s.). Table 3 shows the results for each of the four handler/gentlers. There only is a significant difference between handler/gentler 3 and 4 (MW-U=3; $p<.01$); the result of a significant difference in the handling condition (MW-U=0; $p<.05$).

TABLE 3
Effect of the handler/gentler on weight-gain

	gentling		handling	
h/g	male	female	male	female
1	41.0 (2.3)	-	57.6 (10.7)	-
2	42.1 (7.8)	-	46.3 (8.0)	-
3	-	63.8 (8.2)	-	57.8 (4.7)
4	-	45.6 (11.2)	-	43.8 (1.1)

Discussion

No significant differences were found between the handling and gentling conditions, with respect to weight-gain. In contradiction to our hypothesis handling even had a marginal effect on the weight-gain of male hamsters. Gentling proved to be no function of healing, meaning that the making of ´passes´ or the ´laying-on of hands´ does not effect weight-gain in normal hamsters. The overall difference between handler/gentler 3 and 4 is significant suggesting that handler/gentler 3 effects the weight-gain more than handler/gentler 4, but the effect is due to a difference only in the handling condition. This is in accordance with the remarks made above and we therefore do not ascribe the effects of handler/gentler 3 to a paranormal cause.

EXPERIMENT II

This experiment is a replication of the first experiment, using Lewis rats instead of hamsters. Since no differences were found between handling and gentling as functions of healing in the first experiment the gentling condition is compared with a control condition.

Method

Experimental animals: 20 Lewis rats, 9 males and 11 females. Gentlers: 4 non-healers (different from those who handled/gentled in experiment 1) who participated voluntarily.

Procedure

The rats were randomly divided into four groups. Group 1 consisted of 5 male rats in the gentling condition and group 2 of 4 male rats in the control condition. Group 3 consisted of 7 female rats in the gentling condition and group 4 of 4 female rats in the control condition. The rats were subsequently assigned at random to the gentlers.

TABLE 4
Assignment of rats to gentlers.

g	males		females	
	gentling	control	gentling	control
1	3	2	-	-
2	2	2	-	-
3	-	-	4	2
4	-	-	3	2

The gentling condition is described in experiment 1. The rats in the control condition were handled once a week for weighing; they were housed individually and cages numbered.

Results

Table 5 shows the mean weight and weight-gain of the male and female rats in the conditions gentling and control. Weight-gain in general was not affected by gentling. A negative effect on weight-gain for the

TABLE 5
Effect of gentling on mean weight-gain and standard deviation of Lewis rats

	condition	weight*	weight-gain
male	gentling	208.1 (10.9)	179.7 (10.5)
	handling	221.6 (4.0)	193.7 (4.9)
female	gentling	148.0 (9.3)	115.9 (5.8)
	handling	146.0 (12.4)	114.4 (8.3)

* weight after 42 days (gram)

gentled male rats was found, indicating that the animals in the control condition gained significantly more weight ($t=2.23$; $p<.05$). No effects for the female rats were found. Table 6 shows the results for each of the four gentlers. No overall difference was found between gentler 3 and 4 (MW-U=6; p=n.s.). However, a difference was found between gentler 3 and 4 in the control condition (MW-U=0; $p<.05$).

Discussion

The results of experiment 1 are replicated in this experiment. No effects of gentling on weight-gain were found for the female rats. There is, however, a significant difference in weight-gain for the male rats in favour of the control condition. This trend was also found in experiment 1 for the handling condition, indicating that gentling has a negative effect on weight-gain in male animals while it

TABLE 6
Effect of the gentler on weight-gain

	experimental		control	
gentler	male	female	male	female
1	160.17 (39.9)	-	196.75 (1.4)	-
2	178.39 (8.6)	-	190.74 (7.5)	-
3	-	113.67 (2.9)	-	107.58 (7.3)
4	-	117.61 (9.8)	-	121.24 (6.0)

has no effect in female animals. A similar result is reported by Beijnen and Van Tintelen (in press) for female rats of the Wistar strain. They compared a handling condition (cage cleaning) with a control group. A significant increase in group mean body-weight for rats that were not handled was found. The difference between gentler 3 and 4 was found (as in the handling condition in experiment 1) in the control condition and is not an overall difference. Therefore, we do not ascribe this effect to a paranormal cause.

EXPERIMENT III

In the introduction to this paper we mentioned an experiment (Mogenson and Ehrlich, 1958) which suggested that the effects of gentling would be larger on 'wilder' or more reactive rats than on 'tamer' rats. In experiment 1 and 2 the animals used were tame. For this reason in this experiment we used animals of a wilder strain to replicate experiments 1 and 2.

Method

Experimental animals: 30 rats (Fisher strain), 10 males and 20

females. Gentlers: 3 non-healers who participated voluntarily.

Procedure

The rats were randomly divided into four groups. Group 1 consisted of 5 males in the gentling condition and group 2 of 5 males in the control condition. Group 3 consisted of 10 females in the gentling condition and group 4 of 10 females in the control condition. The male rats were assigned to gentler 1, the female rats were assigned at random to the other 2 gentlers.

TABLE 7
Assignment of rats to gentlers.

g	males		females	
	gentling	control	gentling	control
1	5	5	-	-
2	-	-	5	5
3	-	-	5	5

Experimental and control groups were housed together in one cage and individually recognizable by colour markings. For further details see experiment 1.

Results

Table 8 shows the mean weight-gain of the male and female rats. No significant differences were found between the conditions. Table 9 shows a significant difference between gentler 2 and 3 for the female rats in the gentling condition (MW-U=2.5; $p<.02$). The female rats in the control condition of gentler 2, however, did not gain more weight than the rats in the gentling condition (MW-U=6; p=n.s.). The

TABLE 8
Effect of gentling on mean weight-gain and standard deviation of Fisher rats

	condition	weight*	weight-gain
male	gentling	237.4 (11.7)	162.8 (11.8)
	control	223.0 (14.8)	156.6 (8.6)
female	gentling	140.1 (14.0)	72.6 (12.7)
	control	140.9 (5.9)	74.9 (3.9)

* weight after 42 days (gram)

TABLE 9
Effect of the gentler on weight-gain

	experimental		control	
gentler	male	female	male	female
1	162.8 (11.8)	-	156.6 (8.6)	-
2	-	65.0 (12.2)	-	76.0 (3.6)
3	-	80.2 (5.2)	-	73.8 (3.5)

difference between the gentled and control rats of gentler 3, however, was significant in favour of the gentling condition (MW-U=3.5; $p<.03$).

Discussion

In this experiment we did not find a significant difference between the 2 conditions neither for the male nor for the female rats. From the first two experiments we learned that the male animals showed a lower weight-gain in the gentling condition than in the handling condition (the strongest effect is in experiment 2). In those experiments the animals represented tame strains while in this experiment the animals were of a wild and aggressive strain. The results of this third experiment show a significant difference in weight-gain in favour of the gentled female rats of one gentler, thus supporting the hypothesis of Mogenson and Ehrlich. The results of the two other gentlers however showed no significant differences between conditions.

GENERAL DISCUSSION

When we compare experiment 2 and 3 in which the animals are of different strains (tame/wild) a significant difference is found for each of the four groups. Table 10 shows weight-gains in percentage of original weight of the animals in experiment 2 and 3. The comparison

TABLE 10
Weight-gains in percentage of original weight (rats).

	male		female	
exp.	experimental	control	experimental	control
2	634.1 (32.7)	697.8 (42.9)	374.4 (74.0)	366.1 (34.0)
3	218.7 (19.5)	238.2 (21.5)	108.3 (21.2)	114.0 (9.5)

between the same conditions (condition 1: $t=21.82$; condition 2: $t=18.43$; condition 3: $t=10.11$; condition 4: $t=19.86$) shows a highly significant difference; in all conditions $p=<<.001$. The overall

difference between animals of both experiments is not significant (MW-U=251). The differences in weight-gain are probably due to strain differences. When we evaluate our hypothesis on handling and gentling as functions of paranormal healing we are tempted to conclude that neither handling nor gentling is a function of paranormal healing, with one restriction: in all experiments described above the animals are normal and healthy, while in healing practice human subjects suffer from a (chronic) disease.

EXPERIMENTS WITH SICK ANIMALS

In the first 3 experiments the dependent variable was weight-gain. Weight and/or weight-gain could not be used as a dependent variable for the next experiments. In the experiments described below the animals are injected with a disease, causing malignant cell growth and death. One symptom of the disease is loss of weight. Tumour weight (ascites and small solid tumours) is between 2 and 6% of the total body weight on day 10 after the injection and after 31 days this percentage increases from 25 to 50% (Snel, 1984). It is not possible to discern which part of the weight-increase is due to the tumour or other factors. Weight, therefore, is not a usable variable in these experiments. We introduce a different dependent variable, namely survival time. Survival time is closely related to mortality rate (percentage of survival). With respect to mortality rate an intrigueing observation was made by Spitz in 1946. He found a mortality rate of zero over a three and half year period for a total of 122 children in an institution where mothers or mother-substitutes were available for each child, and a mortality rate of 37% over a two year period for a total of 91 children in an institution in every way comparable with the first, but where it was not possible to give individual care to the children (note 2). Care and attention seem to be of the utmost importance for individuals to develop and live a healthy life. We therefore reformulate our hypothesis into the following: if handling and gentling are functions of paranormal healing we expect a longer survival time for the experimental (gentled) rats.

EXPERIMENT IV

All animals were injected with tumour. Tumour cell: a partially characterized rat tumour cell of thymic epithelial origin (Koninkx et al, 1984, 1986) was passed on intraperitoneally, resulting in voluminous ascites and solid tumours. The crude ascites, containing among other things lumps of- and free tumour cells, was harvested by abdominal puncture. An equal volume of ascites was mixed with plain RPMI-1640 culture medium (Flow Laboratories, Irvine, Scotland) and left on ice for 20 minutes. Washing was repeated twice. After discarding the supernatant 2 ml medium was added and thoroughly mixed, centrifuged (3 minutes, 20.000 g) in a glass capillary and the percentage tumour cells measured, from which the dilution factor was calculated. The suspension was diluted with plain RPMI to a 0.05% concentration of tumour cells to be used in the experiments. Each rat received 0.5 ml of this suspension by intraperitoneal injection on day 0. Mean survival time for juvenile rats (younger than 13 weeks) is 23 + 4 days, young adult rats generally survive the tumour for 25 + 6 days (Snel, 1984).

Method

Experimental animals: 20 Lewis rats, 9 males and 11 females.
Gentlers: 4 non-healers who participated voluntarily.

Procedure

The Lewis rats were randomly divided into four groups. Group 1 consisted of 5 male rats in the gentling condition and group 2 of 4 male rats in the handling condition. Group 3 consisted of 7 female rats in the gentling condition and group 4 of 4 female rats in the handling condition. The rats were subsequently assigned at random to a gentler. Rats were individually housed, gentling was done 10 min/day for 31 days.
Data collection on survival time: the rats were checked every hour from 07.00 h till 01.00 h. Survival time for the rats that died during the night was estimated by examining the usual distinctive features of death.

Results

Table 11 shows the mean survival time in hours of male and female rats. A marginal significant difference between the two conditions for the female rats in the gentling condition was found ($t=1.78$; $p<.10$). Although mean survival time for the male rats in the handling

TABLE 11

Effects of handling and gentling on mean survival time and standard deviation in hours of male and female Lewis rats with an experimental tumour

	condition	mean	standard deviation
male	gentling	558.0	33.9
	handling	575.3	47.5
female	gentling	592.6	86.7
	handling	502.8	37.3

condition was higher the difference between the gentling and handling condition was not significant ($t=-.50$; p=n.s.).

Discussion

Gentling had a marginal effect on the female rats in this experiment. The handling condition for the male rats had a non-significant higher mean survival time, in agreement with the results of weight-gain in the handling and control condition in experiments 1 and 2. The female animals of experiment 1 and 2 showed the same trend as observed in this experiment. Gentling in general does not seem to have a significant influence on the dependent variables; differences are due to individual handler/gentlers.

EXPERIMENT V

This experiment is a confirmation study of experiment 4, using larger groups of animals. We tried to replicate the results of experiment 4. Instead of the handling condition we again introduce a control condition similar to that of experiment 2.

Method

Experimental animals: 120 Lewis rats, 48 males, 72 females. Gentlers: 5 non-healers who participated voluntarily.

Procedure

The Lewis rats were randomly divided into four groups. Group 1 consisted of 24 male rats in the gentling condition and group 2 of 24 male rats in the control condition. Group 3 consisted of 36 female rats in the gentling condition and group 4 of 36 female rats in the control condition. The rats were subsequently assigned at random to the gentlers.

All animals were injected with the experimental tumour; experimental and control groups per gentler were housed in one cage. The rats were individually recognizable by colour markings. The experiment ended on day 31. Animals who survived were excluded from the analysis. In experiments 1 to 4 gentling was done for 10 minutes/day, each gentler gentling at least for 30 minutes/day. This proved to be too long: gentlers could not keep their concentration. In experiment 5 each gentler had twelve animals in his care, therefore gentling time was reduced to 3 minutes/rat/day.

Results

In the analysis the animals who survived were excluded, which means that for the male gentling condition the number of experimental animals was 15, for the control condition 16. For the female gentling

condition the number of experimental animals was 31, for the control condition 26.

Table 12 shows the mean survival time and standard deviation in hours of the rats. There were no differences between the conditions in mean survival time for neither the male (t=-.70; p=n.s.) nor the female rats (t=.32; p=n.s.)

TABLE 12

Effect of handling on mean survival time and standard deviation in hours of male and female Lewis rats with an experimental tumour

	condition	mean	standard deviation
male	gentling	550.4	78.3
	control	569.0	64.9
female	gentling	517.5	57.4
	control	512.8	50.6

Discussion

The gentlers in experiments 4 and 5 had a tough time. They knew the rats were to become (very) ill and eventually die. It is much easier to pity the rats and wish them to die faster than trying to 'heal' through gentling. The gentlers were highly motivated and tried their best in a disciplined way. There was no competition between the gentlers.

The results of experiment 4 were not completely confirmed in this experiment. Larger groups of animals caused the disappearance of the marginal significant effect of the gentling condition for the female rats, assuming handling is comparable to the control condition. Weininger et al (1954) noted that "gentling may be more beneficial

than group living, even when the gentled animals are caged individually". We checked for this possible "housing" effect. The results for weight, weight-gain and survival time, tested between experiments, showed no significant difference whether the animals were housed individually; experimental and handled/control groups were housed in a different cage or whether experimental and handled/control groups were housed together in one cage (results not presented).

FINAL DISCUSSION

We have conducted two series of experiments with non-healers. In the first series the experimental animals were healthy while in the second series they received an experimental tumour. In the first series the dependent variable, to be influenced by handling and gentling, was weight-gain, while in the second series, because weight-gain was not a valid variable, survival time was studied. Although these are major differences one trend emerges: generally gentling neither causes gain of weight in the first series nor extension of survival time in the second series. In the first series of experiments the handling and control condition even produced a higher score for the dependent variable than the gentling condition. This suggests that gentling by non-healers, which resembles the behaviour of healers during the healing process, has no or little effect for the dependent variable. Because the same holds for the second series the inevitable conclusion is that non-healers do not exert an influence on normal and sick animals when imitating ´healing behaviour´. This could mean two things: either suggestion does not influence lower order vertebrates (as stated in the introduction) or handling and gentling by non-healers do not influence the dependent variables in the same way as a paranormal healer is supposed to do. The importance of this study lies in the second conclusion: i.e. when working with animals in an experimental setting a non-healer does not influence the dependent variable. We intend to replicate this study with healers and non-healers because the results from this study indicate that suggestion, as operationalized by handling and gentling, does not ´simulate´ paranormal healing.

NOTES

1) In a pilot experiment we investigated the differential effect of

pre- and post-weaning handling and gentling with three groups of hamsters: group A (n=50) was handled once a day from the day of birth till day 42; group B (n=68) was handled from day 21 till day 42 and group C (n=11) was gentled from day 21 till day 42. No significant differences in mean weight were found between these groups on day 42 (68.1 +/- 3.6, 58.9 +/- 7.5 and 64.6 +/- 8.9 gram respectively).

2) Recent research on both preemies and rat pups supports earlier observations that lack of tactile stimulation - in preemies stroking; in rats licking - seems to slow their growth. Investigators stroked and handled 20 preemies according to a standard routine. The infants gained weight 21 percent faster than a control group that was not stroked. They also had higher levels of neurotransmitters and hormones (Scientific American, 1989, 261, September, 16A/B).

ABSTRACT

The effects of handling and gentling, as described in the literature, was studied in normal and sick animals. Weight-gain was the dependent variable in a series of experiments with normal hamsters and rats; survival time in two experiments with sick rats. A number of non-healers of both sexes (children and adults) cooperated for prolonged periods of time. No overall significant differences were found between the conditions handling/gentling or gentling/control. The results do not support the hypothesis that increased individual attention through handling and gentling by non-healers influences weight-gain or survival time.

REFERENCES

Bechterev, W. ´´Direct influence´ of a person upon the behavior of animals´, J. Parapsych., 1949, 13, 166-176.

Beynen, A.C. and Van Tintelen, G. ´Frequency of cage changing and weight gain of rats´, 1989, in press.

Cowley, J.J. and Widdowson, E.M. ´The effects of handling rats on their growth and behaviour´, British J. of Nutrition, 1965, 19, 397-403.

Gertz, B. ´The effect of handling at various age levels on emotional behavior of adult rats´, J. Comp. Physiol. Psychol., 1957, 50, 613-616.

Griffiths, W.J. and Stringer, W.F. ´The effects of intense stimulation experienced during infancy on adult behavior in the rat´, J. Comp. Physiol. Psychol., 1952, 45, 301-306.

Hall, C.S. and Whiteman, P.H. ´The effects of infantile stimulation upon later emotional stability in the mouse´, J. Comp. Physiol., 1951, 44, 61-66.

Hirsjarvi, P.A. and Junnila, M.A. ´The effects of gentling on open-field behaviour in rats´. In Beynen, A.C. and Solleveld, A.H. (Eds) ´New developments in biosciences: their implications for laboratory animal science´. Dordrecht, Martinus Nijhoff Publishers, 1988, 399-403.

King, J.A. and Eleftherion, B.E. ´Effects of early handling upon adult behavior in two subspecies of deermice, Peronyscum Maniculatus´, J. Comp. Physiol. Psychol., 1959, 52, 82-88.

Knowles, F.W. ´Rat experiments and Mesmerism´, J. Amer. Soc. Psych. Res., 1959, 53, 62-65.

Koninkx, J.F.J.G., Schreurs, A.J.M., Penninks, A.H. and Seinen, W. ´Induction of postthymic T-cell maturation by thymic humoral factor(s) derived from a tumour cell of thymic epithelial origin´, Thymus, 1984, 6, 395-409.

Koninkx, J.F.J.G., Penninks, A.H. and Seinen, W. ´In vitro and in vivo induction of terminal deoxynucleotidyl transferase activity in bone marrow cells by thymic humoral factors derived from a tumor cell of thymic epithelial origin´, Thymus, 1986, 8, 45-58.

Levine, S. ´Emotionality and aggressive behavior in the mouse as a function of infantile experience´, J. Genetic Psych., 1959, 94, 77-83.

Levine, S. and Otis, L.S. ´The effects of handling before and after weaning on the resistance of albino rats to later deprivation´, Can. J. Psych., 1958, 12, 103-108.

Levine, S. and Cohen, C. ´Differential survival to leukemia as a function of infantile stimulation in DBA/2 mice´, Proc. Soc. Exp. Biol. Med., 1959, 102, 53-54.

McClelland, W.J. ´Differential handling and weight gain in the albino rat´, Can. J. Psych., 1956, 10, 19-22.

Mogenson, G.J. and Ehrlich, D.J. ´The effects of early gentling and shock on growth and behaviour in rats´, Can. J. Psych., 1958, 12, 165-170.

Ruegamer, W.R. and Silverman, F.R. ´Influence of gentling on physiology of the rat´, Proc. Soc. Exp. Biol. Med., 1956, 92, 170-174.

Scott, J.H. ´Some effects at maturity of gentling, ignoring or shocking rats during infancy´, J. Abnormal Psych., 1955, 51, 412-414.

Snel, F.W.J.J. ´Ascites and solid tumours resulting from a tumour cell of thymic epithelial origin in the Lewis rat: weight and survival time´, Unpublished results, 1984.

Spitz, R.A. ´Hospitalism: a follow-up report´. In Freud, A. et al (Eds): The psychoanalytic study of the child. Vol.II. Intern. Univ. Press, New York, 1946, 113-117.

Van Bergeijk, J.P., Van Herck, H., de Boer, S.F., Meijer, G.W., Hesp, A.P.M., Van der Gugten, J, and Beynen, A.C. ´Effects og caging and gentling on behaviour, selected organ weights and blood constituents in female rats´, 1989, in press.

Weininger, O. ´Mortality of albino rats under stress as a function of early handling´, Can. J. Psych., 1953, 7, 111-114.

Weininger, O., McClelland, W.J. and Arima, R.K. ´Gentling and weight gain in the albino rat´, Can. J. Psych., 1954, 8, 147-151.

Williams, D.I. and Russell, P.A. ´Open-field behaviour in rats: effects of handling, sex and repeated testing´, British J. Psychol., 1972, 63, 593-596.

F.W.J.J. Snel
Nederlandse Vereniging voor Parapsychologie
Postbus 271
3720 AG Bilthoven
The Netherlands

C. van der Sijde
Univerity of Twente
Department of Education
P.O. Box 217
7500 AE Enschede
The Netherlands

EUR.J.PARAPSYCHOL., V.7, 1988-1989

THE RELATIONSHIP BETWEEN VOLUME AND YEAR OF PAST AND FUTURE ISSUES OF THE EUROPEAN JOURNAL OF PARAPSYCHOLOGY

Sybo A. Schouten
Editor European Journal of Parapsychology

We regret very much that due to unfortunate circumstances, discussed in more detail in the first article of this issue, it has not been possible to ensure regular publication of the European Journal of Parapsychology in recent years.

Since we believe that a European technical journal for research in parapsychology is desirable we intend to continue issuing the European Journal of Parapsychology. The journal used to be published by the Parapsychology Laboratory of the University of Utrecht. In view of the recent developments further publication of the journal will take place in cooperation with the Koestler Chair of Parapsychology of the University of Edinburgh.

For many years the subscription rate of the journal has been kept at the modest level of $10 a year. We would like to keep the subscription rate as low as possible. To maintain that subscription rate we have to reduce printing and mailing cost. Therefore we will combine issues and from now on we will print one issue a year. That implies that starting with volume 8 two issues will constitute a volume.

In the past each volume of the E.J.P. involved four issues which

were printed in a period of three consecutive years, but with overlap in years between volumes. The relationship between volumes of the E.J.P. and years of publication has been as follows:

Volume 1	covers	1975 - 1977
2		1977 - 1979
3		1979 - 1981
4		1981 - 1983
5		1983 - 1985
6		1985 - 1987

In order to maintain the relationship between volumes and years we rate the present issue as volume 7, numbers 2 to 4, i.e. Volume 7, number 2, 3 and 4 will all be contained within one issue of the journal. Volume 6 covered the period 1985 - 1987, and hence volume 7 number 1 and the present issue consitutes the entire volume 7 and covers the years 1987 - 1989.

Printing one issue a year implies that the next issue to be printed in 1990 will be Volume 8, numbers 1 and 2 (1989/1990), followed in 1991 by volume 8, numbers 3 and 4 (1990/1991). Hence in accordance with the past the continuation of E.J.P. volumes will be:

Volume 7	covers	1987 - 1989
8		1989 - 1991
9		1991 - 1993
	etc.	

We hope that reducing the number of issues to one a year will not lead to a decrease in number of articles or pages printed per volume and that the continuation of publication of the journal will contribute to strengthen the position of parapsychology as a scientific discipline in Europe.

EUR.J.PARAPSYCHOL., V.7, 1988-1989

THE THIRTY-THIRD ANNUAL CONVENTION OF THE PARAPSYCHOLOGICAL ASSOCIATION

The 33rd annual convention of the Parapsychological Association will be held 16 to 20 August, 1990, near Washington, D.C.

Anyone may propose a paper or poster (if it will not be published before the convention) to the Program Committee. Topics may be reports of laboratory or field research, or be methodological, theoretical, or historical.

A paper should be equivalent to a full-length journal article. The first sheet must have a centered title, author(s) and affiliation(s), followed by an abstract of no more than 300 words. The text may not exceed 12 single-spaced pages, with no more than five additional pages for figures, tables, and references. Papers must be typed on 8.5 x 11 inch paper, with margins of one inch on all sides, and should adhere to the style of the Publication Manual of the American Psychological Association (3rd edition).

Four copies are required. One copy must be single-spaced and camera-ready, for inclusion in the Convention Proceedings. If convenient, three copies should be double-spaced; if this is inconvenient, the other copies may be single-spaced.

Describe any required visual aids. If there are multiple authors, state who will make the presentation.

Posters are short papers presented on poster board. Proposals for posters must include four copies of the poster. Photocopies are acceptable.

Members and Associates of the Parapsychological Association may propose a symposium, panel discussion, or workshop. Proposals for symposia and panel discussions must include four copies of a summary sheet showing the title, chairperson, participants, order of presentation, and proposed time allotments (up to a total of 90 minutes, including the question period).

Proposals for symposia must also include four copies of a full paper (as above), one camera-ready, from each participant. For panel discussions, designed for spontaneous interactions, each participant's speaking time is limited to five minutes. No papers are needed. Proposals for workshops should include three copies of a summary sheet listing the title, chairperson, participants, and type of activity.

There are two deadlines for all these proposals: March 15 and April 15.

A proposal received by March 15 may be accepted or rejected, but it may also be returned to the author with suggestions for minor revisions that are necessary before it is accepted. April 15 is the last date for acceptance.

The Program Committee consists of Dr. Rick E. Berger, Dr. James E. Crandall, and Dr. Gertrude Schmeidler, Chairperson. Send all proposals to Dr. Schmeidler, 17 Kent Ave., Hastings-on-Hudson, NY 10706, U.S.A.

EUR.J.PARAPSYCHOL., V.7, 1988-1989

CORRECTION TO DR. ULRICH TIMM´S ARTICLE ´WHEN WILL WE BEGIN TO REDUCE ALPHA - AND BETA - ERRORS IN STATISTICAL PSI EXPERIMENTS?´

In the note on page 92 of Dr. Ulrich Timm´s article ´When will we begin to reduce alpha - and beta - errors in statistical psi experiments?´, published in the Volume 7, number 1 issue of this journal, a factor 2 was omitted in the expression given in the first sentence. The first sentence should read: The expression -2Sum ln P is also chi-square distributed, but with df=2k.