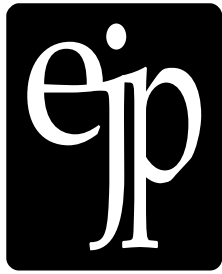




- 
- 90 *Assessing the relationship between ESP task success and vividness of voluntarily generated mental imagery as measured by Betts' QMI*  
**Ian R. Hume & Tony A. Lawrence**
- 117 *Do human fingers "see"? — "Finger-reading" studies in the East and West*  
**Yung-Jong Shiah & Wai-Cheong Carl Tam**
- 135 *The effect of weak magnetic fields on a random event generator: reconsidering the role of geomagnetic fluctuations in microPK studies*  
**Paul Stevens**
- 150 *Interpersonal psi — exploring the role of the experimenter and the experimental climate in a ganzfeld telepathy task*  
**Simon J. Sherwood, Chris A. Roe, Nicola J. Holt & Stuart Wilson**
- Book Reviews
- 
- 173 *Parapsychology in the 21st Century: essays on the future of psychical research*  
by Michael A. Thalbourne & Lance Storm (eds.) (2004)  
**Marios Kittenis**
- 186 *From shaman to scientist: essays on humanity's search for spirits*  
by James Houran (ed.) (2004)  
**Ian R. Hume & Tony A. Lawrence**
- 
- 191 *Obituary. Piero Cassoli*  
**Carlos Alvarado**





Indexed by PsycINFO.

Typeset by EJP. Printed by JEMDigital Print Services Ltd.

Copyright © 2005 European Journal of Parapsychology

**EDITOR: Paul Stevens**, Koestler Parapsychology Unit, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK. Email: Editor@ejp.org.uk

**ASSOCIATE EDITOR: Ian Baker**, Koestler Parapsychology Unit, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK. Email: AssociateEditor@ejp.org.uk

## EDITORIAL CONSULTANTS

**Eberhard Bauer**, Institut für Grenzgebiete der Psychologie und Psychohygiene, Freiburg, Germany

**Dick Bierman**, Psychology, Universities of Utrecht and Amsterdam, Netherlands

**Jason Braithwaite**, Behavioural Brain Sciences Centre, University of Birmingham, UK

**Deborah Delanoy**, Division of Psychology, University College Northampton, UK

**Alan Gauld**, School of Psychology, University of Nottingham, UK

**Anneli Goulding**, Department of Psychology, Göteborg University, Sweden

**Ian Hume**, School of Health and Social Sciences, Coventry University, UK

**Harvey Irwin**, School of Psychology, University of New England, Australia

**Peter Lamont**, Koestler Parapsychology Unit, University of Edinburgh, UK

**Edwin May**, Cognitive Sciences Laboratory, USA

**Chris Roe**, Division of Psychology, University College Northampton, UK

**Simon Sherwood**, Division of Psychology, University College Northampton, UK

**Christine Simmonds**, Rhine Research Center, USA

**Matthew Smith**, Psychology, Liverpool Hope University College, UK

**Fiona Steinkamp**, Koestler Parapsychology Unit, University of Edinburgh, UK

**Jon Sykes**, Glasgow Caledonian University, UK.

**Caroline Watt**, Koestler Parapsychology Unit, University of Edinburgh, UK

**Joakim Westerlund**, Department of Psychology, University of Stockholm, Sweden

**Carl Williams**, Psychology, Liverpool Hope University College, UK

**Stuart Wilson**, Psychology, Queen Margaret University College, UK

**Robin Wooffitt**, Department of Sociology, University of York, UK

---

**European Journal of Parapsychology** (ISSN: 0168-7263) is published twice a year in Spring and Autumn. Subscription enquiries to [subscriptions@ejp.org.uk](mailto:subscriptions@ejp.org.uk) or postal mail to the Editor at the above address. Permission from the Editor and author(s) must be secured before reproducing any EJP content for commercial use. Reproduction of material for nonprofit educational use is permitted.

## INSTRUCTIONS TO AUTHORS

The European Journal of Parapsychology (EJP) is a peer-reviewed scientific journal for research - particularly theoretical and theory-driven empirical work - relating to the field of parapsychology (defined as the study of communication or interaction between organisms and their environment that do not appear to rely on the established sensorimotor channels). The Journal's aim is to stimulate and enhance activity in parapsychology, particularly in Europe, by publishing Articles, Research Notes, Reviews and Comments that offer insight into or criticisms of parapsychological research. The Journal also publishes peer-reviewed Student Research Briefs. Submissions are welcomed on any topic that falls under the heading of parapsychology as defined above. Authors are requested to briefly define any specific terms used (including 'standard' terms such as ESP, PK) in the text of the paper. All manuscripts submitted to the EJP must be under consideration solely by this journal and may not have been previously published in any form (with the exception of translations of non-English articles). Initial submissions should be sent via email to [submissions@ejp.org.uk](mailto:submissions@ejp.org.uk) as an attached file (acceptable formats are ASCII, PDF, RTF, OpenOffice or MS Word).

ARTICLES are reports of experimental, clinical or theoretical research, with introductory comments and discussion of results generally being of direct relevance to the data or ideas being presented. Generally articles should be less than 10,000 words in length, though longer papers may be accepted after prior negotiation with the Editors. After evaluation by the editors, the initial article will be sent to two referees for peer-review.

RESEARCH NOTES are manuscripts of up to 3000 words in the following categories: conceptual replications; attempted replications which contradict the original study; reporting on new methodology or apparatus; new applications of existing methodology; challenges to key elements of an existing study, theory or methodology. Research Notes are not intended as a way of publishing inadequate or incomplete work. The Editors or Referees may also decide that a manuscript submitted as an article would be better published as a Research Note.

STUDENT RESEARCH BRIEFS are summaries of research (maximum 2000 words) conducted for assessment as part of a Bachelors or Masters degree (or equivalent) which would not otherwise be submitted as a full article but where the research is of sufficient interest to be disseminated to the wider parapsychological community. Student Research Briefs should provide a summary of methodology and results of the conducted research. They will be subject to peer-review with the understanding that the scope of the research will be less extensive than that of a full article.

REVIEWS provide critical and analytical review or assessment of previously published work. The EJP welcomes novel and cogent reviews (1500 words maximum), subject to Editorial approval.

COMMENTS are extended letters (1000 words maximum) that address specific issues raised in material published in the journal. All comments will be sent for peer-review, the process being similar to article review. Comments will be published in as timely a manner as possible, accompanied where appropriate with a reply from the original author(s).

Full formatting details are provided on the EJP website <http://ejp.org.uk/>

# European Journal of Parapsychology

Volume 20.2

2005



# Assessing the Relationship between ESP Task Success and Vividness of Voluntarily Generated Mental Imagery as Measured by Betts' QMI: A Review

Ian R. Hume and Tony R. Lawrence  
Department of Psychology, Coventry University

## Abstract

*Many psychic claimants suggest that the capacity for generating vivid visual imagery is conducive to the reception and detection of ESP mediated material. This paper reviews controlled studies which have aimed to test this claim by using Betts' Questionnaire upon Mental Imagery. In the case of restricted-choice studies, a significant negative relationship was evidenced between ESP task success and scores imaging ability. However, the combined studies indicated that their findings lacked homogeneity. For free-response studies, no discernible relationship was found between task success and imaging ability although there was evidence of ESP task success overall. Studies that required participants to use a voluntarily generated mental image as a tool for focus yielded a significant positive relationship between ESP task success and imaging ability. Studies were also homogenous in terms of effect size and direction. Further, there was evidence for a significant psi effect across these studies. It is concluded that if subsequent research is to investigate the relationship between vividness of imaging ability and ESP task success further, participants must not only report their capacity for voluntary imaging ability, they must also use this ability during ESP tasks in an appropriate manner.*

## Introduction

The reports of purported instances of ESP suggest that voluntarily generated mental imagery plays an important role in experiencing such phenomena. White (1964) has noted that many psychics highlight the importance of the capacity to generate vivid mental imagery in the process of their ESP experiences. Morris (1977) reviewed 74 “popular” books on the development of psychic ability. In many instances the generation of mental imagery was advocated. Therefore, it is evident that the relationship between individual differences in voluntarily generated mental imagery and ESP warrants investigation under controlled settings.

The questionnaire most widely employed to try to determine whether a relationship exists between ESP and individual differences in vividness of voluntarily generated mental imagery has been Betts’ Questionnaire on Mental Imagery (QMI: Betts, 1909), which consists of 150 items designed to measure across seven<sup>1</sup> different modalities. Betts found that modality sub-scale correlations ranged from .40 (visual and olfactory) to .78 (olfactory and gustatory) which suggest a general factor of imagery vividness. A revised shortened version (from here on referred to as the ‘Revised Version’) was developed by Sheehan (1967) which consists of only 35 items (five for each of the seven modalities) which reduced completion time from approximately 55 minutes to 10 minutes. Using two different samples, Sheehan reports correlations of .90 and .92 between the original and revised versions. Further, for the Revised Version, using principal component factor analysis a large general factor of imagery vividness was found — with average loadings of .57 for the 35 items. Standard tests of internal consistency also suggest a large general factor of imagery vividness. Using a same sample of 25 undergraduates, Shor, Orne, and O’Connell (1966) found split-half reliability of .91 on one occasion and .93 on another. Juhasz (1972) found an alpha coefficient of .95 on a sample of 67 students.

Richardson (1994) states that the reliability of the Revised Version is adequate when tested over time. In a review of studies concerned with test-retest reliability, he reports that coefficients ranged from .91 after a six-week interval (Evans & Kamemoto, 1973) to .54 after 12 months (White, Ashton, & Brown, 1977).

---

<sup>1</sup>These being the visual modality (consisting of 40 items), auditory, cutaneous, gustatory, kinaesthetic, and olfactory modalities (20 items each), and an ‘organic’ modality (10 items).

In turn, the Revised Version has also undergone revision by White, Ashton, and Law (1978) who developed a randomised version in an attempt to reduce the response set found when items concerned with the same modality are presented next to one another. However, the author is unaware of any published studies related to ESP research which have utilised this most recent revision.

In this review of ESP studies which employed a version of Betts' QMI, the entire history of the following periodicals were searched: *European Journal of Parapsychology*, *Journal of the American Society for Psychical Research*, *Journal of Parapsychology*, *Journal of Scientific Exploration*, *Journal of the Society for Psychical Research*, *Proceedings of the Society for Psychical Research*, and *Research in Parapsychology*. In addition, earlier reviews dealing with ESP and mental imagery by Honorton (1975), George (1981), and George and Krippner (1984) were also consulted.<sup>2</sup>

At this point, it should be mentioned that a high score on the Betts scale indicates poor imagery vividness, whilst a low score is characteristic of an individual who reports high imagery vividness. Thus, a negative relationship between scores on Betts' QMI and ESP success would indicate a positive relationship between increased imagery vividness ability and ESP success. However, for ease of interpretation of the ESP studies reviewed herein, this relationship has been inverted so that a reported positive correlation between scores on the Betts scale and ESP indicates a likewise correlation between imagery vividness and ESP success.

In preparing this review it was decided that restricted-choice and free-response studies would be analysed separately. The reason for this was that Stanford (1993) argued that whilst mental imagery could be an asset — in that it may well serve as a mediating vehicle for ESP — it could also prove to be a liability if it “is so abundant and free-flowing that it serves as a serious source of cognitive noise when the subject is trying to identify and report information about the ESP target” (p. 141). Stanford argued that this may be especially so in the case of free-response tasks.<sup>3</sup> Further, regarding restricted-choice tasks, he could see no reason why an abundance of vivid imagery was necessary

---

<sup>2</sup>Indeed, for a review of other imagery scales that have been employed in ESP research, these two papers are a valuable source of reference.

<sup>3</sup>We can only assume by this that Stanford meant that the target irrelevant imagery would interfere with or mask the target image.

at all. Thus, it would seem from Stanford's view, there should be a negative relationship between individual differences in imagery and ESP task success for free-response studies and no relationship at all between imagery and forced-choice task success. These are claims that can be put to the test in this review.

### **Restricted-choice**

The authors are aware of only five published restricted-choice studies that involved the use of one or other versions of Betts' QMI in an attempt to determine whether a relationship exists between individual differences in the vividness of mental imagery and ESP success. All five studies were designed to test for clairvoyance.

Honorton, Tierney, and Torres (1974) administered Sheehan's Revised Betts' QMI to 50 volunteer participants to determine the vividness of their imaging ability. Two of the researchers each ran 25 participants in individual sessions. Participants' scores on the questionnaire were used to categorise them as either 'strong' or 'weak' imagers via a median split (two scores tied at the median and were thus eliminated). The ESP task involved the use of either of two strategies for psi-guessing: the use of a visual image of the 'target' symbol; and imageless 'guessing'. Strategies were alternated by the participant after each run. There were no significant differences in psi scoring between strategies. However, there was overall significant psi-hitting for 'strong' imagers ( $t = 3.38$ ,  $df = 23$ ,  $p < .005$ , two-tailed), whilst 'weak' imagers produced significant psi-missing ( $t = 3.45$ ,  $df = 23$ ,  $p < .003$ , two-tailed). The difference between 'strong' and 'weak' imagers was also significant ( $t = 3.14$ ,  $df = 46$ ,  $p < .004$ , two-tailed).

Honorton considered mental imagery to be a mediating vehicle for ESP material. Therefore, in this study he reasoned that those who had a greater ability imaging should also be more confident in their target selection. Honorton et al report that correct confidence calls for strong images yielded a mean of 22.7 percent whereas weak imagers gave an average of 17.3 percent. They stated that the difference between strong and weak imagers regarding correct confidence calls was significant:  $t = 2.20$ ,  $df = 39$ ,  $p < .05$  (two-tailed). Because of the differences in hit rate and correct confidence calls for strong and weak images, Honorton et al concluded that there was "rather clear-cut evidence of a relationship between individual differences in mental imagery and psi performance"



(1974, p. 391).

However, Schechter, Solfvin, and McCollum (1975) — and later Stanford (1993) — note that “there was a serious statistical error in Honorton’s assessment...of the difference of correct confidence-call accuracy for strong and weak imagers” (Stanford, 1993, p. 140). In Honorton’s contrast between strong and weak imagers he neglected that strong imagers had a significantly higher hit rate than did weak imagers. Therefore, the higher number of correct confidence calls for strong imagers would be higher than for weak imagers by chance alone. Using the data from the Honorton et al study, Stanford calculated correct confidence call accuracy based on the proportion of hits gained by strong and weak imagers. He found that “strong imagers did not perform significantly above MCE [mean chance expectation] on confidence calls, given their actual hit rate in the ESP task ( $z = .52$ ) [and] ...weak imagers did not perform significantly below MCE in their confidence-call accuracy, considering their hit rate ( $z = -.80$ )” (Stanford, 1993, p. 140). He thus concluded that contrary to the claims of Honorton et al, there was no significant difference between the levels of correct confidence calls for the two groups. Indeed, if we contrast<sup>4</sup> the two  $z$ -scores given by Stanford, we find that  $z = -.20$ ,  $p = .422$  (two-tailed). Therefore, from the results of this study, whilst there is evidence of a positive relationship between imaging ability and ESP, there is no evidence that participants imaging ability significantly increases confidence in correct target identification.

In a response to the conclusions drawn by Honorton et al, Schechter, Solfvin, and McCollum (1975) employed 48 volunteer participants to provide an independent replication of the above study. Again, no significant differences in psi-scoring between strategies were found. However, contrary to the original study, it was found that there was significant psi-hitting for ‘weak’ imagers ( $t = 3.92$ ,  $df = 22$ ,  $p = .002$ , two-tailed) and non-significant psi-missing for ‘strong’ imagers ( $t = 1.10$ ,  $df = 22$ ,  $p > .05$ , two-tailed). The difference on ESP scores between ‘strong’ and ‘weak’ imagers was significant ( $t = 3.21$ ,  $df = 44$ ,  $p = .005$ , two-tailed). Regarding participants correct confidence calling, weak imagers gave an average hit rate of 24.8 percent whereas strong imagers yielded an average of 19.6 percent. Schechter et al report that the difference between correct confidence-calling was not significant and, as with the

---

<sup>4</sup>This is calculated thus:  $z_1 + z_2$  divided by the square root of the number of  $z$ 's (in this instance, two).

difference found in the study by Honorton et al, merely reflects the difference in the proportion of hits gained by the two groups.

Pleshette (1975) conducted a clairvoyance study designed to explore whether a relationship exists between psi-guessing performance on a binary Schmidt random number generator and EEG alpha rhythm activity. Before the psi-task, there was a trial period of alternating two-minute 'tone on', 'tone off', and 'rest' periods. During the tone on periods, participants attempted to keep the tone on by maintaining their EEG activity in the alpha frequency. During the tone off periods, they were required to keep the tone off — again by maintaining alpha activity. During the rest period participants were not required to maintain the presence or absence of a tone (thus feedback was not given for alpha activity regulation during these periods). After a short break, participants were subjected to a further five tone on, tone off, and rest trials, except during these periods they had to make at least five guesses on the random number generator during the alpha feedback trials. Participants were encouraged to only make guesses when they had some degree of confidence. Auditory feedback was given for hits. No significant relationships were discovered between alpha activity and psi-guessing performance. Unfortunately, details regarding participants' overall psi-scores in relation to mean chance expectation were not given.

Pleshette also administered the original Betts' QMI to 20 volunteer participants prior to the study and, on the basis of their scores, they were labelled as either 'strong' or 'weak' imagers via a median split. Pleshette found that those labelled as 'weak' imagers had a significantly higher psi-guessing performance than those labelled as 'strong' imagers ( $t = 2.22$ ,  $df = 18$ ,  $p = .039$ , two-tailed).

Rao, Rao and Rao (1977) presented their 20 volunteer participants with a series of 60 slides via a tachistoscope. Each slide was said to contain a picture of one of ten possible objects. In fact, only thirty slides contained pictures, whilst the other 30 were blank. Thus, the study was designed so that 30 pictures could have been perceived subliminally, whilst the other thirty slides might perhaps have been perceived by ESP, i.e., if a participant's response to a blank slide corresponded with a concealed target picture, then it was considered a hit. Overall, significant psi-hitting was found ( $CR = 2.18$ ,  $p < .05$ , two-tailed). Rao et al used participants' group mean scores on Sheehan's Revised Betts' QMI (which was administered before the ESP task) to divide them into 'vivid' and 'weak' imagers. It was found that both groups scored positively on the

ESP task, with 'vivid' imagers scoring hits to a significant extent ( $CR = 2.109, p < .05$ , two-tailed). However, the difference between the two groups failed to reach significance ( $t = .99, df = 18, p > .05$  two-tailed).

Sargent (1978) employed a modified version of the Revised Betts' QMI to determine the vividness of mental imagery of his 20 participants on two modalities (visual and auditory) and, on the basis of their scores, labelled them as either 'high-imagers' or 'low-imagers'. The nature of the study involved a recall task of word lists (consisting of 16 words), in which the frequency and imageability of the words varied systematically. Participants were given 120 seconds to memorise each word list and then asked to record the words that they remembered on a response sheet. The ESP task required that participants record each letter of the recalled words either above or below a horizontal midline on the response sheet. If the participant recalled less than the maximum 16 words, they were asked, instead of placing the word letters, to place X's in either of the available spaces reserved for each letter. A pre-selected binary random digit determined the target position for each letter (either above or below the line). Thus the ESP test was a binary-choice task. Overall psi-scoring was not significant. However, Sargent found that 'high-imagers' had a significantly lower psi-score than 'low-imagers' ( $CR_d = 2.41, p < .016$ , two-tailed).

It can be seen from these studies that a literary review makes it difficult to determine whether or not a relationship exists between scores on Betts' QMI and ESP success. A summary of statistics derived from the five studies is presented in Table 1 below which gives the  $z$ -scores and effect sizes ( $r$ ) together with the means and standard deviations. It can be seen that the  $z$ -scores range from 2.91 to  $-2.96$  ( $M = -.70, SD = 2.55$ ) whilst computed effect sizes range from .420 to  $-.539$  ( $M = -.158, SD = .447$ ).

An analysis of these studies was performed using the Stouffer method of adding the standard normal deviates ( $Z$ 's) for each study and dividing the sum by the square root of the number of studies, yielding a new Stouffer  $Z$  for the combined studies.<sup>5</sup> Stouffer  $Z = 1.57$ , which gave a non-significant  $p = .116$  (two-tailed). The combined effect size estimate was determined by computing  $r$  and the associated Fisher  $z_r$  for each study and then summing them to find the mean  $z_r$  and its associ-

---

<sup>5</sup>It was intended to determine whether or not there was evidence of psi across restricted-choice studies. However, only one study (Rao, Rao and Rao, 1977) reported adequate statistical details regarding ESP task success.

Table 1: Studies included in the analysis, displaying their sample size ( $N$ ), computed  $z$ -scores, effect-sizes ( $r$ ), means, and standard deviations ( $SD$ )

Study	$N$	$z$ -score	$r$
Honorton, Tierney, & Torres (1974)	48	2.91	.420
Schechter, Solfvin, & McCollum (1975)	46	-2.96	-.436
Pleshette (1975)	20	-2.08	-.465
Rao, Rao, & Rao (1977)	20	1.02	.228
Sargent (1978)	20	-2.41	-.539
<i>Mean</i>		-.70	-.158
<i>SD</i>		2.55	.447

ated  $r$ . The mean effect-size estimate,  $r$ , for these five studies =  $-.177$ . Thus the analysis evidenced a non-significant negative relationship between restricted-choice ESP hits and scores on Betts' QMI. However, it is clear from Table 1 that the range of  $z$ -scores (and effect sizes) across the studies was considerable. Therefore, a test for homogeneity of  $z$ -scores was computed yielding a highly significant result of  $\chi^2_{(4)} = 26.836, p < .01$ , thus indicating that the  $z$ -scores were heterogeneous. Indeed, when the studies were weighted by their  $df$ , the chi-square statistic (based on Fisher's  $z_r$  transformation) increased:  $\chi^2_{(4)} = 28.249, p < .01$ . The results of the inferential statistics in this analysis are presented in Table 2.

Table 2: Computed results of the analysis, both unweighted and weighted by sample size, including the number of studies ( $k$ ), Stouffer  $Z$ , effect size, and  $\chi^2$  test of homogeneity.

	$k$	Stouffer $Z^a$	$r$	$\chi^2^b$
Unweighted	5	1.57, $p = .116$	-.177	26.836, $p < .01$
Weighted	5	.80, $p = .424$	-.107	28.249, $p < .01$

<sup>a</sup>two-tailed

<sup>b</sup> $df=4$

If we refer back to Table 1, it can be seen that the two largest  $z$ -scores in either direction were yielded by the studies of Honorton, Tierney, and Torres (1974) and Schechter, Solfvin, and McCollum (1975). These results undoubtedly play a major role in the heterogeneity of the studies. Further, these two studies also have the largest sample sizes (both more than twice that of the other three studies) which would account for the slight increase in the chi-square statistic when weighted by sample size. It is difficult to account for the diametrically opposing results of these two studies, but an attempt must be made if we are to determine the

relationship between ESP and scores on the Betts'.

Schechter et al report that in their replication study, the same procedure and conditions as the Honorton et al study were adhered to. However, one difference that is clear is the reasoning of the two groups of researchers regarding the relationship between mental imagery and ESP. From the results of their study, Honorton et al conclude that there is "clear-cut evidence of a relationship between individual differences in mental imagery and psi performance" (1974, p. 391). The authors also imply that the reason why there was no difference in psi performance between conditions was because participants adopted the imaging strategy over all six runs, regardless of instructional set. On the other hand, Schechter et al argue that "there is no logical necessity to assume that images had *any* critical role in producing the significant results of [Honorton et al's] study" (Schechter et al, 1975, p. 332). Further, they argue that the type of cue used in each run is irrelevant to ESP task success. Thus, Schechter et al argue that there may be merely a correlation between strength of imagery and ESP, but not a causal link.

Presumably in both studies, the researchers involved were not blind to the hypotheses generated. Therefore, in either study we do not know how much emphasis was placed on the instructions for the two conditions — either inadvertently or otherwise. Further, we do not know whether the participants themselves were aware of what the research hypotheses were, but for Honorton to imply that participants in his study may have used the imaging strategy in both conditions suggests that the participants in his study may at least have had inkling. What we do know is that for both studies (unlike the other three restricted choice studies in this review), participants completed the Revised QMI *after* the ESP trials. Although in both instances feedback on psi task success was withheld from participants until completion of the scale, this would not stop participants from being influenced in their scoring of the questionnaire based on confidence of their success on the psi task — confidence gained from either within or from the response of the sender or researcher, or both. Thus, although conjectural, knowledge of the hypothesis and completion of the scale after the study, coupled with confidence in task success may go a long way in explaining the very different results of these two studies.

A more satisfactory follow-up study could have used researchers and senders who were blind to the hypothesis being tested — with receivers completing the scale before ESP trials and the scores withheld

from the researcher and sender. Until such studies are conducted, it is difficult to determine the relationship between vividness of imagery of the receiver and restricted-choice ESP from the outcome of the above two studies.

In the meantime however, because of the questionable nature of the procedure of these two studies, their undoubted impact on the outcome of an analysis with a handful of studies can be reduced in either of two ways. First, and more drastically, the two studies can be omitted from the analysis. Secondly, they could remain in the analysis but given a lesser weighting than the other three studies that do not share this problem. However, if we adopted the latter approach, uncertainty will always remain as to the magnitude of the relationship between ESP and vividness of imaging ability. To eradicate this uncertainty, we can only include studies where possible artefactual effects have been considered. Therefore, it was decided to omit the two studies from the analysis. Table 3 below gives a summary of descriptive statistics for the three remaining restricted-choice studies in the analysis. As can be seen, compared with the analysis of all five studies, the mean  $z$ -score of -1.16 was larger suggesting a more negative relationship between ESP task success and scores on the QMI (Mean  $r = .259$ ).

Table 3: Studies included in the analysis, displaying their sample size ( $N$ ), computed  $z$ -scores, effect-sizes ( $r$ ), means, and standard deviations ( $SD$ )

Study	$N$	$z$ -score	$r$
Pleshette (1975)	20	-2.08	-.465
Rao, Rao, & Rao (1977)	20	1.02	.228
Sargent (1978)	20	-2.41	-.539
<i>Mean</i>		-1.16	-.259
<i>SD</i>		.1.89	.423

A Stouffer  $Z$  of -2.00 was found, indicating a significant negative relationship between ESP task success and scores on the Betts':  $p = .046$  (two-tailed), with an estimated effect size of  $r = -.283$ . However, eliminating the two studies did little to eradicate heterogeneity within the studies:  $\chi^2_{(2)} = 9.480, p < .01$ . Weighting the three studies by their  $df$  was pointless as sample sizes were the same. A summary of the inferential statistics is presented in Table 4.

We can see that the analysis of these three remaining restricted-choice studies yielded a significant negative relationship between ESP task success and scores on Betts' QMI. Therefore it is unlikely that we

Table 4: Computed results of the analysis, both unweighted and weighted by sample size, including the number of studies ( $k$ ), Stouffer  $Z$ , effect size, and  $\chi^2$  test of homogeneity.

	$k$	Stouffer $Z^a$	$r$	$\chi^2^b$
Unweighted	3	2.00, $p = .046$	-.283	9.480, $p < .01$

<sup>a</sup>two-tailed

<sup>b</sup> $df=2$

can accept the suggestion made by Stanford (1993) that there would be no relationship at all between the two. It would seem that regardless of the amount of imagery involved in the ESP task stimulus, the capacity for vivid mental imaging may serve as a potential source of noise that inhibits the reception and detection of restricted-choice ESP. However, the lack of homogeneity for these three studies leaves this conclusion as rather tentative. Therefore, more research employing restricted-choice methods is warranted if only to see if heterogeneity reduces to a non-significant level as the number of studies increases.

### Free-response

Having reviewed all published restricted-choice studies, it is the turn of free-response studies. Again, analysis of studies designed to determine this relationship will be employed in an attempt to estimate the overall effect size, direction, and statistical significance.

Only five free-response studies were found that investigated the relationship between ESP task success and mental imagery. All five employed a version of Betts' QMI to measure individual differences in imaging ability.

Palmer and Vassar (1974) administered the Revised Betts' QMI to 60 volunteer participants in a clairvoyance study. Primarily, the study was designed to test the efficacy of an Out-of-Body-Experience (OBE) induction technique for its possible psi-facilitating properties. The overall aim of the induction technique was to try to develop a sense of separation of the self from the physical body so that the participant could pass from the control room through the wall into the room where the target material was situated so that it could be identified. Overall, ESP scoring was not significant ( $z = -.20$ ,  $t_{(59)} = 1.85$ ). However, participants who reported having an OBE during testing, scored significantly below chance ( $\bar{z} = -.34$ ,  $t = 2.24$ ,  $df = 20$ ,  $p < .05$ , two-tailed), whilst the remainder who did not report an OBE failed to reach significance levels on ESP testing.

ESP success and scores on the Revised Betts' QMI revealed a significant positive relationship ( $r = .35$ ,  $df = 53$ ,  $p < .01$ , two-tailed). However, it must be noted that participants did not complete the revised Betts scale until after they had received feedback concerning their ESP scores. This being the case, it is difficult to determine whether the relationship between ESP success and scores on the Revised Betts' QMI is genuine or artefactual, i.e. the result of demand characteristics.

In a second study using a clairvoyance task, again designed to test the efficacy of an OBE induction technique for ESP success Palmer and Lieberman asked half of their 40 participants to imagine leaving their bodies in order to locate and identify target material located in another room, whilst the other half were asked to simply report their spontaneous mental imagery during the session (the control condition). Thirteen of the 20 participants in the OBE condition reported having the feeling of being outside their body during the study, compared to only four out of 20 in the control condition. The 17 participants who reported having an OBE during the testing session yielded mean hit rates significantly above chance on the ESP task ( $M = .44$ ,  $t_{(16)} = 2.60$ ,  $p < .02$ ) whereas task success for the 23 non-OBEers was not significant ( $M = -.06$ ,  $p = \text{n.s.}$ ) The difference between these two means was significant:  $t_{(38)} = 2.11$ ,  $p < .05$ ). Thus, this finding contradicts that found by Palmer and Vassar (1974). Overall ESP task success did not deviate from mean chance expectation.

In this study, Palmer and Lieberman (1975) also manipulated the administration of the Revised Betts' QMI to 40 participants in order to determine whether the relationship between scores on the scale and ESP success found in the first study was artefactual or not. Half the participants were given the scale before receiving feedback concerning their ESP score, whilst the other twenty completed the scale after they were given feedback. As with the first study, there was a significant positive relationship between scores on the Revised Betts' QMI and ESP success for participants who completed the scale after feedback ( $r = .59$ ,  $df = 18$ ,  $p < .01$ , two-tailed). However, no such relationship was manifest for those participants who completed the scale before feedback ( $r = -.17$ ,  $p = \text{n.s.}$ ). Thus, Palmer and Lieberman concluded that the relationship between ESP success and scores on the Revised Betts' QMI found in the earlier study was indeed artefactual.

Smith, Tremmel and Honorton (1975) employed ganzfeld stimulation to both senders and receivers in a study comparing the effects



of subliminal versus supraliminal presentation of target material on GESP success. In the subliminal condition, senders were presented with the target picture tachistoscopically for approximately one millisecond prior to ganzfeld stimulation. In the supraliminal condition, the target material was presented to senders for ten minutes (again, prior to ganzfeld stimulation). Betts' QMI was administered to both sender and receiver before the first experimental session commenced. After the ganzfeld session, receivers were asked to perform a content analysis of their mentation and record whether there was the presence or absence of material from ten content categories. Overall, there was significant psi-hitting ( $t = 2.25$ ,  $df = 39$ ,  $p = .015$ , one-tailed). However, receiver ESP task success did not correlate significantly with scores on Betts' QMI in either condition.<sup>6</sup> In the subliminal condition the correlation was  $-.24$  ( $df = 18$ ), whilst in the supraliminal condition it was  $-.10$  ( $df = 18$ ).

Braud and Wood (1977) administered the shortened version of the QMI to thirty participants in a GESP free-response ganzfeld study. To test the effects of feedback on correct target content (by signalling a tone), participants were divided into two groups of fifteen, with only one group receiving feedback during practice trials. On a formal trial before the feedback practice trials, neither group evidenced psi task success. However, in a formal trial after the practice trials,  $t$ -tests revealed that the group that had received feedback evidenced significant improvement over the pre-practice trial ( $t = 2.67$ ,  $df = 14$ ,  $p < .02$ , two-tailed) and showed significant psi-hitting ( $t = 2.50$ ,  $df = 14$ ,  $p < .05$ , two-tailed). Participants in the no feedback condition failed to produce evidence for psi in either pre or post-practice trial tests. Furthermore, there was no significant improvement in psi task success between the two formal trial tests. Whilst the effect size between ESP task success and scores on the shortened QMI was not stated, the researchers report that, collapsed across conditions, there was no significant relationship in either the pre or post-practice sessions.

Finally, in a clairvoyance task, Sargent (1980) administered the visual, auditory, and 'body-imagery'<sup>7</sup> sub-scales of Betts' QMI to thirty participants prior to ESP testing under ganzfeld conditions. The hit rate of 53.3% was highly significant ( $p < .0004$ , one-tailed). However, the re-

---

<sup>6</sup>Further, there was no significant relationship between the senders' scores on Betts' QMI and task success in either condition.

<sup>7</sup>The 'body-imagery' sub-scale was developed from items taken from the cutaneous, kinaesthetic, and 'proprioceptive' sub-scales of the original Betts' QMI.

relationship between ESP success and scores on this modified version of Betts' QMI, whilst positive, failed to reach significance.

As with the restricted-choice studies, the relationship between scores on Betts' QMI and free-response ESP task success is unclear. Hence, an analysis of free-response studies was performed in an attempt to clarify the relationship should one exist. Because the significant positive relationship between scores on the Revised Betts' QMI and ESP success in the study by Palmer and Vassar (1974) was considered to be artefactual, this study was not included in the analysis. Likewise, the results gained for those participants who completed the scale after receiving feedback on ESP scores in the study by Palmer and Lieberman (1975) were also not included in the analysis. In the case of studies such as that of Braud and Wood (1977) when a statistic for an effect is not reported, Rosenthal (1991) notes that there is little that can be done other than to report an  $r$  effect size of .00 (p. 28). As the effect sizes of the few studies considered suitable for the analysis were small, a non-significant result of the analysis was expected. Therefore it was decided that nothing would be gained by including a study in the analysis with an *assumed* effect size of .00 which would reduce any effect further still. For this reason, and for the sake of accuracy, Braud and Wood's study was not included in the analysis.

The study by Smith, Tremmel, and Honorton (1975) employed a repeated measures design to test the difference between subliminal and supraliminal presentation of the target material. Thus, results for each condition were presented. Rosenthal (1991, p. 27) notes that "treating non-independent studies as independent does tend to create errors in significance testing" and therefore recommends that a single effect-size and a single significance level be computed for each study included in a meta-analysis. For this reason, the mean effect-size and the mean level of significance for the two sets of results were calculated, and it was these figures that were entered in the analysis.

It can be seen from Table 5 that the three free-response studies  $z$ -scores ranged from -.76 to .92 reflecting some degree of heterogeneity. However, as can be seen in Table 1.6, unlike the restricted choice studies these scores were not significantly heterogeneous:  $\chi^2_{(2)} = 1.900, p > .05$ . This remained so when studies were weighted by their  $df$ :  $\chi^2_{(2)} = 1.743, p > .05$  (based on Fisher's  $z_r$  transformation).

Table 6 also gives further results of the analysis. The computed Stouffer  $Z$  for the studies in this analysis yielded a non-significant  $p$ -

Table 5: Studies included in the analysis, displaying their sample size ( $N$ ), computed  $z$ -scores, effect-sizes ( $r$ ), means, and standard deviations ( $SD$ )

Study	$N$	$z$ -score	$r$
Palmer & Lieberman (1975)	20	-.76	-.170
Smith, Tremmel, & Honorton (1975)	20	-.76	-.170
Sargent (1980)	29	.92	.170
<i>Mean</i>		-.20	-.057
<i>SD</i>		.97	.196

value (Stouffer  $Z = .35$ ,  $p = .726$ , two-tailed). A very small effect size was computed:  $r = -.057$ . Thus the analysis of free-response studies evidenced a non-significant negative relationship between ESP hits and scores on Betts' QMI. When the three studies were weighted by their  $df$ , this yielded a Stouffer  $Z$  of .05 ( $p = .960$ , two-tailed). The combined effect size of these studies, when weighted by their  $df$ , produced a near-zero negative relationship:  $r = -.023$ .

 Table 6: Computed results of the analysis, both unweighted and weighted by sample size, including the number of studies ( $k$ ), Stouffer  $Z$ , effect size, and  $\chi^2$  test of homogeneity.

	$k$	Stouffer $Z^a$	$r$	$\chi^2^b$
Unweighted	3	.35, $p = .726$	-.057	1.900, $p > .05$
Weighted	3	.05, $p = .960$	-.023	1.743, $p > .05$

<sup>a</sup>two-tailed

<sup>b</sup> $df=2$

Similar to the forced-choice studies reviewed earlier, whilst the direction of the relationship is negative, the effect size is so small that we are forced to conclude that in the case of free-response ESP studies there is no discernible relationship at all between ESP and mental imagery. Thus, regarding the suggestion made by Stanford (1993) that there would be a negative relationship between mental imagery and ESP free-response performance, we can see from the results of this review that there is no support for this.

It was also intended to determine whether there was evidence of a significant psi effect for these three studies. However, Palmer and Lieberman do not report task success for those who completed Betts' QMI before the psi task. Nevertheless, ESP task success two other studies included in this analysis yielded impressive significant effects. Even if we gave Palmer and Lieberman's study a  $z$ -score of zero, the Stouffer

$Z$  for the three studies is 3.08,  $p = .001$  (one-tailed). Thus we can assume that there was evidence of a significant psi effect across these studies but no evidence that the capacity for generating vivid voluntarily mental imagery contributed to this success.

### **Criticisms of the construct validity of Betts' QMI**

Having noted the rather inconsistent results from studies concerned with ESP and scores on Betts' QMI, Honorton (1975) questioned the construct validity of the questionnaire. As very few studies concerned with the relationship between ESP and mental imagery employed Betts' QMI after Honorton's paper, its publication would seem to have been influential<sup>8</sup>. Indeed, it can be seen that the last study to use the QMI in ESP research was published in 198. Therefore, if the results of the above analyses are to be considered valid indicators for the relationship between mental imagery vividness and ESP, an examination of Honorton's criticisms lodged against the questionnaire is necessary. The criticism focuses on two non-parapsychological areas of research with Betts' QMI, which Honorton evidently considered important in his evaluation of its construct validity. These are the scale's relationship with verbal and visual recall tasks and psychophysiological measurements.

#### *Recall*

Honorton cites four studies which failed to find a relationship between scores on Betts' QMI and participant accuracy on some type of visual or verbal recall task. On the basis of these findings, he concluded that "the failure of the test to relate significantly to a variety of verbal and visual recall tasks calls into question its construct validity" (1975, p. 330). It is interesting to note, however, that Richardson (1994, p. 7) points out that:

"Vividness is not another name for accuracy... Whatever it is that we measure with our vividness questionnaires, it is not something that necessarily predicts high and low accuracy of recall. A visual image is not a mental copy of something previously observed; under no circumstances does it enable the imager to read off details as if looking at the original object."

<sup>8</sup>However, it may simply be the case that researchers abandoned this research because of the inconsistency of findings. Nevertheless, if criticisms of a scale have been levelled we must investigate whether or not the criticisms are valid.

Therefore, according to Richardson, it would seem that we cannot draw any inferences concerning the construct validity of Betts' QMI based on the results of tests of recall. Richardson (1994, p. 8) clarifies the concept of vividness of imagery by noting that:

“When deliberately constructed by a vivid imager, a scene from the actual past, or the fantasied past or future, should have more profound effects on present response systems than do similar constructive attempts by a weak imager. Thus, vivid thought imagery also functions to give a *sense of present reality* to whatever is being imaged. The consequences of establishing a *sense of present reality* to the temporary exclusion of *actual reality* is that actual experiential, behavioural, and physiological responses will follow accordingly.”

Thus, at least for Richardson, perhaps Betts' QMI's relationship with psychophysiological measurements may provide a more appropriate test of its construct validity than do recall tasks.

#### *Psychophysiological Correlates*

Unfortunately, Honorton claims that the findings from psychophysiological research are no more encouraging. He cites a study by Gale, Morris, Lucas, and Richardson (1972) as evidence for this claim. Gale et al used the Revised QMI to compare the differences in EEG alpha suppression between vivid and weak imagers under nine different conditions of imagery instruction. On the basis of scores on the QMI, they reported no significant differences in alpha abundance between vivid and weak imagers, and that a significant difference in mean dominant alpha frequency was found in only one of the nine conditions. Thus, the Revised QMI failed to detect individual differences in EEGs during imaging.

However, Gale et al suggested that under all the conditions, weak imagers found the task difficult, and therefore alpha activity was suppressed. Vivid imagers, on the other hand, may have shown alpha suppression either because of the difficulty of the task *or* the richness of their imagery. Thus, Gale et al. conclude that there may have been a common shift in alpha abundance, but for different reasons. Obviously, EEG records can not distinguish between such underlying reasons. Therefore, it could be argued that the construct validity of the QMI can not reasonably be called into question on the basis of this study. Rather, it

is more likely that EEG measures may, in retrospect, have been inappropriate for such an investigation of individual differences in thought imagery. Nevertheless, the results of this study still do not help to determine the adequacy of the construct validity of Betts' QMI.

Since Honorton's critique, there have been a number of studies published that have plotted some form of physiological measurement against scores on Betts' QMI. The majority of these studies reported herein have been reviewed recently by Richardson (1994). However, given that it would appear that the use of Betts' QMI to investigate the relationship between imagery vividness and ESP was abandoned because of criticisms of its construct validity, it was considered necessary to report them here in detail rather than redirect the reader. As the ESP studies have utilised only the original version of Betts' QMI and Sheehan's (1967) Revised Version, only those psychophysiological studies that have also used these two versions of the scale have been reported.

**Heart Rate:** On the basis of their scores on the Revised Betts' QMI, Sutherland and Harrell (1986 – 87) selected 31 vivid and 31 weak imagers from a sample pool of 120 black female undergraduates. The experiment aimed to provide imageable stimulus material that would evoke a detectable physiological response. It was reasoned that individual differences in imaging ability, as measured by the QMI, should evidence significant differential heart rates between the two groups to the stimulus material. All participants were presented with three pre-recorded scripts designed to produce two different emotional states and one neutral state. One scene depicted a fearful situation (a nurse preparing a syringe to draw the participant's blood), the other a racially disturbing scene (where the participant was accused of shoplifting by a group of white individuals), whilst the neutral scene described a pleasant walk in the park. No significant differences in heart rate response were found between vivid and weak imagers to the emotionally charged imagined conditions.

However, the authors note that "there was an attempt to develop extremely vivid scripts for all scenes" (p. 147). Indeed, post-experimental interviews revealed that participants had little difficulty in visualising the scenes. Thus it was concluded that the detailed recordings may have eradicated any differences between the two groups in imaging ability. The authors go on to suggest that future studies should

also vary the vividness of the stimulus material as well as measure individual differences in imaging ability. Only after the interaction of these variables is taken into consideration can the relative contribution of each be evaluated, and in turn support for the construct validity of Betts' QMI be gained or denied.

Hecker and Kaczor (1988) also attempted to find a relationship between self-reported imagery vividness as measured by the Revised Betts' QMI and increases in heart rate to imaged scenes. The 19 sports-oriented female participants in this study were required to imagine four different scenes designed to elicit different emotional responses: neutral (relaxing in a chair), action (working out in a gym), athletic anxiety (pressure to get a hit at the end of a softball game in a tie-break situation), and fear (being involved in a plane crash). The authors report that no significant correlations between QMI scores and heart rate change for any of the four scenes were found. However, it was noted that as a group, the participants in this study evidenced a higher than average imaging ability, indicating a relatively narrow range of scores. Given that this study employed such a restricted sample in imaging ability, coupled with the small sample size, significant relationships between imagery vividness and heart rate should not realistically have been expected.

Hirschman and Favaro (1980) selected 20 participants for a feedback-augmented study of heart rate control on the basis of their combined scores on the auditory and visual sub-scales on the Revised Betts' QMI. Participants were divided into two equal groups of high and low imagers, representing the upper and lower 40% of the QMI distribution. A light signalling system indicated to the participants when they were to attempt to increase and decrease heart rate. A Neuman-Keuls test revealed that, under the increase heart rate condition, high imagers ( $M = 5.4$ ) achieved significantly greater increases in heart rate than did low imagers ( $M = 2.5$ ):  $p < .05$  (two-tailed). However, under the decrease condition, no significant differences were found between the two groups. Post-experimental interviews revealed that under the increase heart rate condition, eight out of ten participants in the low imagery group and nine out of ten in the high imagery group reported using specific images (typically scenes of violence, sex, and anxiety) to increase their heart rate. Conversely, under the decrease heart rate condition, only three high imagery participants and two low imagery participants reported using specific imaged scenes as a strategy in an at-

tempt to reduce heart rate. Moreover, the authors report that in contrast to the scenes employed during the increase condition, the scenes that were imaged for the decrease condition were relatively nondescript.

It can be seen that only one of the above three studies found a significant positive correlation between imagery vividness as measured by Betts' QMI and heart rate. However, it must be pointed out that the authors of the two studies that failed to report significant findings also reported that methodological oversights might well have been responsible for their outcome. Thus, these studies can not be regarded as either a true test of the relationship between imagery vividness and heart rate or adequate tests of the construct validity of Betts' QMI.

**Salivation:** White (1978) employed the sublingual cotton swab technique to show that vivid but not weak imagers could control salivation by imaging the appropriate stimuli. The Revised Betts' QMI was administered to a pool of 289 male psychology undergraduates. From this pool, 30 participants were selected on the basis of their imaging ability: ten with the highest imagery vividness score, ten whose score fell around the sample mean, and ten whose QMI score indicated the weakest imaging ability within the sample.

Participants were asked to rank ten different foodstuffs in order of preference. From these rankings, each participant's 1st, 4th, 7th, and 10th preferences made up the experimental stimuli, which were randomised for order of presentation. It was hypothesised that the amount of salivation during a 30 second imaging period of the stimulus material would increase in a linear trend with the participant's ranking of his preference for the presented foodstuffs. However this hypothesis was only made for those with the ability to generate vivid imagery.

The results supported the hypothesis as it was found that vivid imagers, and to a lesser extent individuals of moderate imaging ability, produced more or less saliva to an imaged foodstuff in accordance with its preference ranking ( $F_{(6,81)} = 2.81, p < .05$ ). However, no such trend existed for weak imagers, as there was a tendency for these participants to salivate the same amount to all four of the imaged foodstuffs regardless of preference.

Thus, this study suggests that vivid imaging ability (or at least a moderate ability) is necessary for the voluntary control of salivation. That this hypothesis was tested and supported using scores obtained



on the Revised Betts' QMI provides additional support for the construct validity of the scale.

**Sexual Arousal:** Harris, Yulis, and Lacoste (1980) administered the Revised Betts' QMI together with a self-report measure of sexual arousability to 200 undergraduate students to try to determine whether a relationship exists between vividness of imaging ability and sexual arousability. Scores on the scales correlated positively and significantly for both males ( $r = .38, p < .001$ ) and females ( $r = .22, p < .05$ ). However, evidence for such a relationship (and support for the construct validity of the Betts) would have been more compelling had there been a physiological measure of sexual arousability with which to plot scores on Betts' QMI rather than reliance upon self-reports. For this reason subsequent studies looking at the relationship between self-report measures of sexual arousal and Betts' QMI (e.g., Meuwissen & Over, 1991) have not been reported.

More satisfactory evidence for the relationship between vividness of imaging ability and sexual arousability was provided by two papers by Smith and Over (1987, 1988) of their study of 66 males. In the first paper they report a significant positive relationship between increases in penile circumference whilst fantasising a self-chosen sexually oriented scenario and scores on the Revised Betts' ( $r = .44, p < .01$ ). Participants' subjective reports of sexual arousal during the fantasy also correlated significantly with their scores on the Betts ( $r = .32, p < .01$ ). The second paper analysed the data in more detail, breaking the Revised Betts' down into its seven sub-scales. Correlations between vividness of imaging ability and physiological indication of sexual arousal ranged from  $.26 (p < .05)$  on the kinaesthetic sub-scale to  $.36 (p < .01)$  on the visual sub-scale.

Koukounas and Over (1997) also used changes in penile circumference as a measure of increase in sexual arousal during imaged fantasy. However, rather than the fantasy being self-chosen, the 36 men were first shown an erotic film segment and then later asked to reproduce the events in the film imaginally. Although participants' scores on the Revised Betts' QMI correlated in the expected direction with increases in penile circumference during imaging ( $r = .24$ ), the relationship failed to reach significance.

**Habituation:** Drummond, White, and Ashton (1978) administered the cutaneous and kinaesthetic sub-scales of Betts' QMI to 14 psychology undergraduates in a study of the effects of imaging ability on habituation. There were seven vivid imagers and seven weak imagers. The experiment comprised of three stages. For the first stage, a 1000 Hz tone at 85 dB was presented every 20–40 seconds for a duration of 0.5 seconds until habituation was considered to be complete (zero response on two consecutive trials). Tones were also presented in the second stage, but in addition, participants were informed that a mild electric shock would accompany some of them (in fact no shocks were actually given). For the third stage, participants were asked to imagine receiving an electric shock every time the tone was presented. No significant differences were found between vivid and weak imagers during the shock threat condition. However, a *post-hoc* Neuman-Keuls test revealed that for the shock imaging condition, weak imagers habituated to the tone more rapidly than did vivid imagers ( $q_{(2,12)} = 5.45, p < .01$ ). Weak imagers also habituated to the tone significantly more quickly during the imaged shock condition than the shock threat condition ( $q_{(2,12)} = 7.58, p < .01$ ), whereas no such difference in habituation between the two conditions was found for the vivid imagers. Thus it can be seen that imaging instructions had a greater impact on vivid imagers than they did on non-vivid imagers.

Of the seven reported studies concerned with the relationship between some direct physiological measure and scores on Betts' QMI, four produced significant relationships — with two of the three non-significant studies acknowledging serious methodological flaws. Thus it can be argued that since Honorton's (1975) critique of the construct validity of Betts' QMI, there is evidence from various physiological correlates to suggest that the construct validity of the scale is adequate. Given the adequacy of the construct validity of the QMI we can accept the results of the analyses of restricted-choice and free-response studies detailed above which suggests that if anything, there is a negative relationship between voluntarily generated mental imagery and forced choice ESP task success. Nevertheless, it is claimed that the capacity for generating vivid mental imagery is important for the reception and detection of ESP material (White, 1964; Morris, 1977).

### **Voluntary mental imagery as a tool for focus**

If we are to understand how imagery generation is important for ESP reception and detection it would be useful to find out from psychic claimants how it is utilised in the process. White (1964) notes that before the more quantitative approach to ESP research, which began at the beginning of the 1940s, investigators tended to pay more attention to the percipient's mental state. Utilising the insights gleaned from these investigators and the experients themselves, White proposed a number of stages which she considered advisable for use by the percipient in subsequent psi research. These stages are as follows: "(1) Relaxation, (2) Engaging the conscious mind, (2A) The Demand, (3) The Waiting, the Tension, and the Release, and (4) The Way the Response Enters Consciousness" (1964, p. 27). The aim of this stage is to fix the attention of the conscious mind so as to prevent it from wandering. To achieve this, many of the experimenters and percipients in White's review encouraged the use of visualisation. This suggests that a high capacity for imagery generation would be efficacious for this stage as a generated image would allow the individual a point of focus. This, we feel, is an important point that has been overlooked by the research reviewed thus far. Although participants self-report measures of vividness of visual imagery was taken by researchers, the participant was not asked to make use of their imaging ability in a way advocated by many psychic claimants. Thus, there is a measure of ability but at no point is this ability been asked to be used when it is needed. There is little wonder that a relationship between this ability and ESP task success has not been forthcoming.

Using rhythmic photic stimulation to drive states of consciousness, Hume, Lawrence and Perez Navarro (2001, cited in Hume, 2003) conducted a free-response telepathy study to assess the relationship between task success and scores on the Randomised Betts'. The QMI was administered before the psi task. During a briefing period, participants were advised to focus upon an image of their choice and maintain their focus upon it during the photic session which was conducted with eyes closed. They were also asked to note any thoughts, images, or impressions they might have had during the session other than their chosen image. Overall, 24 hits were gained over 80 trials, giving a hit rate of 30% ( $\pi = .56$ , 95% CI = .45 to .68,  $z = 1.04$ ,  $p = .149$ , one-tailed). Thus, even though the hit rate was above mean chance expectation, there was

not statistically significant evidence for ESP. However a significant positive relationship was found between task success and voluntary imagining ability as measured by the Randomised QMI ( $r_{pb} = .263$ ,  $N = 76$ ,  $p = .022$ , two-tailed).<sup>9</sup>

In a partial replication of this research, Hume, Lok, More, McLaren and Haymes (2005) found a similar relationship between telepathy task success and scores on the Randomised QMI. Again, the Randomised QMI was administered before the psi task. In this study, the use of photic stimulation was abandoned and instead participants were simply asked to relax with their eyes closed while a red light was shone in their face. The participant was also advised to generate an image of a yellow rose and remain focused upon it and to note any spontaneous thoughts and impressions that they might have during the session. The rate for ESP task success was 40% with 16 out of the 40 participants obtaining a hit ( $z = 2.32$ ,  $p = .02$ , two-tailed,  $\pi = .667$ , 95% CI = .53 - .81). Thus there was evidence for an anomalous transfer of information. However, whilst similar to the previous study in terms of effect size and direction, the relationship between scores on the Randomised QMI and task success failed to reach significance ( $r_{pb} = .278$ ,  $N = 40$ ,  $p = .082$ , two-tailed).

#### *Comparing and combining the two 'imagery focus' studies*

When the two studies were combined, the overall hit rate was significantly above mean chance expectation (Stouffer  $z = 2.38$ ,  $p = .017$ , two-tailed). This effect remained significant when the studies were weighted by their  $N$  (Stouffer  $z = 2.05$ ,  $p = .040$ , two-tailed). The hit rates were also not significantly different from one another ( $z = .91$ ,  $p = .363$ , two-tailed).

The effect sizes for the relationship between scores on the Randomised QMI and task success were also combined. A significant positive relationship was evidenced between these two variables ( $r = .270$ , Stouffer  $z = 2.86$ ,  $p = .004$ , two-tailed). When the studies were weighted by their  $df$ , the effect size decreased marginally to approximately  $r = .265$  but still remained highly significant (Stouffer  $z = 2.84$ ,  $p = .005$ , two-tailed). The effect sizes were then contrasted to determine homogeneity of findings ( $z = .084$ ,  $p = .933$ , two-tailed) suggesting that these two studies are in agreement regarding effect size and direction of relationship. This suggests that the use of mental imagery to focus and

<sup>9</sup>Only 76 participants completed the QMI.

quieten the mind is conducive to ESP task success and that those with a strong capacity for generating vivid mental imagery are more likely to be successful.

### **Conclusions**

There is evidence within this review to suggest that Betts' QMI is a reliable and valid measure of individual differences in the ability to generate vivid voluntary mental imagery. We can see that in the case of restricted-choice studies, once studies with possible artefacts have been removed, if anything the capacity for vivid imaging may have a negative effect on ESP task success. For free-response studies (again with possible artefactual data removed) there appears to be no relationship between task success and vivid imaging ability, even though there is evidence to suggest that there is a significant psi effect across these studies. However, for free-response studies where the generation of an image is used as a tool for focusing the mind, there is a significant relationship between ESP task success and scores on Betts' QMI. Further, there is also evidence of psi across these studies.

Given the findings for these three types of methodology, it would appear that the relationship between ESP and imagery generation is not direct. It is merely part of a process leading to reception and identification of ESP material; there may be other techniques equally as efficacious. Therefore, although using imagery to maintain focus is not the only way to achieve ESP task success, it is perhaps the only way to evidence the effects of voluntary imagery generation in the process leading to success. If subsequent research is to investigate this relationship further, it is evident that participants must not only report their capacity for voluntary imaging ability, they must also use this ability during ESP tasks in an appropriate manner.

### **Acknowledgements**

The first author would like to thank the Society for Psychical Research for providing funding to support this research.

### **References**

Betts, G. H. (1909). *The distributions and functions of mental imagery* (Contributions to Education, No. 26). New York: Columbia University, Teachers College.

- Braud, W. G., & Wood, R. (1977). The influence of immediate feedback on free-response GESP performance during Ganzfeld stimulation. *Journal of the American Society for Psychological Research*, 71, 409-427.
- Drummond, P., White, K., & Ashton, R. (1978). Imagery vividness affects habituation rate. *Psychophysiology*, 15, 193-195.
- Evans, I. M., & Kamemoto, W. (1973). Reliability of a short form of Betts' Questionnaire on Mental Imagery: Replication. *Psychological Reports*, 33, 281-282.
- Gale, A., Morris, P., Lucas, B., & Richardson, A. (1972). Types of imagery and imagery types: An EEG study. *British Journal of Psychology*, 63, 523-531.
- George, L. (1981). A survey of research into the relationship between imagery and psi. *Journal of Parapsychology*, 45, 121-146.
- George, L., & Krippner S. (1984). Mental imagery and psi phenomena: A review. In S. Krippner (Ed.), *Advances in parapsychological research*, Vol. 4 (pp. 64-82). Jefferson, NC: McFarland.
- Harris, R., Yulis, S., & Lacoste, D. (1980). Relationships among sexual arousability, imagery ability, and introversion-extraversion. *Journal of Sex Research*, 16, 72-86.
- Hecker, J. E., & Kaczor, L. M. (1988). Application of imagery theory to sport psychology: Some preliminary findings. *Journal of Sport and Exercise Psychology*, 10, 363-373.
- Hirschman, R., & Favaro, L. (1980). Individual differences in imagery vividness and voluntary heart rate control. *Personality and Individual Differences*, 1, 129-133.
- Honorton, C. (1975). Psi and mental imagery: Keeping score on the Betts Scale. *Journal of the American Society for Psychological Research*, 69, 326-332.
- Honorton, C., Tierney, L., & Torres, D. (1974). The role of mental imagery in psi mediation. *Journal of the American Society for Psychological Research*, 68, 385-394.
- Hume, I. R. (2003). *ESP and mental imagery*. Unpublished PhD dissertation. Coventry University, Great Britain.
- Hume, I. R., Lok, A. W. C., More, D., McLaren, V. L., and Haymes, C. E. (2005). *Mental focus and ESP task success*. Unpublished manuscript.
- Juhasz, J. B. (1972). On the reliability of two measures of imagery. *Perceptual and Motor Skills*, 35, 874.
- Koukounas, E., & Over, R. (1997). Male sexual arousal elicited by film and fantasy matched in content. *Australian Journal of Psychology*, 49, 1-5.
- Meuwissen, I., & Over, R. (1991). Multidimensionality of the content of female sexual fantasy. *Behaviour Research and Therapy*, 29, 179-189.
- Morris, R.L. (1977). The airport project. In J.D. Morris, W.G. Roll, & R.L. Morris (Eds.), *Research in Parapsychology*, 1976. Metuchen, NJ: Scarecrow Press
- Palmer, J., & Lieberman, R. (1975). The influence of psychological set on ESP and out-of-body experiences. *Journal of the American Society for Psychological Research*, 69, 193-214.
- Palmer, J., & Vassar, C. (1974). ESP and out-of-the-body experiences: An exploratory study. *Journal of the American Society for Psychological Research*, 68, 257-280.
- Pleshette, G. (1975). Biofeedback regulation of EEG alpha, mental imagery and psi guessing performance. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in Parapsychology*, 1973 (pp. 23-24). Metuchen, NJ: Scarecrow Press.
- Rao, R. V., Rao, P. V. K., & Rao, K. R. (1977). Mental imagery and ESP as measured by Rao's Subsensory Psi Test (SPT). In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in Parapsychology*, 1976 (pp. 81-83). Metuchen, NJ: Scarecrow Press.

- Richardson, A. (1994). *Individual differences in imaging: Their measurement, origins, and consequences*. New York: Baywood Publishing Company, Inc.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, California: Sage Publications.
- Sargent, C. L. (1978). Interactions between visual imagery and psi. *Journal of the Society for Psychical Research*, 49, 805-823.
- Sargent, C. L. (1980). *Exploring psi in the ganzfeld*. New York: Parapsychology Foundation.
- Schecter, R., Solfvin, G., & McCollum, R. (1975). Psi and mental imagery. *Journal of the American Society for Psychical Research*, 69, 321-326.
- Sheehan, P. W. (1967). A shortened form of the Betts' Questionnaire upon Mental Imagery. *Journal of Clinical Psychology*, 23, 386-389.
- Shor, R. E., Orne, M. T., & O'Connell, D. N. (1966). Psychological correlates of hypnotisability in a special volunteer sample. *Journal of Personality and Social Psychology*, 8, 80-95.
- Smith, D., & Over, R. (1987). Correlates of fantasy-induced and film-induced male sexual arousal. *Archives of Sexual Behaviour*, 16, 395-409.
- Smith, D., & Over, R. (1988). Relationship between male sexual arousal and vividness of imagery in different sensory modalities. *Australian Journal of Psychology*, 40, 303-310.
- Smith, M., Tremmel, L., & Honorton, C. (1976). A comparison of psi and weak sensory influences on ganzfeld mentation. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in parapsychology, 1975* (pp. 191-194). Metuchen, NJ: Scarecrow Press.
- Stanford, R. G. (1993). Learning to lure the rabbit: Charles Honorton's process-relevant ESP research. In K. Ramakrishna Rao (Ed.), *Charles Honorton and the impoverished state of skepticism: Essays on a parapsychological pioneer* (pp. 129-175). Jefferson, NC: McFarland & Company, Inc., Publishers.
- Sutherland, M. E., & Harrell, J. P. (1986-87). Individual differences in physiological responses to fearful, racially noxious and neutral imagery. *Imagination, Cognition, and Personality*, 6, 133-150.
- White, K. D. (1978). Salivation: The significance of imagery in its voluntary control. *Psychophysiology*, 15, 196-203.
- White, K. D., Ashton, R., & Brown, R. M. D. (1977). The measurement of Imagery vividness: Normative data and their relationship to sex, age, and modality differences. *British Journal of Psychology*, 68, 203-211.
- White, K. D., Ashton, R., & Law, H. (1978). The measurement of imagery vividness: Effects of format and order on the Betts' Questionnaire upon mental imagery. *Canadian Journal of Behaviour Science*, 10, 68-78.
- White, R. A. (1964). A comparison of old and new methods of response to targets in ESP experiments. *Journal of the American Society for Psychical Research*, 58, 21-56.

# Do Human Fingers “See”? — “Finger-Reading” Studies in the East and West

Yung-Jong Shiah\* and Wai-Cheong Carl Tam\*\*

\*Koestler Parapsychology Unit, Department of Psychology,  
The University of Edinburgh, and Chien Kuo Technology University, Taiwan

\*\*Chung Yuan Christian University, Taiwan

## Abstract

*The “finger-reading” effect refers to successful touch identification of apparently flat targets on paper, where the participant is unable to see, or feel, any normal sensory cues to aid touch identification. Studies of this have been running for over 10 years in Taiwan. A quarter of children, after finger-reading training, appeared to be able to determine the identity of targets by means of directly touching a flat target varying in four different colours printed by an ink printer on paper. In the West, one study indicates that the fingers might read printing on paper without sight, while six studies find that fingers alone can discriminate colours on paper. However, a discussion of methodological issues follows, which points out the deficiency of well-controlled conditions in all the finger-reading studies reviewed. This leads to a conclusion that fraud has not been entirely ruled out – suggesting unreliable finger-reading results. In addition, this finger-reading effect has never been replicated. It is thus not safe to assume that exceptional abilities were in fact successfully measured. It is suggested that finger-reading needs to be further explored under stringent conditions, especially in children.*

## Introduction

Recently, attempts have been made to explore possible exceptional human abilities in Chinese societies. Si-Chen Lee, a professor of the

---

Correspondence details: Yung-Jong Shiah, Psychology Department, The University of Edinburgh, Edinburgh, Scotland, EH8 9JZ, United Kingdom. Email: s0239482@sms.ed.ac.uk. This paper is dedicated to Professor Robert L. Morris, the first author’s late supervisor, who sadly died before publication.



Electrical Engineering Department at National Taiwan University and also the President of the most prestigious university in Taiwan, began to focus on finger reading from 1993 by way of a developed training paradigm. The first author was a member of his research team from 2000 to 2002. Following this, since 2003, the first author has studied at the Koestler Parapsychology Unit to learn about applying more rigorous scientific methodology to the exploration of exceptional human abilities, with a view to looking at possible explanations for the finger-reading effect.

A quarter of children, after finger-reading training, appeared to be able to determine the identity of targets by means of directly touching a two-digit number or a complex character varying in four different colours printed by an ink printer on paper. Some Chinese children seemed able to do this when other senses such as vision were ruled out. This "touch" effect has also been reported occasionally in Western experiments for more than a century. According to Novomeysky (1965), the first report of finger reading was published in Russian scientific literature in 1898. Since then, several studies were conducted to explore this effect. The results of Western studies apparently showed positive results. However, poor experimental design was used in most of the studies (Gardner, 1996). In addition, many of the participants who claimed to have this ability were found to be cheating e.g., peeking at targets (Gardner, 1966). The last Western study of finger reading was undertaken in 1992. Since then, no further research about finger-reading effect has been conducted in the West.

In fact, the finger-reading effect is also subject to criticism in Taiwan (Du, 2005). Does finger-reading ability really exist? Were the successful studies merely replicating errors and were they open to fraud? Could it just be an example of a performance using tactile cues? Before addressing those questions, we had better take a look at finger-reading studies in the East and the West. In what follows, finger-reading studies in the East and West will be reviewed and their limitations will be scrutinized. Future research will be suggested.

#### *Studies of "finger reading" in Chinese society*

On 11th March 1979, a boy aged 12 was reported by the Sichuan Daily in mainland China as seemingly possessing an "ear reading" capability, i.e. he was able to recognise characters written on a piece of paper screwed into a ball and put into his ear (Chien, 1981; Eisenberg,

1985; Gardner, 1996). Since then, hundreds of Chinese children have been reported as appearing to possess this ability. Sometimes a folded paper involving Chinese characters was placed into children's hands or armpits. One of the more recurrent claims of possession of exceptional ability was for a finger-reading capability (Lee, 1998; Wang et al., 1989). Empirically, it was further reported that this ability could be induced by intensive training. For example, out of forty children of ages ranging from five to fourteen, 15 appeared to show this touch effect after between three and ten training sessions (Chen et al., 1989). In this training programme, children were instructed to use their fingers directly to touch a paper with written Chinese characters. It was even claimed that children seemed able to read characters within folded paper after more training. The children reported that the targets seemed to appear in their minds as a real visual image even when other senses, such as vision, were ruled out. The researchers assumed this to be a demonstration of something like extra-sensory perception (ESP) (Lo et al., 1989; Shao et al., 1982; Tien, 1994; Wang et al., 1989). ESP is a general term used for all paranormal abilities that cannot be explained by "rational" terms (Irwin, 2004).

This finger-reading effect caught the attention of Si-Chen Lee. He gathered a research team to study this touch effect and developed a four-days-a-week, two-hour finger-reading programme to study these phenomena in Taiwan from 1993 onwards (Lee, 1998, 1999, 2002, 2003; Lee & Chang, 2001; Lee, Chen, & Tang, 2000; Lee, Tang, & Kuo, 2004; Tang, Lee, & Hsu, 2000). Briefly, the research team conducted a variety of training and testing procedures, and found that children, aged between seven and thirteen, were the easiest to train. Si-Chen Lee trained adults at first as well, but they seemed to benefit little from this training process and failed to show any positive results. It appeared to be very hard for adults to learn how to visualise targets during the training process.

The training procedures can be illustrated briefly as follows: First, the children were given imagery exercises. The children were trained by letting them touch a paper directly which bore a two-digit number or a complex character printed in four different colours from an ink printer. This training included a "dark" condition in which the paper with its character was put into a dark bag where it could not be seen. Then they were asked to imagine that they could see the numbers, characters or words while touching them. The children were encouraged to practise

touching and visual imagery during this training process. The procedural training details will be covered and discussed later.

216 participants, aged seven to thirteen, were recruited from different elementary schools during the years 1996 to 2004. The average success rate at recognition (by  $p < .05$  criterion) by means of directly touching an unseen paper with a two-digit number or a complex character varying in four different colours was approximately 24% (41 out of the 173 participants who went through the whole training programme). The dropout rate was about 20% (43 participants). The major reason for leaving the training programme was that the children felt the programme was somewhat tedious and time-consuming.

The children for whom the techniques seemed to be successful reported that visual experiences had accompanied their successful trials. They reported visual images appearing as if from the real world. They reported seeing a "transparent screen" like a mist, with a floating patch or pattern overlaying their field of vision. Some of the children experienced the coloured targets as a distinct form of imagery like an "opaque screen" masking the normal visual image. The quality of the screen reported by participants seems important for this touch effect; for example, children appeared to recognise easily complex characters or other complex symbols after seeing an opaque screen. The experience of this opaque screen in the mind correlated highly with correct recognition. It might be trained by touching a complex character, producing a more complex visual image display in the children's minds. The shortest training time was only 20 minutes.

#### *Studies of "finger reading" in Western society*

Novomeysky (1965) conducted a study employing 80 participants and found that participants distinguished well between colours presented in pairs just by touching, without seeing them. After two or three weeks of exercises, one-sixth of participants learned to recognise five to seven colours just by touching paper. In 1919, the finger-reading effect was investigated by the French novelist, poet and dramatist Romain Rolland. Rolland's book was translated into English, entitled *Eyeless Vision*, in 1924. Rolland investigated French women who claimed they could read without seeing, being blindfolded (Duplessis, 1975; Gardner, 1966).

One piece of evidence for tactile-colour sensitivity was replicated by Nash (1969): participants significantly distinguished black and red

paper by touching without seeing ( $p < .05$ ). Later, in response to Gardner's (1966) criticism of the lack of control of the peeking problem, a head box constructed of 3/8-inch thick plywood was employed (Nash, 1971). This box fitted over the participants' heads and rested on their shoulders and came completely under the chin to fit snugly around the neck. He found the same positive results as in his previous study of 1969 ( $p < .001$ ).

A similar result was found among blind people. A 30 year old blind woman, who had been totally blind since the age of 18, was found to discriminate four colours on paper with a significant result ( $p < .001$ ) (Moss, Gray, Hubacher, & Bush, 1972). Both blind and normally sighted people were found to be able to discriminate colours by touch on paper (Duplessis, 1978).

### *Overall findings*

All studies reviewed suggest that fingers might be able to detect colours on paper. With respect to recognising printing, only one Western study, but all the Eastern studies found significant results.

One of the major differences between Western and Eastern studies is that Lee developed formal procedures targeted at children for developing finger-reading ability. According to this training paradigm, a visual experience accompanied with the correct answers was suggested to play the key role in helping participants successfully identify targets. This might indicate that reporting seeing a visual screen might be a good predictor of finger-reading ability. It is worthwhile investigating this claim. If finger-reading ability is real, one might expect its manifestations to be predictable. However, it should be noted that well-controlled conditions are of particular importance while conducting this investigation, as will be discussed later.

The other difference is that only children, and not adults were recruited as participants in the Eastern studies. As noted before, children seemed to perform better than adults did in Lee's finger-reading studies, revealing a reason for exploring ESP and children. Many parapsychological studies have focused on examining the relationship between adults and ESP, with little research examining children and ESP. In this regard, little is known about the topic of ESP in children. The cause of children performing better than adults did in Lee's studies is unknown. Indeed, over the years, researchers have noticed that children might be a potential group for demonstrating ESP ability (Bourgeois & Palmer,

2002; Rhine, 1965). Although it is not easy to conduct child parapsychological experiments, children might provide us with remarkable performances and phenomena (Alvarado, 2001). For example, some studies show a negative relation between ESP performance and aging (Bourgeois & Palmer, 2002; Shargal, 1987; Spinelli, 1977; Van Busschach, 1959).

### **Problems and limits of the previous studies**

#### *Methodological problems*

The very important issue of experimental controls developed and discussed in parapsychological studies provides a good checklist with which to examine finger-reading studies (Kennedy, 2004; Lamont & Wiseman, 1999; Milton, 1996; Morris, 1987, 1999, 2001; Steinkamp, Milton, & Morris, 1998; Wiseman & Morris, 1995).

**Methodological problems in the Eastern studies:** No satisfactory explanations of the phenomena were given, nor were the procedures described in sufficient detail in published reports done in Mainland China. Thus, the whole process cannot be evaluated. It is not clear whether the researchers ruled out fraudulent techniques such as those used in performance magic. For example, in 1981, children were caught peeking by scientists during finger-reading tests (Gardner, 1996).

Before examining Lee's finger-reading training paradigm, we should take a look at his training procedures. Lee's training procedures have never been mentioned in detail in any published journal. Most of the procedures depend on the first author's observations during the time when he worked with Si-Chen Lee, who helped to clarify some of the described procedures.

**Summary of Lee's finger-reading training procedures:** The stimuli consisted of 5cm × 8 cm rectangular pieces of paper. In the middle of each paper was a two-digit number in one of four different colours (black, green, blue and red) printed by an ink printer. There were two-digit numbers from 10 to 99. Confounding numbers, or "double chance numbers," such as 16 and 91, 19 and 61, 18 and 81, 66 and 99, 69 and 96, 68 and 89, 86 and 98, were excluded, so there were 76 numbers used in all. The trial samples were always prepared by a research assistant who did not participate in the finger-reading training process. They were

folded twice and all put into a big envelope. Each sample was only used once in all procedures. In Training 2, the digit was replaced by a Chinese character. Sometimes, for example, in some special conditions, the stimuli were drawn on a 5cm × 10cm or 3cm × 10cm rectangular piece of white paper (Lee, 1998; Lee et al, 2000). Written or printed on the paper was a Chinese character or an English word or a symbol or mathematical formula.

A specialised black bag, used for handling photographic negatives, was employed as a barrier against sensory leakage. Two cuffs are snugly fitted around the participant's forearms and the bag has two layers, each with its own zipper. Hardly any light could enter the bag, as was empirically shown by a light detector. The participant could move and feel around freely within the bag. The purpose of the bag was to prevent the participants, experimenter and co-experimenters from seeing or peeking at the targets.

A three-stage, specialised training procedure was used in the experiments. Participants were first required to participate in 'warm-up training', where they watched a 30-minute videotape describing this "touch" phenomenon, such as how to identify the target. First of all, participants were required to sit and close their eyes and breathe deeply with a calm and peaceful mind for at least ten minutes. Then participants were required to practise image-making. The experimenter showed an object, such as a red apple, to the participants who were asked to look at the apple very carefully and remember every detail of it. Then they closed their eyes to visualise the apple exactly as they perceived it. Next, they visualise the apple changing its colour three or four different times i.e., through green, blue and black.

Participants then moved onto training procedure 1 which involved directly touching a two-digit number. The experimenter usually drew ten samples randomly from the big envelope and put them on the co-experimenter's chair. Then, the co-experimenter clenched one sample in his or her fist and put it the bag, and then closed the zippers. Participants must not see the target during this process. Next, participants put their hands into the two sleeves of the black bag and the sleeves were tied up. Participants were then required to open the folded samples and use their fingers to feel the targets. Participants were asked to focus on touch and to imagine that they can see the numbers while touching. There were no time restrictions and participants were free to use whatever scanning force and speed they chose. They removed their hands

to write down the answer whenever they had told the co-experimenter what they saw, and the co-experimenter had recorded their response too. The co-experimenter then took out the training item from the black bag and showed the number to the participant. Thus, participants received feedback and the co-experimenter recorded if the participant’s response was correct. Usually, children could try 20 items in one session within two hours.

Finally, participants who had a statistically significant performance level and subjectively report experiencing a subjective visual experience, usually a transparent screen in their mind, were invited to attend Training procedure 2, which involved directly touching a complex target (a Chinese character). This training procedure is the same as the training procedure for directly touching a target (a two digit number), but the stimulus was now a Chinese character. Sometimes, in special conditions, one experimenter and several co-experimenters carefully watched the participant in Training procedures 1 and 2.

**Inadequate controls in Lee’s procedures:** The first issue in Lee’s procedures is the problem of randomisation, as weak randomisation procedures are considered a serious problem (Bierman, Broughton, & Berger, 1998; Brugger & Taylor, 2003; Diaconis, 1978). A target should be selected randomly from target pools. The experimenter randomly drew several samples from the envelope (samples pool) and gave them to each co-experimenter. Plainly, this randomisation is inadequate.

Sensory leakage is also an issue in Lee’s procedures, and this refers to participants obtaining information from sensory other than from extrasensory (Irwin, 2004). Usually, one co-experimenter worked with two participants, or sometimes three participants. The co-experimenter could not carefully observe each participant’s responses and behaviour. The authors suggest that at least one experimenter and one co-experimenter or more work with each participant.

The experimenter put samples on the co-experimenter’s chair and the participants cannot see the samples. Although the sample — a small piece of paper — was folded twice to prevent seeing or peeking, a remote possibility existed that the experimenter or co-experimenters might see the mark from the outside. The authors suggest that the samples should be put into an envelope before each trial.

The production of stimuli should be standardized in both proce-

dures. Detailed information on how targets are prepared should be given. A tactile cue might be present due to different printing quality, especially in written samples. The procedures had not been examined by an expert in detecting fraud, so they may be open to cheating.

There is also an issue surrounding the participants and potential recording problems. The authors suggest that participants who have a history of nerve or brain injury, finger trauma, or learning disability (including dyslexia), diabetes (because of associated peripheral neuropathy) and callouses on their finger tips should be excluded. These factors might affect tactile learning results (Goldreich & Kanics, 2003; Vega-Bermudez & Johnson, 2004).

Although over two hundred children have taken part in Lee's finger-reading training, their psychological traits and demographical background have not been studied. Such information might provide useful explanations for the finger-reading effect. After discovering which variables best predict the finger-reading effect, we could be in a position to discuss which assumptions or theories are closely related to explaining the phenomena.

It is not clear if records of participants were double-checked by at least two different researchers/co-experimenters to avoid calculation error. Only individual scoring was analysed and not all the participants' trials were reported. All the trials for each participant should be clearly noted, as well as the method of analysis.

**Methodological problems in the West:** Likewise, Western finger-reading studies did not provide fully detailed information of randomisation procedures. They exhibit the same problems with participants and recording as described above. Sensory leakage is also a serious problem. For example, blindfolds have been found to provide only a rather weak control (Gardner, 1996). Wearing a pair of blindfolds was used in Novomeysky's and Romans's studies but this still provided only a weak safeguard against cheating, because it was possible for blindfolded people to see down through tiny openings made by muscular contortions or eye twitching. For this reason, it is not clear that Romains' investigation ruled out cheating.

Claims that the finger-reading studies lacked sufficiently tight controls to rule out trickery were often reported, with peeking being an especially common problem. For example, according to Gardner's report



(Gardner, 1996), one 13 year-old boy in a 1937 study claimed that he could name playing cards without seeing them. However, J. B. Rhine, the famous parapsychologist at Duke University, tested this boy with opaque goggles and found him to be cheating by peeking over the bridge of his nose. In 1962, a 22 year-old Russian epileptic patient claimed to be able read while blindfolded, but she, too, was caught cheating by scientists. Also, in another study (Zubin, 1965), a 15 year-old girl claiming to have this kind of ability was tested. She wore a blindfold taped to her face around its edge and was found to have tensed the muscles in the areas of her blindfold to cause a very tiny opening allowing peeking down the side of the nose.

**Common problems and limits of the previous studies:** Actually, methodological issues are a very serious problem in all finger-reading studies. In addition to the described problems, not one of reviewed studies provided fully detailed information about its safeguards. For example, measures to prevent cheating, such as a possible access to target pools, changing experimental records and replacing targets, should be implemented. Details of the materials and how targets are kept secure between being taken out of storage and being used in experiments must be noted. Clearly, bad methodological design has been a major problem in all finger-reading studies. Still worse, some of the participants have figured out how to cheat.

Regarding the printing quality, it was suggested we could accurately identify touch recognition in terms of about three levels of intensity (Geldard, 1960). We can detect a very small difference of particle sizes with thresholds between .0024 and .0033 mm and the difference of ridge height thresholds was between .00095 and .002 mm (Miyaoka, Mano, & Ohka, 1990). Different printed colours might cause different levels of touch intensity, providing a tactile cue to detect different colours, especially when the participants only had to discriminate two colours on the same trial in Novomeysky's study. The details of how the samples were obtained were not fully noted in all previous finger-reading studies. In this case, the possibility of tactile cues cannot be excluded.

Finally, there is an issue surrounding replication. If the finger-reading effect cannot be replicated reliably, it will lose credibility. Many researchers (Alcock, 2003; Burns, 2003; Jeffers, 2003; Milton & Wiseman,

1999, 2001) have pointed out that no sufficient evidence has proved the existence of ESP. The results of ESP performance have been found elusive, weak, unreliable and lacking in quantity (Kennedy, 2001). This leads to other problems, such as unpredictability, lack of progress and failure to propose coherent explanatory theories (Alcock, 2003). Likewise, the finger-reading effect is now facing the problem of replication. No-one has replicated Lee's finding using his training paradigm.

The reviewed studies indicate that fingers might be able to recognise colours on paper, but are vulnerable to poor methodology as above. The methodological quality of a study is an important criterion for its inclusion in a meta-analysis (Rosenthal, 1995). For this reason, the authors suggest that all of these studies cannot be selected in any meta-analysis.

In summary, replication of the finger-reading effect with respect to recognising colours or print is wanting.

### **Incomplete potential mechanisms and explanations**

Attempts to explain the finger-reading effect have been made. One of the very important questions was "Can our skin see or perceive radiation?" For instance, perceiving light or radiation has been suggested as a possible normal explanation of the finger-reading effect. In one experiment (Barrett & Rice-Evans, 1964), the participants were given a dim and low-level visible light condition (.00012 lumens/cm<sup>2</sup>). It was of 3.5 times the intensity of the black condition. No participant showed a significantly improved performance.

To the best of our knowledge, retinal photoreceptors are only found in the eye's rod and cone cells, with an exception that photoreceptors, which contain light-absorbing photopigment, are found not restricted to just rod and cone cells in the salmon's eyes (Soni, Philp & Foster, 1998). It is suggested that, evolutionarily, pigment cells in the skin may be precursors of the photoreceptor cells in the eyes (Arnheiter, 1998). However, the human skin only unconsciously responds to light, especially the ultraviolet-B (UVB) light (290–320nm) and ultraviolet-A (UVA) light (320–400 nm) wavelengths, resulting in the production of vitamin D and thus affecting skin pigmentation. The level of skin pigmentation works to prevent UVB radiation damage (Slominski, Tobin, Shibahara, & Wortsman, 2004). Light-absorbing photopigment, reactive to 400-700nm, has not been found in human skin. The existing evidence indicates that human skin cannot "see".

Regarding detecting radiation, everything has its own radiation. For example, in a paper with printed characters, the printed targets and the paper involve different materials. Thus, it is logical to infer that they have different radiation levels, which might, perhaps, lead to different levels of radiation feedback. Fingers might be able to detect the differences in radiation reflected by colours. To test this hypothesis, several attempts have been made. A higher level (60 watt lamp) testing box comprising two compartments separated by a sheet of frosted glass was used (French, 1965). Then a stack of 72 cards was put in the box. Black and white were used for the cards. The participant put one hand inside the box to go through the pack of cards and then guessed its colour under two conditions: one with the light on and the other one with light off. No positive results were found. In another study (Passini & Rainville, 1992), blind and blindfolded participants were tested to see if they could discriminate colours on boxes in normal light condition, but the result did not support this idea.

Although many workers (Jacobson, Frost & King, 1966; Markous, 1966; Nash, 1969, 1971; Novomeysky, 1965; Youtz, 1966) support the hypothesis of that human fingers might be able to detect radiation, all studies exhibit methodological problems. In Nash’s and Novomeysky’s studies, their methodological problems are as above. It was not clear if the experimenter was ignorant of the targets used in Jacobson’s study. In Markous’s study, only three of six participants used an aluminium box to prevent peeking. Youtz has not yet published a full account of his work, though he did use a blindfold as a safeguard. Again, they all did not provide fully detailed information of the experimental process and safeguards.

This hypothesis has not been proved. Presumably, we need to investigate whether this finger-reading effect is measurable, then it could be appropriate to explore the basic properties of it in order to develop assumptions or theories to explain it.

### **Future research**

The finger-reading effect has never been proved. To solve this problem, the answer is simply to run finger-reading experiment under well-controlled conditions. The authors suggest adopting the paradigm originally developed by Si-Chen Lee to further explore this finger-reading effect, which now has been modified and is being tested by the authors.

The finger-reading procedures are developed from Chinese culture. One might ask whether it can be applied in Western culture? Needless to say, no studies of this issue have been undertaken. To answer this question, the authors would initially make the assumption that ESP might be an ability common to all humans. It is a good strategy that researchers observe what is going on when finger-reading studies are conducted in Western society.

If the finger-reading effect is true, the assumptions would be as follows:

1. Our fingers might be able to detect printing with a very low elevation, even a nearly zero elevation, probably through unknown functions in the fingers. This would be a new and astonishing discovery about sensory abilities.
2. The finger-reading effect might involve some new means of perception beyond those presently understood.
3. In fact, no one has produced any plausible or satisfactory explanation for the finger-reading effect or any new means of communication. The most difficult aspect is whether to attribute it to the first assumption or the second assumption. This effect might involve *both* exceptional tactile ability and some new means of communication.

With regard to the assumption one, the limit of relief recognition needs to be assessed first. For example, the finger-reading task used in Lee's studies was ink-printed text, which is in a range of 1–20 microns (.001–.02 mm in elevation). Usually, the paper absorbs most of the ink. One might expect that this ink-printed text is near zero in elevation. However, the true elevation of the text still needs to be precisely determined. To our knowledge, the relief recognition task and Braille reading are the most similar to this tactile touching task in finger-reading studies. In a previous study, an elevation of .5 mm has been shown to lead to correct recognition of letters in normal sighted adult people (Vega-Bermudez, Johnson, & Hsiao 1991). However, there are no studies of recognition using printed text and rarely have studies been conducted on the limits of relief recognition. In other words, the elevation between .49 mm and zero has not been explored so far. But there is some evidence that people may have abilities within this range. Braille characters are a good measure for spatial acuity, because they have been

devised to assess the ability to resolve fine spatial form (Craig & Johnson, 2000). It was proposed that Braille pattern recognition is based on shapes outlined by the dots (Loomis, 1981), which can be considered as a relief recognition. One previous study has shown that normal sighted people can distinguish Braille patterns of .3 mm in elevation, while people who are blind from an early age can identify Braille patterns of .2 mm in elevation (Grant, Thiagarajah & Sathian, 2000). Assuming that Braille pattern recognition is similar to a relief recognition, the authors hypothesise that a relief recognition of .3 mm may be discovered. The authors are currently testing these hypotheses. It is vital to explore the limits of relief recognition since the results of these experiments will serve three important functions. Firstly, the limits of the tactile recognition of alphanumeric figures can be determined, as previously research has concentrated only on the recognition of Braille figures. Secondly, the value below threshold of tactile relief recognition will be applied to produce the touching samples used in later experiments, with the aim of ruling out the possibility of tactile cues. Finally, it can be regarded as a control experiment comparing the later finger-reading training experiment, since the former experiment will not involve any training procedures.

Any new sensory function of fingers or a new means of communication will need to be reconsidered and further explored, if fingers identifying printings with an elevation much below threshold is found in later research. If it is real, further investigation into underlying mechanisms can be done later, as studying it can tell us about exploring exceptional performance and how to enhance this. Theories might then be developed to account for the finger-reading effect.

In summary, the authors suggest that the limits of tactile relief recognition needs be determined. The finger-reading ability needs to be further explored under well-controlled conditions, especially in children.

### **Acknowledgements**

Preparation of this article was supported by the Society for Psychological Research and Professor Si-Chen Lee. We are grateful to Wendy Martin and to anonymous referees for their helpful comments.

## References

- Alcock, J. E. (2003). Given the null hypothesis a chance - Reasons to remain doubtful about the existence of psi. *Journal of Consciousness Studies*, 10, 29-50.
- Alvarado, C. S. (2001). *Children in the history of parapsychology*. Paper presented at the Parapsychology Association 44th Annual Convention, New York, USA.
- Arnheiter, H. (1998). Evolution biology: Eyes viewed from the skin. *Nature*, 391, 632-633.
- Barrett, S. M., & Rice-Evans, P. (1964). Sensitivity of hands to visible light. *Nature*, 203, 993.
- Bierman, D. J., Broughton, R. S., & Berger, R. E. (1998). Notes on random target selection: The PRL autoganzfeld target and target set distributions revisited. *Journal of Parapsychology*, 62, 339-348.
- Bourgeois, R. L., & Palmer, J. (2002). *A computerized ESP test for children: Is age a factor for ESP?* Paper presented at the Parapsychology Association 45th Annual Convention, Paris, France.
- Brugger, P., & Taylor, K. I. (2003). ESP - Extrasensory perception or effect of subjective probability? *Journal of Consciousness Studies*, 10, 221-246.
- Burns, J. E. (2003). What is beyond the edge of the known world? *Journal of Consciousness Studies*, 10, 7-28.
- Chen, S. -L., He, M. -Y., Wang, T., Zheng, L. -M., Shao, S. -Y., Chen, Z. -C., et al. (1989). *The general problems of special sensory function* (Vol. 1). Tsn-Du: Si-Chun Education Publication (in Chinese).
- Chien, S. -S. (1981). Developing the basic researches of human science. *Journal of Nature*, 4, 483-488 (in Chinese).
- Craig, J. C., & Johnson, K. O. (2000). The two-point threshold: Not a measure of tactile spatial resolution. *Current Directions in Psychological Science*, 9, 29-32.
- Diaconis, P. (1978). Statistical problems in ESP research. *Science*, 201, 131-136.
- Du, L. (2005). Taiwan: New university president has links to paranormal research. *Science*, 308, 1852.
- Duplessis, Y. (1975). Jules Romain - Forerunner of current research on dermo-optic sensitivity. *Parapsychological Review*, 6, 20-21.
- Duplessis, Y. (1978). Dermal-optic perception. *Parapsychological Review*, 9, 22-27.
- Eisenberg, D. (1985). Second encounter: The sisters Wang. In *Encounter with Qi: Exploring Chinese Medicine* (pp. 79-90). London: Ebeneger Balis & Son Ltd.
- French, C. N. (1965). Tactile-vision: Thermal texture cues in the discrimination of black and white. *Nature*, 208, 1352.
- Gardner, M. (1966). Dermo-optical perception: A peek down the nose. *Science*, 151, 654-657.
- Gardner, M. (1996). Eyeless vision. In S. Gordon (Ed.), *The Encyclopedia of the Paranormal* (pp. 224-264). New York: Prometheus Books.
- Geldard, F. A. (1960). Some neglected possibilities of communication. *Science*, 131, 1583-1588.
- Goldreich, D., & Kanics, I. M. (2003). Tactile acuity is enhanced in blindness. *The Journal of Neuroscience*, 23, 3439-3445.
- Grant, A. C., Thiagarajah, M. C., & Sathian, K. (2000). Tactile perception in blind Braille readers: A psychophysical study of acuity and hyperacuity using gratings and dot patterns. *Perception & Psychophysics*, 62, 301-312.

- Jacobson, J. Z., Frost, B. J., & King, W. L. (1966). A case of demooptical perception. *Perception and Motor Skills*, 22, 515-520.
- Jeffers, S. (2003). Physics and claims for anomalous effects related to consciousness. *Journal of Consciousness Studies*, 10, 135-152.
- Kennedy, J. E. (2001). Why psi so elusive? A review and proposed model. *Journal of Parapsychology*, 65, 219-245.
- Kennedy, J. E. (2004). A proposal and challenge for proponents and skeptics of psi. *Journal of Parapsychology*, 68, 157-167.
- Lamont, P., & Wiseman, R. (1999). *Magic in theory*. Hertfordshire: University of Hertfordshire Press.
- Lee, S. -C. (1998). Mechanism of finger reading (the third eye) and physiological measurements. *Journal of Chinese Somatic Science*, 8, 105-113 (in Chinese).
- Lee, S. -C. (1999). Unification of mind and material: The macroscopic quantum phenomena. *Journal of Chinese Somatic Science*, 9, 124-128 (in Chinese).
- Lee, S. -C. (2002). The dialogue with the spirit. *Buddhism and Science*, 3, 68-79(in Chinese).
- Lee, S. -C. (2003). Traveling around the visible and invisible worlds. *Buddhism and Science*, 4, 32-44 (in Chinese).
- Lee, S. -C., & Chang, L. -S. (2001). The connecting model between keywords and information field in finger reading experiments. *Buddhism and Science*, 2, 60-77 (in Chinese).
- Lee, S. -C., Chen, C. -T., & Tang, D. (2000). The observation of extraordinary phenomena when special keywords were tested in the experiment on character recognition by fingers. *Buddhism and Science*, 1, 8-17(in Chinese).
- Lee, S. -C., Tang, D., & Kuo, H. (2004). The possible connection between information field and dark matter and dark energy. *Buddhism and Science*, 5, 47-58 (in Chinese).
- Lo, L. -R., Wang, T., He, M. -Y., Chen, Z. -C., Shao, S. -Y., Zheng, L. -M., et al. (1989). *The detecting process of the special sensory function on multistory samples. Creating Somatic Science (1)*. Tsn-Du: Si-Chun Education Publication (in Chinese).
- Loomis, J. M. (1981). On the tangibility of letters and Braille. *Perception and Psychophysics*, 29, 37- 46.
- Markous, W. L. (1966). Cutaneous color sensitivity: Explanation and demonstration. *Psychological Review*, 73, 280-294.
- Milton, J. (1996). Establishing methodological guidelines for ESP studies: A questionnaire survey of experimenters' and critics' consensus. *Journal of Parapsychology*, 60, 289-334.
- Milton, J., & Wiseman, R. (1999). Does psi exist? Lack of replication of an anomalous process of information transfer. *Psychological Bulletin*, 125, 387-391.
- Milton, J., & Wiseman, R. (2001). Does psi exist? Reply to Storm and Ertel. *Psychological Bulletin*, 127, 434-438.
- Miyaoka, T., Mano, T., & Ohka, M. (1990). Mechanisms of fine-surface-texture discrimination in human tactile sensation. *Journal of the Acoustical Society of America*, 105, 2485-2492.
- Morris, R. L. (1987). Minimising subject fraud in parapsychology laboratories. *European Journal of Parapsychology*, 6, 137-149.
- Morris, R. L. (1999). Experimental systems in mind-matter research. *Journal of Scientific Exploration*, 13, 561-577.
- Morris, R. L. (2001). Research methods in experimental parapsychology: Problems

- and prospects. *European Journal of Parapsychology*, 16, 8-18.
- Moss, T., Gray, J. I., Hubacher, J., & Bush, B. (1972). Skin vision and telepathy in a blind subject. In W. G. Roll, R. L. Morris & M. J. D. (Eds.), *Research in Parapsychology 1972* (pp. 82-84). Metuchen, New Jersey: The Scarecrow Press.
- Nash, C. B. (1969). Cutaneous perception of color. *Journal of the American Society for Psychical Research*, 63, 83-87.
- Nash, C. B. (1971). Cutaneous perception of color with a head box. *Journal of the American Society for Psychical Research*, 65, 83-87.
- Novomeysky, A. S. (1965). The nature of the dermo-optical sense. *International Journal of Parapsychology*, 7, 341-367.
- Passini, R., & Rainville, C. (1992). The dermo-optical perception of color as an information source for blind travelers. *Perceptual and Motor Skills*, 75, 995-1010.
- Rhine, L. E. (1965). *Hidden Channels of the Mind*. New York: William Morrow.
- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, 118, 183-192.
- Shao, S. -Y., Wang, T., He, M. -Y., Lo, L. -R., Chen, Z. -C., Chen, S. -L., et al. (1982). Special sensory function: The image appearing process of recognizing the folded papers. *Journal of Nature*, 5, 274-275 (in Chinese).
- Shargal, S. (1987). Children's ESP scores in relation to age. In D. H. Weiner & R. D. Nelson (Eds.), *Research in Parapsychology 1986* (pp. 51-53). Metuchen, New Jersey: Scarecrow Press.
- Slominski, A., Tobin, D. J., Shibahara, S., & Wortsman, J. (2004). Melanin pigmentation in mammalian skin and its hormonal regulation. *Physiological Reviews*, 84, 1155-1228.
- Soni, B. G., Philp, A. R., & Foster, R. G. (1998). Novel retinal photoreceptors. *Nature*, 394, 27-28.
- Spinelli, E. (1977). The effects of chronological age on GESP ability. In J. D. Morris, W. G. Roll & R. L. Morris (Eds.), *Research in Parapsychology 1976* (pp. 122-124). Metuchen, New Jersey: Scarecrow Press.
- Steinkamp, F., Milton, J., & Morris, R. L. (1998). A meta-analysis of forced-choice experiments comparing clairvoyance and precognition. *Journal of Parapsychology*, 62, 193-218.
- Tang, D., Lee, S. -C., & Hsu, L. -C. (2000). Comparison between normal color sense and ESP color sense in color adaptation task: A case study. *Journal of Chinese Somatic Science*, 10, 21-28 (in Chinese).
- Tien, W. -S. (1994). Special vision and eyes. *Journal of Chinese Somatic Science*, 4, 4-8 (in Chinese).
- Van Busschhach, J. G. (1959). An investigation of extrasensory perception in school children. *Journal of Parapsychology*, 17, 210-214.
- Vega-Bermudez, F., & Johnson, K. O. (2004). Fingertip skin conformance accounts, in part, for differences in tactile spatial acuity in young subjects, but not for the decline in spatial acuity with aging. *Perception & Psychophysics*, 66(1), 60-67.
- Vega-Bermudez, F., Johnson, K. O., & Hsiao, S. (1991). Human tactile pattern recognition: Active versus passive touch, velocity effects, and patterns of confusion. *Journal of Neurophysiology*, 65, 531-546.
- Wang, T., He, M. -Y., Chen, S. -L., Zheng, L. -M., Shao, S. -Y., Chen, Z. -C., et al. (1989). *The appearing image process of the special sensory function*. *Creating Somatic Science* (1). Tsn-Du: Si-Chun Education Publication (in Chinese).



- Wiseman, R., & Morris, R. L. (1995). *Guidelines for testing psychic claimants*. Hertfordshire: University of Hertfordshire Press.
- Youtz, R. P. (1966). Dermal-optical perception. *Science*, 152, 1108.
- Zubin, J. (1965). Dermooptical perception: A cautionary report. *Science*, 147, 985.

# The Effect of Weak Magnetic Fields on a Random Event Generator: Reconsidering the Role of Geomagnetic Fluctuations in MicroPK Studies

Paul Stevens

Koestler Parapsychology Unit, Department of Psychology  
The University of Edinburgh

## Abstract

*This exploratory study looked at the effect of geomagnetic fluctuations on a random event generator (REG) commonly used as a target system in microscopic psychokinesis (microPK) studies. It is suggested that the geomagnetic fluctuations may themselves be considered as a signal which affect the REG through a process of stochastic resonance. The effect of different geomagnetic activity "signals" on REG output was thus measured during simultaneous application of different levels of magnetic field "noise". As might be expected if stochastic resonance were involved, the REG output ( $z^2$ ) for each of the geomagnetic activity "signals" showed a maximum at a non-zero level of magnetic field "noise". This level was consistent for 3 of the 4 geomagnetic "signals" used (maximum for  $K = 0, 2, 3$  when noise field was  $1 \mu\text{T}$ ), with the highest geomagnetic "signal" ( $K = 4$ ) showing a separate peak for a  $25 \mu\text{T}$  noise field. As the noise fields were applied in a random schedule over non-consecutive days, this would appear to be a robust finding. The role of magnetic fields should therefore be considered, either as a direct effect on REGs or as part of a microPK mechanism, in any theories attempting to model microPK effects.*

## Introduction

Many studies have reported results (see Radin & Nelson, 1989 for summary) which appear to support the existence of microscopic psychokinesis (microPK), wherein a statistical relationship is found between an agent's mental state or behaviour and the activity of a target system. The target system is usually some form of random event generator (REG), the most commonly used type being based on the electronic noise inherent in semiconductors, amplified and digitised to give a output sequence of random numbers. To date, no mechanism has been identified by which such effects might be explained. In the absence of an accepted microPK theory, researchers have attempted to look for physical variables that might affect experimental results. Knowledge of such variables could eventually help theorists identify potential microPK mechanisms, and in the short term might lead to solutions to, or at least explanations for, the lack of replicability that plagues parapsychological research.

One potential physical variable is the amount of activity of the Earth's (geo) magnetic field, which cover a range of frequencies from 0 to 1 kHz (Parkinson, 1983). At one end of this range is the static dipole field generated by the Earth's internal dynamics; at the other are the rapid changes due to interactions between the charged layer of the atmosphere (the ionosphere) and the ground. Between these are the pulsations due to charged particles entering the geomagnetic field from space. The overall activity of the geomagnetic field is monitored using sensitive magnetometers and presented either in terms of the ongoing absolute magnitude<sup>1</sup> or the relative change in a given time period. The most commonly used measures in parapsychology studies are geomagnetic indices (such as the 3-hourly K, Kp and ap indices, or the daily Ap index) which represent standardised local and global disturbance levels of the geomagnetic field.

Many studies have reported a correlation between the effects measured in perceptive psi experiments and geomagnetic field activity, with the general finding that success, however it is defined, decreases as geomagnetic activity increases. For example, apparent extrasensory perception was inversely related to global magnetic activity while the participant was in a dream state (Persinger & Krippner, 1983) and also while in

---

<sup>1</sup>Technically this is actually a quantity known as *flux density*, but for the purposes of this discussion, magnitude is a close enough concept.

the mild sensory-deprivation Ganzfeld state (Dalton & Stevens, 1996). However, the picture is less clear with microPK experiments. Gissur- arson (1992) found a significant negative correlation between the direc- tion of REG output and local field fluctuations, while Nelson and Dunne (1986) reported a non-significant positive correlation between the global activity and REG output.

This relationship, if it represents a real effect, might occur in a va- riety of ways, but in general it seems that the physical fields (primarily magnetic) associated with geomagnetic activity measures must interact either with the microPK agent, the microPK mechanism or the target system. The first case, that geomagnetic field activity affects the psi agent, has some merit in that there is an increasing body of research showing that weak magnetic fields in the geomagnetic frequency range can alter human behaviour (Stevens, 2001; Bell et al, 1994; Warnke, 1994). However, without knowing what it is the psi agent is doing to bring about a microPK effect, it is difficult to evaluate this possibility. It is also hard to conceptualise how sensitivity to magnetic fields (a recep- tive ability) might relate to the apparently emissive ability of microPK.<sup>2</sup> The second case, that geomagnetic field activity interacts with the mech- anism underlying psi effects is even more difficult to evaluate in that the mechanism has not yet been identified. The third case, that the field interacts with the target system itself is an intriguing idea in that, if a magnetic field can affect the functioning of a target system, then this suggests that such fields could also be involved in the production of microPK effects. In this paper, I will look in more depth at this third possibility.

### *Magnetic Fields and microPK*

That magnetic fields might be involved in some way with microPK is not a widely studied idea, mainly due to the small magnitude of ge- omagnetic or biologically-generated magnetic fields in comparison to man-made sources. However, there have been a few studies which have addressed this idea. Chauvin and Varjean (1990) found that application of a weak magnetic field to a random mechanical cascade device was associated with a reduced mean during microPK trials. Stevens (1999) conducted a remote microPK study where the REG output showed a

---

<sup>2</sup>This is true only if microPK represents an action of the agent on the target system, rather than the extrasensory selection process that Decision Augmentation Theory (May et al, 1995) adherents would suggest. However, if microPK is indeed a perceptive ability akin to ESP, then this does not explain why the geomagnetic relationship sometimes shows the reversal in direction.

positive correlation with a 3 hourly global geomagnetic activity index (ap) in both microPK trials and in control data. That both studies found a correlation between microPK success and ambient magnetic activity suggests that there is an interaction between the agent's attempts to influence the target system and that field. However, the Stevens study also suggests that the magnetic field must directly affect the REG, as a significant correlation, albeit weaker, was also found in the control data. To confuse the matter further the two studies found correlations in opposite directions: the Chauvin and Varjean study found decreased means with the applied field, whereas Stevens found increased means for increased geomagnetic activity.

Part of the confusion could lie in the way the interaction occurs. By reporting correlations between (geo) magnetic field activity and REG output, the implicit assumption is that it is a linear relationship. But the lack of consistency could imply otherwise, suggesting a more complex form of interaction.

One concept that may be useful is the phenomenon of *stochastic resonance* (SR) wherein a noisy system is driven by an external signal that would be normally be considered too weak to affect it (Wiesenfeld & Moss, 1995; Gammaitoni et al, 1998). Some characteristics of that signal (such as amplitude, signal-to-noise ratio, coherence, or other measure of performance) are, counter to intuition, actually improved by the presence of the noise. Essentially, the noise randomly boosts the weak signal by sometimes giving it enough extra energy to have a significant effect. This means that, with SR, a system becomes a more sensitive detector of a weak signal as more noise is added, at least up to a point: it is optimally sensitive at some non-zero level of input noise. If some measure of the Signal to Noise ratio is plotted against the input noise, this results in a characteristic profile (see figure 1).

### *Stochastic Resonance*

It should be noted that SR is not an esoteric phenomenon rarely found in nature but one which appears in many physical systems, from the periodic recurrences of Earth's ice ages (Wiesenfeld and Moss, 1995) to quantum tunnelling effects (Grifoni et al, 1996). It has also been observed in semiconductor diodes similar to those used as the underlying noise source in many REGs (Jung & Wiesenfeld, 1997).

So how might this help in understanding a relationship between magnetic fields and microPK? As mentioned earlier, we know that mi-

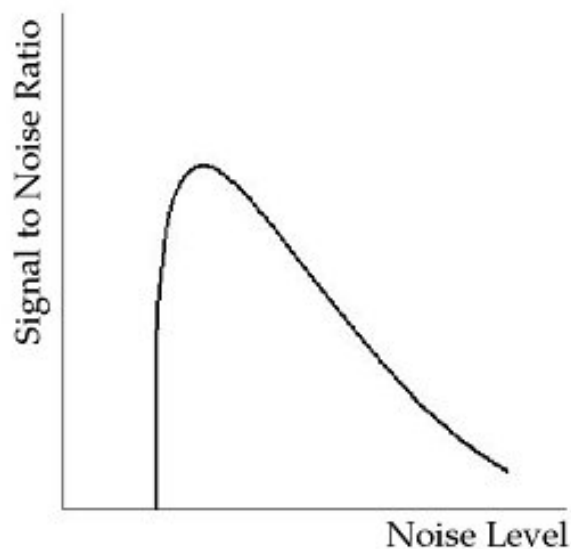


Figure 1. Characteristic profile of a system exhibiting stochastic resonance, showing a maximum value at a non-zero level of noise

croPK effects alter as magnetic field variance increases, whether that field is geomagnetic or artificially generated. A further clue comes from the Stevens (1999) study, which also found a significant, albeit weaker, correlation with geomagnetic activity in the REG control data when there was *no microPK attempt* i.e., the magnetic field alone appears to affect REG activity and so it makes more sense for it to be conceptualised as a “signal”<sup>3</sup> than as “noise” as the latter would not give the observed consistent/statistically significant bias to the REG output. Given the magnitude of variation for the geomagnetic field is extremely small — typically less than 10 nT — we would not expect them to have a noticeable direct effect (e.g., current induction) on the types of REG used. But SR tells us that even small signals can interact with a simultaneous source of noise to produce an effect which would not be seen if only the signal or the noise were present. So if we view the magnetic field variations as our signal, then we need only to find a suitable source of noise to look at the system in terms of SR. Obviously REGs have an inherent amount of noise as this forms the basis of the random event generation, so this alone might give an SR effect to a geomagnetic signal —

<sup>3</sup>The definition of a signal can get complex and to some extent depends on the context in which the term is used. For the purposes of the arguments presented here, a signal may be seen as variations in some physical process which conveys information. Thus the truly random variations generated by quantum processes within an isolated REG would not contain information or be considered to be a signal. The ordered oscillations of the geomagnetic field or of an artificially-generated magnetic field would contain information (e.g., the frequency of the field) and so could be seen as a signal.

something that should be taken into account in future geomagnetic correlation analyses of studies involving an REG. This would mean that, rather than the usual conceptualisation of the geomagnetic field as interference, it could actually be playing a major role in producing the observed effects on the REG under a microPK protocol: some aspect of the changes in the geomagnetic field comprises a signal that the REG system is sensitive to. Any effects are usually small compared to the inherent variability of the REG output, but when a third factor is added — a source of external noise — the geomagnetic signal's effects are enhanced i.e., a good working model might be:

$$\textit{Signal (Geomagnetic variation)} + \textit{Noise} = \textit{REG output measure}$$

So what would constitute a source of signal-amplifying noise in the system? It could come from a variety of changes in the natural environment to which the REG is sensitive, such as cosmic rays events or tectonic stress electromagnetic fields; it could also be artificial, due to electrical devices in the vicinity of, or even part of, the study. Of most interest to researchers interested in microPK work, it could potentially be of biological origin — the exogenous fields generated by the ostensible microPK agent's heart, brain or muscle activity (which, for example, generate magnetic fields in the pT to nT range, at a range of frequencies below 100 Hz). Whatever the source, consistent effects will only be seen under the right combination of conditions, where the properties of the noise are resonant with the properties of the signal and the target system is itself sensitive to the amplified signal. It should also be noted that, just because the GMF signal is magnetic in nature, this does not require the noise to be magnetic too, as SR has been observed where signal and noise are dissimilar in nature (e.g., Richardson et al, 1998), but the simplest model would be to start with an all magnetic field interaction, especially as we are concerned with an electronic system.

So this could suggest an alternative way of looking at the effects seen in some microPK studies: the REG essentially acts as a detector of weak magnetic fields. If this is the case, then plotting the output of a REG for different levels of weak magnetic field “noise” in the presence of various geomagnetic ‘signals’ should at least show the characteristic SR signature shown in figure 2, indicating whether this approach is a useful one to pursue.

## Method

### *Procedure*

An Orion serial port REG (Orion Technologies<sup>4</sup>) connected to a PIII 600 MHz PC was placed in the centre of a pair of .20m radius copper-wire wound coils in the Helmholtz condition (parallel coils with the centre-to-centre distance equal to the radius). Coil input currents were produced from twin power supply units (Coutant model LA200.2), with the frequency and amplitude being determined by the signal characteristics from the PC's sound card output (Creative SoundBlaster Live! Value). Control periods involved identical computer operation to experimental periods, but the output to the sound card was an empty file (i.e., zero amplitude). Field strength had been tested prior to the experimental sessions using a flux-gate magnetometer (model 428C, Applied Physics Systems, Mountain View, CA) for all but the pT work, as these fields were well below the sensitivity of the magnetometer.

An earlier unpublished exploratory study had been conducted looking at the possibility of direct effects of a number of artificially generated geomagnetic-magnitude fields on the REG, compared to the null-field control periods. For the current analysis, it was *a priori* decided to use the data from this study which covered fields with a sinusoidally modulated frequency of 40 Hz at magnitudes of 1, 25 and 50  $\mu$ T. This frequency was originally chosen as it corresponded to a frequency which is present throughout human brain-wave records (the "thalamo-cortical loop") and it was also the first harmonic of the fields used in an earlier study which successfully found effects of weak magnetic fields on human physiology (Stevens, 2001).

For any one session, the PC repeatedly selected (using the Microsoft QuickBASIC v4.5 pseudo-random algorithm) one of four conditions — corresponding to either of the 3 fields or the control condition. Selection probability was weighted so that there were approximately as many control periods as experimental. One session lasted for approximately 10 hours, wherein the PC took 1000 samples each of 175 bits at random intervals (between 1 and 30 seconds). Each 1000 samples (one trial) was reduced to a single mean and this was saved to disk along with information about which field was present. Experimental trials were randomly interspersed with control trials, taken over 7 non-consecutive days chosen purely due to the practicalities of when the experimental

<sup>4</sup><http://rng.interact.nl/>



rooms were unused by other studies. Of the 3535 experimental trials, 1155 were in the 50  $\mu\text{T}$  field, 1150 in the 25  $\mu\text{T}$  field and 1230 in the 1  $\mu\text{T}$  field. There were 4229 control trials.

As I was also interested in looking at field strengths comparable to those generated by a biological system, further data were collected comparing the REG output under the null-field control condition to a 1 pT (the magnitude of the field generated by the human brain in a normal waking state: Ioannides, 1994) 40 Hz sinusoidal magnetic field. As a much weaker field was used which required a finer degree of control, the PC this time pseudo-randomly selected to turn on or off the input current to the coils. This current was produced from a digital frequency generator, pre-calibrated to produce a 40 Hz 1pT field. This time, one session lasted for 2.5 hours, wherein the PC took 1000 samples each of 200 bits at random intervals (between 1 and 30 seconds), these 1000 samples representing one trial. This trial was then reduced to a single mean value and saved to disk along with information as to whether the field was on or off. Of the 2000 trials recorded over 10 non-consecutive days, 1004 were in the experimental condition and 996 were in the control condition.

### *Analysis*

A measure was needed which would represent an REG response to the effect of the geomagnetic signal. For ease of comparison to microPK studies, it was decided to use the square of the z-score i.e., a measure representing the absolute magnitude of the deviation from expected mean values for the number of events generated by the REG. So for each of the session means for each field type, the mean values were converted into squared z-scores (i.e., expressed in terms of their squared standard deviation). This removed the direction of any effect observed as this was not currently relevant and instead gave an indication of the magnitude of change in REG activity.

Geomagnetic field activity measures were then taken from the British Geological Survey's geomagnetic monitoring station, sited 40 miles away at Eskdalemuir, UK via their website at [http://www.nmh.ac.uk/gifs/on\\_line\\_gifs.html](http://www.nmh.ac.uk/gifs/on_line_gifs.html). The index taken was K, a baseline-corrected measure which represents the largest range of local activity measured in a 3 hourly period. The K index ranges from 0 to 9, with each digit indicating activity which is approximately a factor of 2 greater than the previous digit (Parkinson, 1983). K was chosen

rather than any other geomagnetic index as it better reflected changes in the local field activity than global indices.

## Results

Columns 3-6 of Table 1 show the squared z-scores representing the mean output of the REG under each type of magnetic field, each of the columns representing the level of geomagnetic field activity (based on the K-index) under which the REG data was collected. The descriptor column gives some indication of naturally occurring magnetic fields which would have the same magnitude as the applied field.

Table 1: Output of random event generator ( $z^2$ ) for different levels of magnetic field “noise” (nT), split by geomagnetic activity “signal” (K-index)

Magnitude of Magnetic Field “Noise” (nT)	REG Output ( $z^2$ )			
	K = 0	K = 2	K = 3	K = 4
0	$5.84 \times 10^{-4}$	$2.22 \times 10^{-3}$	$5.06 \times 10^{-4}$	$2.45 \times 10^{-4}$
0.001	0.00	$3.75 \times 10^{-8}$	$1.08 \times 10^{-7}$	$5.44 \times 10^{-8}$
1000	$1.34 \times 10^{-2}$	$1.56 \times 10^{-2}$	$1.12 \times 10^{-2}$	$7.01 \times 10^{-4}$
25,000	$2.59 \times 10^{-3}$	$7.64 \times 10^{-4}$	$1.50 \times 10^{-3}$	$2.06 \times 10^{-2}$
50,000	$6.49 \times 10^{-8}$	$4.91 \times 10^{-8}$	$1.34 \times 10^{-7}$	$8.64 \times 10^{-4}$

Values in table 1 were then plotted (see figure 2) and found to show multiple peaks at 1000nT and a single peak at 25000nT. The former are found for K values of 0, 2 and 3, whereas the latter peak is found only for the highest K value of 4. Best fit curves have been added purely to aid identification of K-index profiles and should not be taken to indicate continuous distributions.

## Discussion

This study suggested that the REG target system used in many microPK studies might be a sensitive detector of geomagnetic fluctuations, possibly through a process of stochastic resonance with intrinsic and/or external noise. To test the basic idea, an artificial source of noise was generated via Helmholtz coils, comparing REG output under different levels of magnetic field “noise” in response to different geomagnetic activity “signals”. As might be expected under a case of stochastic resonance, the REG outputs showed maximum deviation at a non-zero level of added noise. This level was consistent for 3 of the 4 geomagnetic signals used (maximum for K = 0, 2, 3 when noise field was 1  $\mu$ T), with

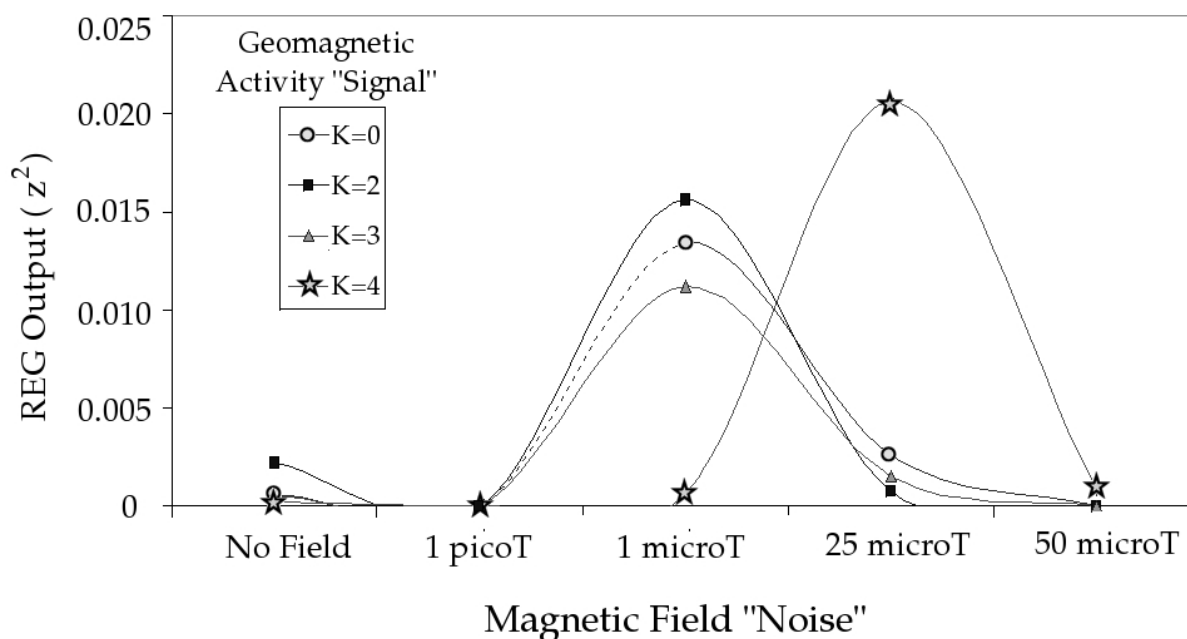


Figure 2. Variation of Random Event Generator output ( $z^2$ ) as the applied magnetic field magnitude (the *noise*) is changed. Each line represents a different level of geomagnetic field activity (the *signal*) in terms of the standard K-index

the highest geomagnetic signal ( $K = 4$ ) showing a separate peak at  $25 \mu\text{T}$ . As the noise fields were applied in a random schedule over a number of different days, this would appear to be a robust finding due to a real interaction and not simply due to chance variations. The effect size, based on  $(z \cdot N^{-0.5})$ , for the largest effect found in this study (the  $25 \mu\text{T}$  field at  $K = 4$ ) is around  $8 \times 10^{-3}$ . For the 1 pT (the "human brain" magnitude) field, it is  $1.3 \times 10^{-5}$ . For comparison with microPK studies, the meta-analysis by Radin and Nelson (1989) found an overall effect size of around  $3 \times 10^{-4}$ . Thus it does appear that geomagnetic field fluctuations in combination with external field can cause effects of a magnitude comparable or exceeding those seen in microPK studies. Even in the absence of an external noise field, the maximum effect size (for  $K=2$ ) was  $6.9 \times 10^{-4}$ , presumably (assuming an SR process) based on the signal interacting with the intrinsic noise of the REG itself.

Does this have implications for REG-based microPK work? It certainly indicates that there are physical factors that should be taken into account, and implies that magnetic fields can play a role (possibly including those generated by the human body) in producing REG deviations of a magnitude comparable to those seen in microPK studies. In

studies where the independent measure is simply a general biasing of the REG output, it could mean that the effects actually have little to do with the ostensible microPK agent — they just happened to be part of a study that coincided with high geomagnetic activity. This would be especially true of any studies which make use of a theoretical control baseline rather than one empirically determined under similar experimental conditions. However, it cannot be used to explain away the many studies where the REG output shows a correlation with agent intention (Radin & Nelson, 1989) or where the microPK effect is much greater in magnitude than any geomagnetic correlation seen in concurrently-collected control data (Stevens, 1999). However, there remains an option which might be involved in some instance of microPK that has not previously been considered: that microPK itself acts as source of noise to selectively enhance a geomagnetic signal. This is in many ways a simpler proposition than the more obvious choice of considering microPK to be a signal: it would not require any degree of fine control on the part of the agent but would instead need a relatively gross alteration in their emitted noise (which could conceivably be as simple as the magnetic fields relating to cortical arousal). Such a concept may be especially useful in instances where there does not appear to be any intention to cause a microPK effect on the part of the agent (e.g., Stanford, Zenhausern & Dwyer, 1975) but where the REG still appears to be affected when subjects are present.

Alternatively, it would also be possible to incorporate the ideas presented above into a selection model such as DAT. In that case, the selection would not necessarily be just of fortuitous “extra-chance” sequences in the random bit-stream but instead could be the selection of times when the signal and noise fields are suitable to produce the desired output. Alternatively, a blend of selection and influence could operate, with the signal fields being selected for and the noise field generated by the microPK agent.

Obviously this study was exploratory in nature: much of the data was originally collected with the idea of simply looking at direct effects of weak magnetic fields on the REG but, on becoming interested in the phenomenon of stochastic resonance, I decided to look for SR-like effects and added in a further data-set to cover bio-range magnitude fields. Because of this, and as a first foray into a new area of research, the current study has several limitations which should be addressed in any further research. Firstly, the applied magnetic fields all showed full-

amplitude variation and so aren't necessarily realistic when compared to natural equivalents. While the magnitudes of the noise fields used are encountered in the environment on a daily basis, the associated frequencies may be different. For example, the main component of the geomagnetic field has a magnitude of around 50,000nT, but this is static. Variations do occur in a range covering 40 Hz, but the associated magnitudes for these are much smaller (.05 nT). Fields of 1  $\mu$ T are commonly found in an urban environment but are more likely to be at frequencies of 50Hz (the mains electricity operating frequency). However, remember that the noise field is just that: noise. If SR is involved, then the frequency of the noise field is less important than the magnitude as it is just giving a "boost" to the signal at intervals, whatever those intervals are.

Secondly, the K-index used as a measure of the geomagnetic "signal" is a very gross measure and can really only give a rough indication of the variability of the ambient fields over a large area. Any site-specific variations (man-made or natural) are unlikely to be reflected in this. A better measure where possible would be to take local measurements using a magnetometer.

Thirdly, the design of the REG itself is not ideal for this kind of work. The Orion REG is based on two NAND-gated Zener diodes. This not only drastically reduces the amount of information available about the activity of the source randomness (the variations in electrical currents within the diodes due to electron tunnelling, thermal effects and surface noise) down to a binary digit (i.e., greater than/less than threshold value), but was also specifically designed to minimise the effects of direct interaction with external signals (the NAND gate returns a zero if both diodes show a simultaneous event). Unfortunately, this type of "safeguard" is found in all the REGs commonly used in parapsychology research. Future work of this kind will probably benefit from being based on a REG with an output showing a more direct mapping on to the random source.

However, limitations aside, the findings do offer some new perspectives on microPK research that may further our understanding of the phenomena. The idea that some effects seen in microPK studies may be due to an SR type interaction (with or without the involvement of the ostensible agent) suggests some reasons why previous microPK studies have shown inconsistent results. Hopefully it might also offer new ideas when looking at past research from this perspective. For example,

might the old Bio-PK indeed be an interaction effect, as suggested by the adoption of the new DMILS (Direct Mental Interaction between Living Systems) acronym? Perhaps the two people involved could be seen as both contributing to the observed results: one contributing a signal, the other noise.

There is also the idea that some instances of microPK might directly relate to the biomagnetic fields of the people involved. Although in this study the noise fields that were of "human brain" magnitude did not show obvious effects, this could simply be that the combinations of noise and signal were not the right ones (note that the strongest field also showed no effect). That any of the weak fields showed an effect suggests that biomagnetic fields cannot be ruled out of the microPK equation. Future work taking some direct measure of the participant's magnetic output (or at least some related measure such as global EEG activity) might be in a better position to evaluate any effects. There is also nothing to prevent the biomagnetic fields themselves (or some other as yet unrecognised physical field) being the signal if another source of suitable noise is present.

I am not trying to suggest that the ideas presented can account for all microPK effects. In addition to the cases discussed above, it is also hard to see how, according to current knowledge of the physical propagation of magnetic fields, an SR interaction could account for remote effects. Instead, I am presenting these results as a possible area of research that has been neglected but which I feel will help further our comprehension of psi effects. Based on our current understanding of the spatial and temporal parameters that appear to characterise psi effects, current physics seems to offer little chance of explaining the underlying mechanism. However, this apparent lack of explanatory power has led to a dearth of research looking at possible areas of physical interaction which might affect psi effects. It is often assumed that all microPK effects are the result of some unitary mechanism, whereas it may be that the findings are an amalgam of several different mechanisms. Without investigating the possibilities, even if they only apply to a subset of past results, we are unlikely to progress in our understanding. Such thinking has led to a disregard for the idea that electromagnetism in some form might be involved with psi effects, despite the widely acknowledged psi-geomagnetic relationship. I hope that results such as the above will lead us to reconsider.

## Acknowledgements

I would like to thank the two anonymous referees for their helpful comments. An early version of this paper was presented at the 2000 conference of the Society for Scientific Exploration, Amsterdam, Netherlands.

## References

- Bell, G.B., Marino, A.A., & Chesson, A.L., (1994). Frequency-specific responses in the human brain caused by electromagnetic-fields. *Journal of the Neurological Sciences*, 123, 26-32.
- Chauvin, R. & Varjean, B. (1990). Is it possible to strengthen the psi effect using a very weak magnetic field? *Journal of the Society for Psychical Research*, 56, 96-97.
- Dalton, K. & Stevens, P. (1996). Geomagnetism and the Edinburgh automated Ganzfeld, *European Journal of Parapsychology*, 12, 23-34
- Gammaitoni, L., Hänggi, P., Jung, P. & Marchesoni, F. (1998). Stochastic resonance. *Reviews of Modern Physics*, 70, 223-287.
- Gissurarson, L.R. (1992). The psychokinesis effect: geomagnetic influence, age and sex differences. *Journal of Scientific Exploration*, 6, 157-165.
- Grifoni, M., Hartmann, L., Berchtold, S. & Hänggi, P. (1996). Quantum tunneling and stochastic resonance. *Physical Review E*, 53, 5890-5898.
- Ho, M.-W., Popp, F.A. & Warnke, U. (Eds.). *Bioelectrodynamics and Biocommunication*, World Scientific.
- Ioannides, A.A. (1994). Estimates of Brain Activity Using Magnetic Field Tomography and Large Scale Communication within the Brain, in Ho, Popp, & Warnke, (1994).
- Jung, P. & Wiesenfeld, K. (1997). Too quiet to hear a whisper. *Nature*, 385, 291.
- May, E., Utts, J., & Spottiswoode, S.J.P. (1995). Decision Augmentation Theory: Towards a Model of Anomalous Mental Phenomena. *Journal of Parapsychology*, 59, 195-220.
- Nelson, R.D. & Dunne, B.J. (1986). Attempted Correlation of Engineering Anomalies with Global Geomagnetic Activity. *Research in Parapsychology*, The Scarecrow Press, pp. 82-85.
- Parkinson, W. (1983). *Introduction to geomagnetism*. Scottish Academic Press: Edinburgh.
- Persinger, M. & Krippner, S. (1989). Dream ESP experiences and geomagnetic activity, *Journal of the American Society for Psychical Research*, 83, 101-116.
- Radin, D. & Nelson, R.D. (1989). Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics*, 19, 1499-1514.
- Richardson, K.A., Imhoff, T.T., Grigg, P. & Collins, J.J. (1998). Using electrical noise to enhance the ability of humans to detect subthreshold mechanical cutaneous stimuli, *Chaos*, 8, 599-603.
- Stanford, R.G., Zenhausern, T.A. & Dwyer, M.A. (1975). Psychokinesis as psi-mediated instrumental response. *Journal of the American Society for Psychical Research* 69, 127-133
- Stevens, P. (1999). Remote psychokinesis. *European Journal of Parapsychology*, 14, 68-79.
- Stevens, P. (2000). Noise, physics and psi: new ideas for research. *International Journal*

*of Parapsychology*, 11, 63-72.

Stevens, P. (2001). Effects of 5s exposures to a 50 mT, 20 Hz magnetic field on skin conductance and ratings of affect and arousal. *Bioelectromagnetics*, 22, 219-22.

Warnke, U. (1994). Electromagnetic Sensitivity of Animals and Humans: Biological and Clinical Implications, in Ho, Popp and Warnke, (1994).

Wiesenfeld, K. & Moss, F. (1995). Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUID. *Nature*, 373, 33-36.



# Interpersonal Psi — Exploring the Role of the Experimenter and the Experimental Climate in a Ganzfeld Telepathy Task

Simon J. Sherwood\*, Chris A. Roe\*,  
Nicola J. Holt\* & Stuart Wilson\*\*

\*Centre for the Study of Anomalous Psychological Processes,  
The University of Northampton

\*\*Department of Psychology, Queen Margaret University College, Edinburgh

## Abstract

*This study examined the relationship between participants' pre-session perceptions of mood and expectations of success, rapport and the quality of the interactions with the experimenter, and the subsequent outcome of the ganzfeld telepathy sessions. A total of 38 ganzfeld telepathy trials were conducted by three different experimenters using an automated free-response testing system. Participants consisted mainly of volunteers from the local community and the experimenters' friends and acquaintances. Overall psi task performance, based upon ratings of the target relative to the dummy video clips, was slightly and non-significantly below chance (mean  $z = -.015$ ,  $SD = .81$ , one-sample  $t_{(37)} = -.11$ ,  $p = .90$ , two-tailed  $r = -.002$ ); there were no significant differences between the experimenters. The findings concerning potential links between subsequent psi task performance and participants' and experimenters' ratings of their pre-session personal feelings and expectations and their perceptions of the rapport and quality of interaction between all parties are somewhat unclear and, in some cases, contradictory. More qualitative analyses of, for example, filmed interactions between participants and psi-inhibitory and psi-conducive experimenters might identify more reliable predictors of psi performance.*

## Introduction

There seems to be a common belief that some experimenters (such as Charles Honorton, William Braud, Marilyn Schlitz and Kathy Dalton) are 'psi-conducive,' whereas others (such as James Crumbaugh, Donald West and John Beloff) appear to be 'psi-inhibitory' (cf. Irwin, 1999, pp. 84-86). Such consistent differences in findings could be due, at least in part, to properties inherent in the experimenter him/herself (e.g., see Kennedy & Taddonio, 1976; Palmer, 1997) and/or partly due to the experimental environment and climate in which the research is conducted. White (1977, p. 273) noted that "The experimenter has been a neglected variable in parapsychological research ... [yet]... there could hardly be a more significant area of investigation than the role of the experimenter." Gardner Murphy (1949) suggested that there is no such thing as a gifted participant as such, but rather how well a participant scores on a psi task depends upon the person who does the testing and the nature of the experimental conditions. However, recent Ganzfeld research suggests that "the experimenter effect ... may at least in part be a matter of choosing the right subjects" (Parker, Frederiksen & Johansson, 1997, p. 15), with those experimenters who are ostensibly more successful simply adopting a recruitment strategy that goes beyond the readily available — but perhaps not notably interested, able or experienced — pool of psychology students.

Smith (2003a) has summarised four possible explanations for the experimenter effect: (1) experimenter error — where successful experimenters make motivated but unconscious errors during their experiments; (2) experimenter fraud; (3) experimenter-participants interaction; (4) experimenter psi. Whilst we acknowledge the possibility of experimenter error, experimenter fraud and experimenter psi, we plan to focus our attention on the experimenter-participants interaction. A warm social ambience is considered to be very important in a ganzfeld context (Bem & Honorton, 1994; Dalton, 1997; see also Milton & Wiseman, 1999). The experimental environment and the interactions between the experimenter and the participants are considered to be important factors (e.g., Delanoy, 1997; Schmeidler, 1988; White, 1976, 1977) but have not received much experimental attention in terms of establishing how important these actually are, what specific aspects are the

greatest contributors and how these might be harnessed.

Successful experimenters in the field of parapsychology might be more highly motivated to obtain positive results and more likely to believe in the possibility of psi and therefore have greater expectations of success (see Delanoy, 1997; Rhine, 1948; Schlitz, 1987; Smith, 2003a; Taddonio, 1976; White, 1977; Wiseman & Schlitz, 1997, 1999). In a recent test of this hypothesis, Smith (2003b) found medium to large positive correlations between experimenters' belief that psi is possible, that it can be demonstrated in a laboratory setting, belief in their own psi ability and mean ratings of their perceived psi-conduciveness, as rated by their peers. According to Smith, (2003b, p. 125), "this suggests that it is these aspects of attitudes towards psi that are most important in predicting success as a psi experimenter (at least when success is measured by fellow researchers)." Of course, it is possible that these self ratings are a *consequence* of research success rather than a *cause* of it, or that some hidden factor affects them in such a way as to produce an apparent association, so we must be wary of reading causation into such correlations. In a more direct assessment, Sharp and Clark (1937) used several experimenters who differed *a priori* in their attitude towards psi. Experimenters with positive attitudes to psi elicited above chance performance (Sharp's mean hit rate = 5.36; Davidson's = 5.88; where *MCE* = 5.00) whereas an uncommitted experimenter elicited approximately chance performance (Berger's = 4.86) and a sceptical experimenter obtained below chance performance (Myers = 4.30). However, typically for a quasi-experiment, it is difficult to know whether these experimenters may have differed from one another in other ways that might have influenced participant scoring. Two recent EDA DMILS studies found that a psi-proponent experimenter obtained significant findings whereas a psi-sceptical experimenter following the same procedure did not (Wiseman & Schlitz, 1997, 1999). The authors of these studies noted that two possible explanations for these differential results were that they had used their own psi abilities or had acted differently towards their participants prior to the start of the experiment.

It is well-known in mainstream psychology that participants look to the experimenter for cues to what is expected of them — such cues are referred to as 'demand characteristics'. These cues might be picked up from the environment, what the experimenter says about the experiment and how this is expressed, and from the nonverbal body language and behaviour of the experimenter. The fact that interper-

sonal expectancy effects, such as the expectations of experimenters or teachers, seem to affect the subsequent performance of participants and students, respectively, has been well-established (see Rosenthal, 1966; Rosenthal & Rubin, 1978). In a meta-analysis of 135 studies, approximately 75% of which focused on research with teachers, Harris and Rosenthal (1985) identified 16 behaviours that are important potential mediators of this relationship, with effect sizes ranging from .12 to .32. Such mediating variables included a warm socioemotional climate, having longer and more frequent interactions, providing encouragement and praise, smiling and obtaining eye contact. The climate is one of the main factors in Rosenthal's four-factor theory that was strongly supported by the meta-analytic findings. It is difficult to manipulate deliberately the manner in which experimenters interact with their participants but nevertheless there is some suggestive evidence that experimenters who attempt to be informal, friendly, and supportive — or at least are perceived more favourably by their participants — might get more favourable results (Honorton, Ramsey & Cabbibo, 1975; Pratt & Price, 1938; Schneider, Binder & Walach, 2000; Watt & Baker, 2002). The pre-experimental briefing session is likely to be a key point during which the experimenter's attitudes and expectations are communicated and a rapport is or is not established with the participants (Smith, 2003a). Experimenter orientation and preparation also seems to be an important factor (Delanoy, 1997; Schlitz, 1987; Watt, Wiseman & Schlitz, 2002) albeit one that has received little empirical investigation. Some experimenters might spend more time briefing and getting to know their participants and make more of an effort to establish rapport than others (e.g., see Smith, 2003a; Watt et al, 2002). Potentially psi-conducive practice might include making participants feel comfortable and valued with positive expectations and a motivation to succeed in the experimental session (Delanoy, 1997). Woodruff and Dale (1950) were the first to include measures of the experimenter-participant attitudes towards one another. Surprisingly, higher ESP task performance was associated with participants who gave lower ratings for experimenters. It should be noted however, that 'low' ratings were actually neutral rather than very negative, and may indicate that disinterest or antipathy is desirable. For Dale, there was a difference in participant scores such that those to whom she gave higher liking ratings subsequently performed significantly better at the psi task. Parker, Millar and Beloff (1977) attempted to assess the relationship between the rapport established be-

tween the experimenter and participants (during the initial interactions involving the explanation of the study and experimental instructions) and subsequent ganzfeld performance. Rapport was assessed by blind independent judges who listened to tape-recordings of the interactions for the three experimenters. Although there were significant differences in the amount of rapport established by the different experimenters, this did not have a significant effect on performance.

Smith (2003a) summarised the results of studies (Edge & Farkash, 1982; Schmeidler & Maher, 1981) that have investigated the perceived personality and behavioural characteristics of 'psi-conductive' and 'psi-inhibitory' experimenters as follows:

"Supposedly 'psi-conductive' experimenters tended to be rated as more flexible, friendly, free, likeable, warm, enthusiastic and playful whilst 'psi-inhibitory' experimenters were rated as more rigid, cold, overconfident, tense, irritable, egoistic and unfriendly. An attempted replication found 'psi-conductive' experimenters to be perceived as more active, nervous and enthusiastic with 'psi-inhibitory' experimenters being perceived as more poised, egoistic, cold and confident (Edge & Farkash, 1992, pp. 171-2). Thus, although in both studies there was a good deal of variance between judge's ratings, there would appear to be some agreement among potential participants that 'successful' psi experimenters come across as more enthusiastic, warmer and less egoistic than do their less successful counterparts." (Smith, 2003a, p. 77).

However, an earlier study by Parker (1975, see also Parker, 1977), in which experimenters, later classified as either 'psi-conductive' or 'psi-inhibitory' by three judges, completed a 16 Personality Factor (16PF) measure, found no significant differences between these two groups in terms of extraversion, warmth and sociability, confidence adequacy (self-security) or tenderness and sensitivity. This study differed from the aforementioned studies in that the experimenters completed an assessment of their own personality characteristics. Perhaps the results depend upon whether the measures of the experimenter's personality are completed by the participants or the experimenter; perhaps the participants' perception is more important.

In summary then, it seems worthwhile to continue to look at the relationship between actual participants' pre-session perceptions of their

mood and expectations of success, their rapport and the quality of the interactions with the experimenter, and the subsequent outcome of the sessions. In terms of the participants' ratings, it is vital that participants feel that they can be completely honest about their perceptions of their experimenter and thus, we would argue, it is important that the participant is assured that the experimenter will remain blind to the ratings they have given. In our study, participants placed their completed unseen ratings in a sealed envelope that was posted to an independent researcher (SW) who would analyse these data. At no time would the experimenters see the interaction ratings from the individual participants. We agree with Smith (2003a) that a qualitative analysis of videotaped interactions between the experimenters and participants during this briefing session might shed further light on which aspects might be important and thus we planned to videotape this as part of this study, although these data are yet to be analysed and will not be reported here. Due to the difficulties associated with trying to manipulate experimenters' behaviour deliberately and the artificial nature of this, we decided simply to videotape our experimenters' natural behaviour.

## **Method**

### *Design*

This study adopted a correlational design in which experimenter, sender and receiver ratings of their current pre-session mood, relaxation-tension, pessimism-optimism, confidence of success in the experiment, and five measures of the rapport/interaction between the experimenter and the participants, were correlated with a measure of performance at a telepathy ESP task. The measure of performance was pre-specified as the  $z$  score of the target clip's similarity rating. A series of three pilot sessions and 40 trials were planned. However, a computer hardware failure after trial 38 resulted in the premature completion of data collection for this experiment. Two sessions were run but were excluded from the study: one because the sender failed to attend the session so it was run as a clairvoyance trial; one because the receiver needed to visit the bathroom during the mentation period and thus violated security protocols.

### *Apparatus*

This study used an automated ganzfeld computer system developed by Dr Paul Stevens and written in Microsoft Visual Basic v6 that presented video material via the API for Media Player v7. Video clips are stored digitally as MPEG files, labelled 1a, 1b, 1c etc. Three separate monitors for the experimenter, sender and receiver are controlled by the experimenter's PC. Security measures within the program lock the experimenter out of the system completely during a session so that it is not possible to switch to another application or access the computer except by aborting the session.

The target set consists of 116 minute-long digital video clips arranged in 29 sets of 4. These were the same clips used in our previous ganzfeld study (Roe, Sherwood & Holt, 2003) but some of the sets had been re-arranged to make them more orthogonal. These digital clips have mainly been produced at the University of Northampton and are drawn from popular television programmes and commercial films. Copies of the target pool are available on DVD from the first author upon request. Randomisation is achieved using the Visual Basic pseudo-random algorithm (rnd), having seeded it using the timer at the start of the program (RANDOMIZE TIMER). Once the "Start" button has been pressed, the computer first selects a target set, then selects one of the 4 clips within that set. The order of presentation of the four clips at judging is similarly randomised.

All trials were completed using specialist facilities in the Psychology Building at the University of Northampton. The receiver room is sound attenuated and is separated from a public corridor by two lockable doors. The sender's and receiver's rooms are separated by approximately 38m. A security camera is located outside the sender's room so that any activity there can be monitored by the experimenter and automatically video-recorded.

### *Materials*

The Participant Information Form (PIF) is a 56-item measure that was constructed for general use with parapsychological research at the University of Northampton and includes questions concerning biographical and contact details (11-items); religious and parapsychological background (5 items); computer experience (2 items); practice of mental/physical disciplines (2 items); belief in luck (2 items); clumsi-

ness and punctuality (2 items); competitiveness (1 item); absorption (2 items); sleep and dreams (4 items); imagination and fantasy-proneness (3 items); creativity (2 items); and physical and mental health (1 item). The remaining items relate specifically to knowledge, belief and experience of anomalous phenomena including telepathy, clairvoyance, precognition, psychokinesis, 'communication with the dead' and out of body experiences (19 items). The form concludes with questions about hypnagogic/hypnopompic experiences in a range of modalities (10 items) and an open question inviting descriptions of personal anomalous sleep-related experiences. Copies of all in-house measures are available from the first author on request.

Participants also completed the short extraversion and neuroticism subscales of the EPQ-R<sup>1</sup> (Eysenck, Eysenck & Barrett, 1985). Each subscale has 24 items with a dichotomous yes/no response format. The 18-item Australian Sheep-Goat Scale (ASGS, Thalbourne & Delin, 1993), with a 5-point Likert scale ranging from strongly agree to strongly disagree, was also completed.

Given the apparent importance of experimenter-participant interactions and expectations of success, at the end of the pre-session briefing, the experimenters and participants completed a short Interaction Questionnaire. This contained nine questions, concerning their personal feelings and expectations and perceptions of the quality of the interactions between experimenter and participants, which they were required to answer by giving ratings on 7-point scales.

1. How would you rate your current mood? (Negative-Positive)
2. How do you feel at this moment? (Relaxed-Tense)
3. How do you feel about the prospect of participating in this experiment? (Pessimistic-Optimistic)
4. How confident are you that today's experiment will be a success? (Not at all confident-Extremely confident)
5. How would you describe the quality of rapport that you have with the other Participant? (Extremely poor-Extremely good)
6. How would you describe the quality of rapport that you have with the Experimenter? (Extremely poor-Extremely good)

---

<sup>1</sup>An analysis of these personality variables will not form part of this paper



7. How would you rate the quality of the interaction between experimenter and participants? (Very cold-Very warm, 'Rehearsed'-Spontaneous, Very negative-Very positive)

Post-ganzfeld measures included a Sender Strategy Questionnaire that asked about the type of sending strategies used, whether this was active or passive, holistic or atomistic, focused on target clip or on the receiver, realistic or associative, and continuous or episodic. A Receiver Questionnaire asked about the receiver's experience.

### *Participants*

An opportunity sampling method was used to draw 38 pairs of participants (mean age of senders = 38.8 years [range = 19 – 71], 9 males and 29 females; mean age of receivers = 39.3 years [range = 19 – 72], 24 males and 15 females). Participants were not drawn from the pool of psychology undergraduates at the University of Northampton, but rather consisted mainly of volunteers from the local community and the experimenters' friends and acquaintances. Participants were not selected on the basis of prior belief or experiences, or personality and attitudinal dimensions that may predict psi performance (although such variables were measured), but presumably were sufficiently interested in or open to the notion of ESP to volunteer. In most instances, each participant provided his or her own sender but in some circumstances individuals were paired up by the experimenters. Lab personnel did not serve as participants. The mean ASGS belief score for receivers in this sample was 58.1 ( $SD = 14.4$ ) and for senders was 50.5 ( $SD = 15.8$ ). The receivers' belief mean is slightly above the score that would be achieved by checking the neutral midpoint for each item ( $M = 54$ ), suggesting that the sample are moderate believers; the senders' mean is slightly below this value, suggesting that they are moderately sceptical. Among this sample, eight senders and twelve receivers had previously participated in formal parapsychological studies; nine senders and 16 receivers had previously participated in casual testing; 18 senders and 22 receivers had practised a mental discipline such as meditation; and 14 senders and 20 receivers had practised a physical or spiritual regimen such as yoga or tai chi. All three authors acted as experimenters in the running of trials, with NH conducting 14 trials, CR 13 trials and SS 11 trials.

## *Procedure*

Prospective participants were sent an information sheet illustrated with photographs that described the nature of the study. This provided a rationale for the ganzfeld paradigm, outlined the stages of the experimental procedure, focusing on the roles of the experimenter, sender and receiver. Thus participants were made fully aware of all aspects of the experiment so that those who were not comfortable with the procedure had the opportunity to withdraw from the study. Prior to the trial, participants (senders and receivers) completed a battery of measures. A video player was set to record the input from the security video camera as the experimenter prepared for the session and continued recording until after the session was completed. Participants were greeted on arrival and escorted to a reception room that had been specially prepared with comfortable chairs, a coffee table, rugs and curtains in order to make participants feel as comfortable and relaxed as possible prior to the trial. This room contained a wall-mounted microphone and two opposite wall-mounted but unobtrusive cameras, which focused upon the seated experimenter and participants, respectively. Signals from the cameras and microphone were recorded on a video recorder in the experimenter's room. Participants were reminded about the cameras and potential recording of the briefing session and asked to consent to this before entering the room (all participants gave their consent). As recommended by Honorton et al (1990), experimenters encouraged an informal, friendly and positive atmosphere, discussing the procedure and answering any questions arising while sharing refreshments. The briefing included the background and rationale for ganzfeld ESP research plus a description of the procedure. Participants were then given a guided tour of the facility, as recommended by Honorton et al (1990) and Dalton (1997), as the roles of sender and receiver were again explained. After the briefing, the participants decided on who would be sender/receiver and then both the participants and the experimenter completed the interaction questionnaire. Completed responses were placed in a sealed envelope such that the experimenter did not see the participants' responses and vice versa; this envelope was then sent to an independent researcher (SW) who would later conduct the analysis of these data. Participants were made aware that their experimenter would never get to see their individual ratings and thus they were to give completely honest answers.

With the assistance of the sender, the experimenter prepared the receiver for the ganzfeld and wished them success. The receiver was seated in a reclining chair and encouraged to relax. They were invited to remove their shoes and cover themselves with a blanket if desired. The receiver wore headphones that had a microphone attached through which they could communicate with the experimenter and be heard by the sender. Halved ping-pong balls were placed over their eyes and held secure with micropore tape. A red light, positioned immediately in front of them at a distance that was comfortable for them (typically one metre), was shone onto the receiver's face. The receiver was then locked in the room (unless they were uncomfortable with this) and the sender was guided back to their room.

Once the experimenter had returned to the experimenter's room and established contact with the receiver the trial commenced. The receiver began by listening to and following a series of progressive relaxation instructions. The sender watched a randomly-selected video clip that was played fifteen times with one minute intervals between plays (to allow them a break and/or time in which they could focus upon sending details of the clip without being distracted by it). Drawing materials were provided for the sender should they wish to sketch elements of the target clip during these 'quiet' periods. During this thirty-minute sending/mentation period the receiver listened to white noise being played through their headphones and reported on any impressions or sensations that they experienced. The experimenter listened to the receiver's mentation via headphones from the experimenter's room and took notes. The sender could also hear any comments made by the receiver during the mentation period.

Following the mentation period, the experimenter read the receiver's mentation back to them and asked if there was anything further that they would like to add or elaborate upon. The receiver was then asked a series of questions regarding their experiences in the ganzfeld. Simultaneously, the sender completed a questionnaire concerning their interaction with the target and the sending strategies employed. At the judging stage the receiver was asked to remove their eye-shields but was encouraged to remain in a relaxed state as they watched four video clips, giving each one a rating (0–99) in terms of its similarity to their ganzfeld mentation. After viewing all four clips, they were able to view any or all of them as many times as they wished and to alter their ratings if necessary. The sender was able to listen to the clip soundtracks and

the interaction between the receiver and experimenter during the judging stage, but did not view the dummy clips. Once the receiver was satisfied with their ratings, these were confirmed and saved as a permanent record. Only after the data were saved was the target clip revealed and replayed. The experimenter then collected the sender and accompanied him/her to the receiver’s room where they convened with the receiver for a discussion and debriefing session. A copy of the trial data record was printed off and signed by all parties to confirm the details of the session.

### Results

Our primary psi measure was pre-specified as the z score of the target clip’s similarity rating. To facilitate future meta-analyses, however, Table 1 also gives the number of direct hits and ranks data overall and for each experimenter.

Table 1: Target rank frequencies

Experimenter	Rank				Rating	
	1	2	3	4	<i>M<sub>z</sub></i>	<i>SD<sub>z</sub></i>
NH	4 (28.6%)	5 (35.7%)	3 (21.4%)	2 (14.3%)	.048	.74
CR	3 (23.1%)	5 (38.4%)	1 (7.7%)	4 (30.8%)	.002	.96
SS	1 (9.1%)	4 (36.4%)	2 (18.1%)	4 (36.4%)	-.115	.77
Overall	8 (21.1%)	14 (36.8%)	6 (15.8%)	10 (26.3%)	-.015	.81

To assess our prediction that participants would award a similarity rating to the target that was higher than the average rating for the three dummy clips, z scores were calculated. The distribution of z scores is given in Figure 1. The hypothesis was not supported. The overall mean z score was actually slightly negative ( $M_z = -.015$ ,  $SD = .81$ ) but did not differ significantly from zero (one-sample  $t_{(37)} = -.11$ ,  $p = .90$ , two-tailed<sup>2</sup>  $r = -.002$ ). A Kruskal-Wallis test revealed no significant difference in the z-score for the target rating between trials conducted by the three experimenters ( $\chi^2 = .20$ ,  $df = 2$ ,  $p = .906$ ).

#### *Pre-session perceptions of senders, receivers and experimenters*

Note that due to an oversight, one experimenter (CR) did not complete any experimenter ratings prior to the session, only sender and re-

<sup>2</sup>This prediction was originally one-tailed, but in the opposite direction to that found. The two-tailed probability is given for information only.

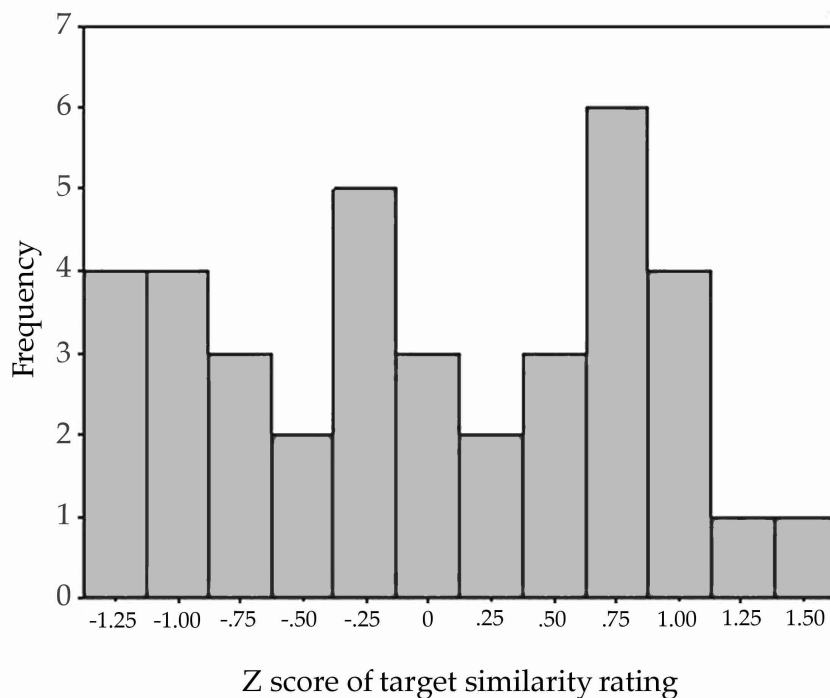


Figure 1. Frequency histogram of target similarity rating z scores

Table 2: Mean receivers' ratings presented separately per experimenter

Receiver (All ratings on scale 1-7)	NH Max N = 13	CR Max N = 12	SS Max N = 11	Overall Max N = 36
Current mood (Negative-Positive)	5.46	5.50	5.82	5.58
Relaxed-Tense	2.92	3.00	2.27	2.75
Pessimistic-Optimistic	5.08	5.33	5.60	5.31
Confidence of success (Not at all-Extremely)	4.58	4.33	4.82	4.57
Quality of rapport with other participant (Ext. poor-Ext. good)	5.69	6.08	6.00	5.92
Quality of rapport with experimenter (Ext. poor-Ext. good)	6.23	5.25	5.55	5.69
Quality of interaction between experimenter and participants:				
Very Cold-Very Warm	6.23	6.17	6.27	6.22
Rehearsed-Spontaneous	5.85	5.58	4.64	5.39
Very Negative-Very Positive	6.08	6.08	6.18	6.11

Table 3: Mean senders' ratings presented separately per experimenter

Receiver (All ratings on scale 1-7)	NH Max N = 13	CR Max N = 12	SS Max N = 11	Overall Max N = 36
Current mood (Negative-Positive)	5.46	5.25	6.18	5.61
Relaxed-Tense	2.54	2.83	2.27	2.56
Pessimistic-Optimistic	5.31	5.00	4.91	5.08
Confidence of success (Not at all-Extremely)	4.23	4.42	4.09	4.25
Quality of rapport with other participant (Ext. poor-Ext. good)	5.92	6.25	6.36	6.17
Quality of rapport with experimenter (Ext. poor-Ext. good)	5.69	5.25	5.82	5.58
Quality of interaction between experimenter and participants:				
Very Cold-Very Warm	6.31	6.17	6.27	6.25
Rehearsed-Spontaneous	5.38	5.00	4.70	5.06
Very Negative-Very Positive	6.08	6.17	6.27	6.17

Table 4: Mean experimenters' ratings presented separately per experimenter

Receiver (All ratings on scale 1-7)	NH Max N = 13	CR	SS Max N = 11	Overall Max N = 24
Current mood (Negative-Positive)	4.08	—	5.82	4.88
Relaxed-Tense	3.54	—	2.82	3.21
Pessimistic-Optimistic	4.54	—	5.29	4.80
Confidence of success (Not at all-Extremely)	4.77	—	4.82	4.79
Quality of rapport with other participant (Ext. poor-Ext. good)	—	—	—	—
Quality of rapport with experimenter (Ext. poor-Ext. good)	—	—	—	—
Quality of interaction between experimenter and participants:				
Very Cold-Very Warm	4.85	—	5.45	5.13
Rehearsed-Spontaneous	4.85	—	4.36	4.63
Very Negative-Very Positive	4.77	—	5.64	5.17

ceiver ratings. Ratings from one pair of participants could not be analysed because they had failed to indicate which of them was the sender and which the receiver. One set of ratings was never received by the independent data analyser.

#### *Personal feelings and expectations*

Across all three experimenters, both senders and receivers tended to be in a positive mood, fairly relaxed, optimistic about the prospect of participating in the experiment but not particularly confident that the session would be a success (see Tables 2–4). In fact, confidence of success was the aspect that was most lacking in the participants. The two experimenters, for whom data are available, were slightly more confident of success than their participants.

#### *Perceived rapport and quality of interaction*

Across all three experimenters, both senders and receivers tended to perceive the interaction between the experimenter and the participants as being warm, fairly spontaneous and positive (see Tables 2–4). The two experimenters, for whom data are available, had similar perceptions albeit not quite as positive.

#### *The relationship between the pre-session perceptions of senders, receivers and experimenters and subsequent ESP task performance*

Spearman's rho correlation coefficients were calculated between each of the receiver, sender, and experimenter ratings and the trial outcome, in terms of the z-score of the rating given to the target relative to the others in the set, for each of the three experimenters. The findings were mixed across the three experimenters, which suggest that the data should perhaps not be combined in order to calculate overall correlation coefficients, but these are reported for information. There was no correction for multiple analyses.

#### *Personal feelings and expectations*

There were no consistent findings across all three experimenters for the receivers' ratings (see Table 5). For CR and SS, more negative receiver moods were associated with higher target ratings ( $r = -.568$ ,  $-.441$ , respectively). For NH and SS, higher ratings of receiver tension were associated with higher target ratings ( $r = .281$ ,  $.269$ , respectively) but the reverse was true for CR ( $r = -.198$ ).

Table 5: Receivers' ratings per experimenter — Spearman's correlations with z-score target rating (two-tailed)

Receiver (All ratings on scale 1-7)	NH Max N = 13	CR Max N = 12	SS Max N = 11	Overall Max N = 36
Current mood (Neg-Positive)	.097 (.752)	-.568 (.054)	-.441 (.174)	-.335* (.046)
Relaxed-Tense	.281 (.352)	-.198 (.537)	.269 (.424)	.150 (.382)
Pessimistic-Optimistic	-.024 (.938)	.153 (.635)	-.281 (.431)	-.038 (.830)
Confidence of success (Not at all-Extremely)	.493 (.104)	-.118 (.715)	.183 (.589)	.184 (.289)
Quality of rapport with other participant	-.082 (.791)	-.547 (.066)	-.172 (.614)	-.289 (.088)
Quality of rapport with experimenter	.212 (.487)	-.467 (.126)	-.526 (.096)	-.285 (.092)
Quality of interaction between experimenter and participants:				
Very Cold-Very Warm	.060 (.846)	-.366 (.241)	.074 (.830)	-.125 (.469)
Rehearsed-Spontaneous	.346 (.248)	-.144 (.654)	-.262 (.437)	-.063 (.713)
Very Negative-Very Positive	.425 (.147)	-.030 (.927)	-.083 (.807)	.076 (.660)

Table 6: Senders' ratings per experimenter — Spearman's correlations with z-score target rating (two-tailed)

Receiver (All ratings on scale 1-7)	NH Max N = 13	CR Max N = 12	SS Max N = 11	Overall Max N = 36
Current mood (Neg-Positive)	.068 (.825)	-.190 (.555)	.353 (.287)	.033 (.847)
Relaxed-Tense	-.246 (.419)	-.416 (.179)	-.171 (.616)	-.268 (.114)
Pessimistic-Optimistic	.358 (.230)	.370 (.236)	.664* (.026)	.432** (.008)
Confidence of success (Not at all-Extremely)	.168 (.584)	.571 (.052)	.511 (.108)	.398* (.016)
Quality of rapport with other participant	-.052 (.865)	.247 (.439)	.178 (.600)	.147 (.391)
Quality of rapport with experimenter	.028 (.927)	.480 (.114)	.100 (.771)	.178 (.298)
Quality of interaction between experimenter and participants:				
Very Cold-Very Warm	.167 (.586)	.190 (.554)	.427 (.191)	.227 (.183)
Rehearsed-Spontaneous	.377 (.204)	.161 (.617)	.459 (.182)	.309 (.071)
Very Negative-Very Positive	.331 (.270)	-.168 (.602)	.044 (.897)	.077 (.654)



Table 7: Experimenters' ratings per experimenter — Spearman's correlations with z-score target rating (two-tailed)

Receiver (All ratings on scale 1-7)	NH Max N = 13	CR	SS Max N = 11	Overall Max N = 24
Current mood (Neg-Positive)	-.015 (.962)	—	.338 (.309)	.083 (.701)
Relaxed-Tense	-.517 (.070)	—	-.325 (.329)	-.332 (.113)
Pessimistic-Optimistic	.223 (.464)	—	.075 (.873)	.105 (.659)
Confidence of success (Not at all-Extremely)	.086 (.780)	—	.093 (.785)	.080 (.709)
Quality of rapport with other participant	.595* (.032)	—	-.017 (.959)	.256 (.226)
Quality of rapport with experimenter	—	—	—	—
Quality of interaction between experimenter and participants:				
Very Cold-Very Warm	.406 (.169)	—	.207 (.542)	.133 (.535)
Rehearsed-Spontaneous	.038 (.903)	—	.149 (.662)	.091 (.674)
Very Negative-Very Positive	.096 (.756)	—	.268 (.425)	.096 (.655)

Findings were mixed in terms of receivers' optimism ratings and confidence of success, but it is interesting to note a large positive correlation ( $r = .493$ ) between confidence ratings and target ratings z-scores for NH.

The senders' ratings (see Table 6) were the most consistent of the three roles involved with the experiment, at least in terms of the findings associated with their personal feelings and expectations. For all three experimenters, more relaxed, more optimistic and more confident senders tended to be associated with higher target ratings, particularly for CR and SS who were associated with some large effect sizes in this regard.

The only consistent trend for the two experimenters for whom data were available (see Table 7) suggested, with medium and large effect sizes ( $r = -.325, -.517$ ), that greater relaxation reported by experimenters was associated with higher target ratings. More optimistic and more confident experimenter ratings were also associated with higher target ratings, though the effect sizes were generally very small ( $r = .075$  to  $.223$ ).

*Perceived rapport and quality of interaction*

Again there were no consistent findings across all three experimenters for the receivers' ratings (see Table 5). There was some indication that for NH, receivers' ratings of more positive ( $r = .346$ ) and more spontaneous ( $r = .425$ ) interactions between the experimenter and the participants were associated with higher target ratings, but for CR and SS poorer rapport with the experimenter or other participant or colder or more rehearsed interactions were associated with higher target ratings.

Although there were some inconsistencies across the experimenters, the general overall trend for senders' ratings (see Table 6) suggested that greater perceived rapport and warmer, more spontaneous and more positive interactions between the experimenter and the participants were associated with higher target ratings relative to the dummy clips, with small to medium effect sizes in most cases.

For both experimenters for whom data were available, the trend suggested that perceptions of warmer, more spontaneous and more positive interactions with the participants were associated with higher target ratings (see Table 7), although the effect sizes were generally small (range .038 to .406). It is interesting to note that for NH there was a large positive correlation between quality of rapport with the other participants and target ratings ( $r = .595$ ) but almost no similar relationship for SS ( $r = -.017$ ).

## Discussion

Overall, as with our previous ganzfeld study, there was no overall evidence for ESP and receivers did not give higher ratings to the target relative to the ratings given to the dummy video-clips; in fact such ratings were slightly lower. There were negligible differences between the experimenters' results with NH's and CR's results approximating mean chance expectation and SS's being below chance expectations.

The findings concerning potential links between participants' and experimenters' ratings of their pre-session personal feelings and expectations and their perceptions of the rapport and quality of interaction between all parties are somewhat unclear and in some cases, even contradictory. Perhaps this is due to particular idiosyncrasies on the part of the individual experimenters in terms of how they tend to behave with participants or interactions between the moods/personalities of

all those involved. Perhaps this is not surprising given that there is no evidence for psi in this experiment either overall or for any individual experimenter. Our feeling is that such a quantitative approach to assessing the link between experimenter and participant interactions has been rather unsuccessful and perhaps it is, by definition, unable to adequately tackle the complexity and richness of human verbal and non-verbal interactions. One potential problem, pointed out by a referee of this paper, is that the wording of our interaction item, "How would you rate the quality of the interaction between experimenter and participants?" seems to suggest an interest in the two-way interaction between experimenter and participants and yet one of the response scales provided, "Rehearsed-Spontaneous," suggests more of a one-way focus on the experimenter's interaction with the participants. We agree with the referee that this item could usefully be reworded as "How would you rate the quality of the experimenter's interaction with the participants?" We recognize that the low power ( $N = 11-13$ ) for the analyses for individual experimenters is also a problem. The complexity of potential relationships and interactions among the variables of interest could perhaps be investigated by more multivariate analyses – this is something that we plan to look into. One of the problems with this kind of research is that "what seems cold and off-putting to one person may seem agreeably warm to another" (Schmeidler, 1997, p. 85). We are hopeful that a qualitative analysis of the videotapes of the pre-session interactions between our participants and ourselves as experimenters will be more fruitful and more useful in terms of identifying levels of rapport and identifying good and bad practice in terms of putting participants at ease and encouraging positive expectations of success in the subsequent experimental session. We support the endeavours of other researchers who have taken or recommended such a more qualitative approach (e.g., Schlitz, 1987; Smith, 2003a; Watt et al, 2002; White, 1997), particularly where they involve comparisons between experimenters who are unequivocally recognised within the parapsychological community as being psi-inhibitory or psi-conducive (cf. Wiseman & Schlitz, 1997, 1999).

If any of the putative associations investigated here are found to predict psi performance reliably, they seem most likely to be found in the senders' ratings, which indicate that trials were slightly more successful the more relaxed, optimistic, and confident of success senders were and the greater the perceived rapport and quality of interaction

they had with the experimenter and the receiver. Particularly in the ganzfeld context, the receiver might appear to have the more important role given that more preparation is required for them and therefore more time invested in them. Perhaps it is important to ensure that equal (more?) attention is paid to the senders during the pre-session briefings, notwithstanding that at least in some circumstances it seems possible for ESP to occur without the presence of a sender (e.g., Lantz, Luke & May, 1994). Perhaps the receiver is more sensitive to the experimenter and this is why their results are more variable. There is also a suggestion, albeit a non-significant one ( $r = -.332$ ,  $p = .113$ , two-tailed), that the more relaxed the experimenters feel before the session the higher the ratings given to the target clip at the end of the sessions. One aspect that both CR and SS found difficult, at times, was focusing fully on sessions that were scheduled during busy teaching and administration days when they may have otherwise been preoccupied; some sessions were scheduled in the evenings and on weekends, but the scheduling was mostly determined by the availability of the participants. This seems to provide further support for the claim that more attention needs to be paid to pre-session preparation by the experimenters (e.g., Delanoy, 1997; Schlitz, 1987; Watt et al, 2002).

Based upon the participants' and experimenters' ratings, it is clear that all those involved tended to be fairly relaxed and in a positive mood at the end of the pre-session briefing and perceived the quality of the interactions to have been warm, fairly spontaneous and positive. It is possible that NH might have had a better rapport with some participants because of her contact with them prior to the session; NH was responsible for arranging the majority of the sessions. Neither the participants nor the experimenters rated their confidence that the session would be a success as being extremely high, though the experimenters' ratings tended to be higher than the participants' but still only with a mean of between 4 and 5 on a 7-point scale. Although there was a warm social ambience here, it would seem that high expectations of success at this particular task were lacking as far as the participants were concerned (and to some extent from the experimenters — see Dalton's, 1997, recommendations). Perhaps extremely highly-motivated experimenters with extremely high expectations of success who do not question the existence of psi at all, and who have readily integrated it into their daily lives, are required (e.g., see Schmeidler, 1987; White, 1976). It may well be a case of pre-selecting experimenters as well as participants

with 'the right stuff.'

### Acknowledgements

We would like to thank the Perrott-Warrick Fund for their support of this project and the anonymous referees of this paper for their helpful comments.

### References

- Bem, D. J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin*, *115*, 4-18.
- Dalton, K. (1997). Is there a formula to success in the ganzfeld? Observations on predictors of psi-ganzfeld performance. *European Journal of Parapsychology*, *13*, 71-82.
- Delanoy, D. L. (1997). Important psi-conducive practices and issues: Impressions from six parapsychological laboratories. *European Journal of Parapsychology*, *13*, 63-70.
- Edge, H., & Farkash, M. (1982). Further support for the psi-distributed hypothesis [Abstract]. In W. G. Roll, R. L. Morris, & R. A. White (Eds.), *Research in parapsychology 1981* (pp. 171-172). Metuchen, NJ: Scarecrow Press.
- Eysenck, S. B. G., Eysenck, H. J., & Barrett, P. (1985). A revised version of the Psychoticism scale. *Personality and Individual Differences*, *6*, 21-29.
- Harris, M. J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin*, *97*, 363-386.
- Honorton, C., Berger, R., Varvoglis, M., Quant, M., Derr, P., Schecter, E., & Ferrari, D. (1990). Psi communication in the ganzfeld: Experiments with an automated testing system and a comparison with a meta-analysis of earlier studies. *Journal of Parapsychology*, *54*, 99-139.
- Honorton, C., Ramsey, M., & Cabibbo, C. (1975). Experimenter effects in extrasensory perception. *Journal of the American Society for Psychical Research*, *69*, 135-149.
- Irwin, H. J. (1999). Experimental research on extrasensory perception. In *An introduction to parapsychology* (3rd ed.) (pp. 59-111). Jefferson, NC: McFarland.
- Kennedy, J. E., & Taddonio, J. L. (1976). Experimenter effects in parapsychological research. *Journal of Parapsychology*, *40*, 1-33.
- Lantz, N. D., Luke, W. L. W., & May, E. C. (1994). Target and sender dependencies in anomalous cognition experiments. *Journal of Parapsychology*, *58*, 285-302.
- Milton, J., & Wiseman, R. (1999). Does psi exist? Lack of replication of an anomalous process of information transfer. *Psychological Bulletin*, *125*, 387-391.
- Murphy, G. I. (1949). Psychical research and human personality. *Proceedings of the Society for Psychical Research*, *48*, 1-15.
- Palmer, J. (1997). The challenge of experimenter psi. *European Journal of Parapsychology*, *13*, 110-125.
- Parker, A. (1975). A pilot study of the influence of experimenter expectancy on ESP scores. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in Parapsychology 1974*, (pp 42-44). Metuchen, NJ: Scarecrow Press.
- Parker, A. (1977). Parapsychologists' personality and psi in relation to the experimenter effect [Abstract]. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research*

- in parapsychology 1976* (pp. 107-109). Metuchen, NJ: Scarecrow Press.
- Parker, A., Frederiksen, A., & Johansson, H. (1997). Towards specifying the recipe for success with the Ganzfeld. *European Journal of Parapsychology*, *13*, 15-27.
- Parker, A., Millar, B., & Beloff, J. (1977). A three-experimenter ganzfeld: An attempt to use the ganzfeld technique to study the experimenter effect. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in parapsychology 1976* (pp. 52-54). Metuchen, NJ: Scarecrow Press.
- Pratt, J. G., & Price, M. M. (1938). The experimenter-subject relationship in tests for ESP. *Journal of Parapsychology*, *2*, 84-94.
- Rhine, J. B. (1948). Conditions favoring success in psi tests. *Journal of Parapsychology*, *12*, 58-75.
- Roe, C. A., & Sherwood, S. J., & Holt, N. J. (2003). Interpersonal psi: Exploring the role of the sender in Ganzfeld GESP tasks. *Proceedings of the Parapsychological Association 46th Annual Convention*, 237-256.
- Rosenthal, R. (1966). *Experimenter effects in behavioural research*. New York: Appleton-Century-Crofts.
- Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: the first 345 studies. *Behavioral and Brain Sciences*, *3*, 377-415.
- Schlitz, M. (1987). An ethnographic approach to the study of psi: Methodology and preliminary data [Abstract]. In D. H. Weiner, & R. D. Nelson (Eds.), *Research in parapsychology 1986* (pp. 103-106). Metuchen, NJ: Scarecrow Press.
- Schmeidler, G. R. (1988). *Parapsychology and psychology: Matches and mismatches*. Jefferson, NC: McFarland.
- Schmeidler, G. R. (1997). Psi-conducive experimenters and psi-permissive ones. *European Journal of Parapsychology*, *13*, 83-94.
- Schmeidler, G. R., & Maher, M. (1981). Judges responses to the nonverbal behaviour of psi-conducive and psi-inhibitory experimenters. *Journal of the American Society for Psychological Research*, *745*, 241-257.
- Schneider, R., Binder, M., & Walach, H. (2000). Examining the role of neutral versus personal experimenter-participant interactions: An EDA-DMILS experiment. *Journal of Parapsychology*, *64*, 181-194.
- Sharp, V., & Clark, C. C. (1937). Group tests for extrasensory perception. *Journal of Parapsychology*, *1*, 123-142.
- Smith, M. D. (2003a). The role of the experimenter in parapsychological research. *Journal of Consciousness Studies*, *10*, 69-84.
- Smith, M. D. (2003b). The psychology of the "psi-conducive" experimenter: Personality, attitudes towards psi, and personal psi experience. *Journal of Parapsychology*, *67*, 117-128.
- Taddonio, J. L. (1976). The relationship of experimenter expectancy to performance on ESP tasks. *Journal of Parapsychology*, *40*, 107-114.
- Thalbourne, M. A., & Delin, P. S. (1993). A new instrument for measuring the sheep-goat variable: Its psychometric properties and factor structure. *Journal of the Society for Psychological Research*, *59*, 172-186.
- Watt, C., & Baker, I. S. (2002). Remote facilitation of attention focusing with psi-supportive versus psi-unsupportive experimenter suggestions. *Journal of Parapsychology*, *66*, 151-168.
- Watt, C., Wiseman, R., & Schlitz, M. (2002). Tacit information in remote staring research: The Wiseman-Schlitz interviews. *The Paranormal Review*, *24*, 18-25.

- White, R. A. (1976). The limits of experimenter influence on psi test results: Can any be set? *Journal of the American Society for Psychical Research*, 70, 333-369.
- White, R. A. (1977). The influence of the experimenter motivation, attitudes and methods of handling subjects in psi test results (pp. 273-304). In B. B. Wolman (Ed.), *Handbook of parapsychology*. New York: Van Nostrand Reinhold.
- White, R. A. (1997). Suggestions for exploring and recording the inscape of psi researchers. *European Journal of Parapsychology*, 13, 96-109.
- Wiseman, R., & Schlitz, M. (1997). Experimenter effects and the remote detection of staring. *Journal of Parapsychology*, 61, 197-207.
- Wiseman, R., & Schlitz, M. (1999). Experimenter effects and the remote detection of staring: An attempted replication. *Proceedings of Presented Papers of the 42<sup>nd</sup> Parapsychological Association Annual Convention*, 471-279.
- Woodruff, J. L., & Dale, L. A. (1950). Subject and experimenter attitudes in relation to ESP scoring. *Journal of the American Society for Psychical Research*, 44, 87-112.

## Book Review

Marios Kittenis

Koestler Parapsychology Unit, Department of Psychology  
The University of Edinburgh

**A Review of “Parapsychology in the 21st Century:  
Essays on the Future of Psychical Research”, Edited  
by Michael A. Thalbourne and Lance Storm (2004)**

This collection of essays takes the turn of the century as an opportunity to present a varied sample of contemporary work from the field, and invites the authors to look ahead at the possible directions parapsychology might take in the future. While partly an attempt at an assessment of the state of parapsychology at the present, unlike other books which accomplish this in much greater depth and detail, it has a unique emphasis on looking ahead to envision where the field might go from here.

As Fiona Steinkamp comments in a later chapter, “there are at least two problems in trying to write about the future of any discipline. The first one is that whatever one writes will, in retrospect, probably appear misguided once the future has arrived. The second problem is that it seems rather presumptuous for any one person to prescribe to others the directions in which that discipline should move”. No matter how fraught with difficulties however, such attempts are nevertheless worthwhile, as they allow investigators to temporarily forgo the customary restrictions on creative speculation, and may produce just the right kind of conceptual cross-fertilisation that is necessary for innovative progress in any field.

The forward by Brian Josephson comments on the ever-present resistance to parapsychology from the scientific mainstream, not only in



its claims, but as a scientific endeavour in its own right. This is echoed in the first chapter by Dean Radin, who traces this attitude to the social context in which science is inextricably embedded, suggesting that “the past teaches us that scientific resistance to psi is far more than intellectual disagreement. Psi is an idea that challenges tightly held ideological assumptions”, among which he identifies as primary the challenge the psi hypothesis poses to our customary distinction between the objective and the subjective, and the threat that comes with this blurring of psychic boundaries to our personal and social egos, and with them to some of the very foundations of modern civilisation, such as personal identity and responsibility, and the social identities of patriotism and nationalism. He supports his observations by reflecting on some less well-known facets in the social history of parapsychology, and projects various possible scenarios for the future. Among these are the possibility of a widespread acceptance and utilisation of psi as an ability and a technology, along with the more dystopian possibilities of widespread misapplication and abuse of such technology (not that unlikely given our dubious track record in using other technologies we have so far learned to master). Noting that the idea of psi is not necessarily dangerous, but that its widespread application may well be, Radin contemplates that “perhaps society at large really does act as a sort of ‘reality governor’ preventing us from learning too much about psi, or ... too quickly”.

However uncomfortable these challenges may make us feel, as Josephson remarked in his forward, “science is not about being comfortable”, and Radin predicts that “by the time the world has comfortably accommodated psi, it will not be thought of in today’s terms at all”, and optimistically suggests that “new views of time and causality in physics, and the adoption of modern physics into the biological sciences, will begin to make some aspects of parapsychology much more palatable to the mainstream”. His vision of the future appears to be one of competing worlds, some of which are wonderful while others are disastrous, while ultimately however, he trusts that “evolutionary necessities will force us past short-term myopic fears”.

The second chapter by Robert Morris is a reprint from the *Journal of Parapsychology*, and presumably what inspired the title for this book. The essay revisits 10 potential difficulties facing parapsychology which the author first identified in 1990 and which he still considered to be relevant at the turn of the century, and offers six strategies for para-

psychology to adopt as it enters the 21st century. Among potential difficulties Morris includes some of the same social factors identified by Radin in the previous chapter e.g., that parapsychology threatens fixed beliefs about how the world works, and that it has often been labelled a pseudo-science by philosophers and sociologists of science. However, Morris also points out some more fundamental difficulties that are intrinsic to the subject-matter of parapsychology. Among these is the apparent difficulty in generating and testing theory-based hypotheses, and the fact that parapsychology typically studies complex, open systems, whose boundaries of influence cannot be precisely delineated. He comments that due to these characteristics, "we may need to focus more on strategies for evaluating the output of definable psi conducive systems, foregoing at least at the start the systematic exploration of specific causal linkages", incorporating "research procedures likely to produce effect sizes sufficiently strong that serious testing of models, of theory-driven hypotheses, can be done". As for strategies for the future, he suggests that we need to evaluate more completely what we have learned so far, partly by the careful, intelligent use of meta-analysis; that we should learn more from our negative results; and that we should focus on measures with a good track record in terms of effect sizes and consistency. As parapsychology is particularly affected by its social context, he advises that we need to be more effective in interacting with the media, and to work towards breaking down the divisions between "sceptic" and "researcher". Finally, he suggests that "as we attract more interest from experts in other areas, we need to integrate more effectively with them and their expertise", and also that we should be more proactive in offering aspects of our own expertise as it relates to the work of others. It is perhaps especially important now for investigators to contemplate these suggestions, as they represent a lifetime's experience in the field by a colleague whose breadth of knowledge and soundness of judgement we will all surely miss.

In the third chapter, William Braud first casts a glance back at the beginnings of psychical research, before making the case for a widening of the domain of contemporary psi research beyond the "big four" phenomena of telepathy, clairvoyance, precognition and psychokinesis. He regrets the narrowing of the range of phenomena that parapsychology has been prepared to deal with since its beginnings as psychical research, arguing that over the years it has become "concerned more and more with less and less". He suggests that "our field can profit from

an expansion of its subject matter and from a larger contextualisation of what is studied in psi research”, mentioning the relevance of exceptional human experiences, non-ordinary and transcendent experiences and arguing for their inclusion into our investigatory gambit, and proposes the growing field of transpersonal studies as a “suitable larger rubric in which to situate our psi studies”. Braud also urges a widening of our time-windows when reviewing past research, which would not only honour the work of earlier contributors, but which could also “uncover important clues to psi functioning that otherwise might be lost ... and prevent the redundant reinvention of many intellectual ‘wheels’”. He also argues for a deeper and more thoughtful assessment of our conclusions and assumptions, such as the often assumed temporal and spatial independence of psi, drawing attention to the uncomfortable possibility that much of our research findings which we assume to be revealing inherent features of psi functioning, may in fact be a reflection of experimenter beliefs, expectations and selective attention, and the way these define the experimental designs that we use. In the last section, reiterating his call for psi researchers to adopt a more inclusive approach in their work, Braud warns against exclusively mimicking the methods of the natural sciences, and encourages us to also be open to learning from philosophy, history and other humanities and the arts, as well as from various esoteric and wisdom traditions which often describe processes closely resembling those studied in psi research. He urges investigators to place more emphasis on exploring the life impacts of psi and other non-ordinary experiences, as well as the meanings and interpretations people attribute to them, and calls on researchers to become more intimately familiar with their subject-matter from a first person perspective. His suggestion that we explore the limits of psi and how it may be attenuated seems a particularly worthwhile goal, one that has largely been ignored so far in the near-compulsive search for positive results. He predicts that success in this may alleviate some of the fear of psi, which he sees as possibly related to decline effects and the apparent elusiveness of psi.

The next chapter by Adrian Parker, entitled “Psi and Altered States of Consciousness”, initially adopts a chronological perspective and demonstrates how intertwined the historical threads of these two topics have been, from early psi research and the first recorded instances of spontaneous use of remote viewing and ganzfeld-like techniques, to contemporary laboratory procedures which incorporate such methods

of altered states induction as a formal part of their protocol. Notable landmarks along this route include the various attempts at using hypnosis to facilitate psi-conducive states, the Maimonides dream-ESP program and replication attempts that followed, and the development of the ganzfeld technique and its most recent incarnation, the real-time digital ganzfeld technique (or RTDGT). The rest of the review is largely devoted to the RTDGT, which is somewhat unfortunate, in that the original theme of altered states and psi seems to be forgotten in the process, although the discussion of RTDGT is engaging in its own right. For example, the unique quality of temporal information this protocol can provide, allows for Parker's observations of some unusual correspondences between mentation reports and time-displaced or non-target material, leading to interesting questions regarding the apparent conflict between real-time causal accounts of psi, and the apparently time-independent acausal features that it occasionally demonstrates. This divergence in topic however, together with some notable omissions in the overview of altered states of consciousness and psi, such as the lack of mention of parapsychological studies making use of meditation and psychoactive substances, makes one wonder if this essay would not have been more accurately described under a different title. However, as much as readers expecting an essay dedicated on the title topic may be disappointed, they will be compensated by the concluding section which provides a collection of observations on the past, present and future of psi research that are worth persevering for. The closing statement by Parker also seems to summarise the views of several other authors in this book: "The greatest hindrance to future research is no longer the replication problem, but the need to produce theories that relate to mainstream psychology. To be able to study the phenomenon under at least semi-controlled conditions without denaturalising it, and then to relate this knowledge to neuroscience, is the challenge of parapsychology in the twenty-first century".

The following chapter by Suitbert Ertel describes the Ball Drawing Test, which, unlike my initial impression, has nothing to do with measures of artistic aptitude in the graphical representation of spherical objects. The BDT is instead a simple procedure for testing psi, which Ertel maintains has the distinct advantage of combining ecological validity with the ability to conduct large numbers of trials easily and inexpensively. It involves participants blind-drawing numbered balls from an opaque bag, after they guess and record a number in advance. The num-

ber on the ball that is subsequently drawn is also recorded, and the correspondence between these pairs of numbers is compared against mean chance expectation after a sufficiently large sequence of draws. Ertel describes a number of trials done by participants alone at home and under controlled conditions, the first of which show far more impressive hit rates compared to the latter, which well exceed traditional card guessing results. He naturally considers the possibility of participants tampering with their home records, but noting that the hit rates of the most successful subjects recovered gradually in controlled conditions after an initial drop, sometimes approaching their home test levels, he is more inclined to conclude that a relaxed, familiar home atmosphere is more conducive to successful performance, while the psychological pressures of laboratory controls can be detrimental. He proposes adopting this test primarily in alone-at-home conditions, with occasional lab controlled validation trials with the rationale that as long as hit rates recover with continued testing, one can rule out suspicions of subjects tampering with their home results. The BDT has other attractive features other than its simplicity and inexpensiveness in terms of the materials and man-hours investment needed, such as the ease with which it makes itself suitable for psychometric evaluation, and the fact that participants' activity during the test combines motor activity with tactile, visual and auditory sensory perceptions, thus engaging more than solely mental faculties and making it more akin to everyday actions. Most investigators however would find the level of control against tampering in alone-at-home trials unacceptable, and laboratory trials seem to show modest hit-rates on average. Although the author is certainly right in saying that a test capable of showing a consistent effect can be very valuable in understanding the phenomena involved, even if it requires conditions lacking the best level of controls, the BDT as proposed seems best suited for exploratory studies and demonstration/training purposes, as it cannot provide the level of safety against fraud or error expected of formal studies. Perhaps this problems can be easily rectified with some relatively minor modifications of the test however; for example, balls fitted with inexpensive Radio Frequency Identification (RFID) tags, as used in shop anti-theft systems<sup>1</sup>, could be used in conjunction with a bag or other container fitted with a tag reader used to identify the number of the ball drawn from it each time. A small electronic device can then en-

---

<sup>1</sup>For more information see:<http://en.wikipedia.org/wiki/RFID>

crypt and record this number in a tamper-proof memory card. Such a device could also be used to register the participant's choice of number before the ball is selected, and the experimenter could then retrieve the data once the bag is returned, confident that these accurately represent the outcome of the trials. Such modification does not require the presence of the type of laboratory controls that Ertel justifiably points out as potentially psi-inhibitory, and maintains all the original advantages of the BDT technique. Such a system could well fulfil Ertel's assertion that the BDT "may serve as a standardized psi test with ordinary psychometric properties (validity, reliability)", something which would be invaluable in clarifying the factors involved in psi-processes by using variations in research design.

Such a test capable of providing consistent results, which can be used to methodically explore the nature of psi phenomena, can be said, with very little exaggeration, to be the "holy grail" of contemporary psi research. As William Braud and others in this volume warn however, it is all too easy to become fascinated with looking for the next new shiny tool to add to our investigatory toolbox, and lose awareness of the wealth of knowledge our older tools have already uncovered. Fiona Steinkamp compensates for some of this lack of awareness of the past in the following chapter, by providing a thorough review of the forced-choice ESP literature from 1880 onwards, and in doing so she provides a valuable resource for examining factors potentially associated with ESP scoring. Identifying such variables is necessary for designing the much sought-after universally replicable experiment, "or, failing that, the experiment that can be replicated by experimenters of a clearly defined type". Some of her conclusions are surprising considering certain widely held assumptions in the psi research community, thus revealing "how little we know even about the most promising variables". For example, although the associations of forced-choice ESP scoring and low neuroticism, extraversion and good social adjustment appear to be quite consistent, Steinkamp calls for a re-examination of the sheep-goat literature, and argues that the basic issue of whether psi declines over distance and time needs to receive more attention in order to be understood better, an understanding which is essential for viable model-building purposes. She suggests that among the more easily implementable findings are "to use trial-by-trial feedback, to test people individually, to avoid using goats or super-sheep, and possibly to use shorter runs". In closing, she calls for a greater interaction with mainstream researchers,

and suggests that “if parapsychology is to walk among the sciences, it has to learn to walk with a direction and follow what evidence there is”.

Stanley Krippner and Gerd Hövelmann open the next section of the book which deals primarily with theoretical issues, putting forward their eleven recommendations for future psi research they first proposed in 1986, by presenting a critique of a critique of their original article, while adding another six new recommendations. These are all thoughtful and insightful suggestions and their advice is well worth contemplating; after all, much of what they suggest could be seen as basic common sense, and with that being a virtue all too often undervalued or overlooked in ambitious academic endeavours, their observations are particularly refreshing in this context. Although much of what they say is echoed in other essays, such as the ones by Bob Morris and William Braud, one can also find some truly novel observations. They call, for example, on researchers to utilise parapsychology’s uniqueness, as seen in its success in developing “methods and techniques of investigating anomalous events that, by and large, are more sophisticated than those to be found in any comparable field”. They remind us that “psi phenomena are not the only anomaly that needs to be taken seriously by scientific investigators in the 21st century”, and propose that lending our expertise can be of valuable help in this respect.

The next two chapters are written by the book’s editors, Michael Thalbourne and Lance Storm, and are both dealing with the theory of psychopraxia, with Thalbourne providing a clarification of some of the terms introduced in the theory, and Storm providing a critique of the theory. Thalbourne’s essay is not intended as an introduction to the theory of psychopraxia, but as a discussion of some issues not covered in depth in previous publications, and will be of most value to those already somewhat familiar with the theory. Several authors in this book have lamented the general lack of theoretically-driven predictions in the field, and attempts such as Thalbourne’s to formulate such predictions are certainly commendable. At the same time, there is justified resistance to the introduction of new terminology and theoretical constructs, even if most parapsychologists are often dissatisfied with the terminology already in use. Storm provides a psychological and philosophical critique of some of the terms introduced by the theory, thus testing its fitness for survival, while also contributing to its continued evolution by pointing out its weaknesses and necessary improvements. The long-term survival of the theory of psychopraxia however, is something that

will only be decided by the extent to which it proves capable of successfully guiding future research, and of contributing to our greater understanding of psi phenomena.

The next essay by Christine Hardy is entitled “Tackling the mind-matter problem from a consciousness perspective”, in which she views consciousness, through its creation of meaning, as an organising and negentropic force in the universe. She presents her *semantic fields theory* in which she suggests a semantic dimension where mind and matter are deeply intermingled, and she considers all cognitive processes, not only psi, to “present anomalies to classical physics and quantum mechanics because they reveal a more complex dimension that instantiates nonlocal properties escaping space-time constraints”. She also conceives a *law of connective-dynamic emergence*, where “totally distinct types of systems or processes . . . through internal self-organisation and a connective dynamic, are able to adjust between them, to interact dynamically and coevolve, and henceforth to create an emergent system with emergent properties”. As she sees these dynamics to also exhibit goal-directedness, Hardy goes on to consider the implications for intentional re-organisation in complex, living, mind-brain systems, and speculates on the possibilities for such systems to intentionally self-organise into new and more adaptive forms. One possibility she considers, as an adaptive evolutionary response to the multiple challenges facing our species at its current stage of technological and cognitive development, is for a collective emergence of latent psi abilities in the human population. She predicts that this may give rise to a novel type of collective consciousness, similar to the concept of the *noosphere* as envisioned by Teilhard de Chardin, and points to the Global Consciousness Project as a potential gauge of such an emergence.

Vernon Neppe and John Palmer devote their essay on the concept of Subjective Paranormal Experiences, making the case for the importance of studying the purely subjective elements of such events in the manner that psychology studies other unusual experiences, such as hallucinations and *déjà-vu*, by using detailed phenomenological analyses and classification of events. They suggest that a greater understanding of the subjective elements of paranormal experiences will be of significant value to other disciplines, especially to psychiatry, and as such will encourage a greater interaction and integration of parapsychology with other human sciences. As they point out, the basic terminology of the medical model in psychiatry most often labels paranormal experiences



in psychopathological terms, and they see a need to “educate psychiatrists and psychologists about basic approaches in parapsychology, and to re-establish diagnostic nomenclature and narrow what is regarded as abnormal”. They provide the example of hallucinations, which although usually considered abnormal and pathological, are in fact surprisingly common in the general population, to suggest that no matter how strange a patient’s experiences, they need not be regarded as abnormal “unless they distinctly interfere with the patient’s functional and coping skills”. They proceed to describe in some depth how such a phenomenological approach may be applied in the cases of out-of-body and near-death experiences, and call for parapsychology to develop a detailed classification system for such experiences, which will enable it to work with educating others in this area.

In the final section of the book dealing with sociological and phenomenological issues, Lance Storm presents a socioempirical perspective on scepticism about psi, examining in greater depth a theme touched by several authors in this book. Discussing first the problem of specification of psi phenomena, he comes to the conclusion that this is no more or less a burden for parapsychologists than it is for investigators in other disciplines when they honestly confront their subject-matter, unlike the general impression given by sceptics that parapsychology is a special case. Discussing Kuhn’s notion of “paradigmatic incommensurability”, according to which there may be no common measures by which ideas from one paradigm could be understood from within the context of another, Storm points out that parapsychologists will commit to multiple or plural paradigmatic viewpoints. He suggests that, “in the past, this approach would have suggested a discipline that was fragmented (if not unscientific), but now it may be regarded as a major strength for parapsychology”. Finally the chapter discusses the three factors of experience (of psi events), belief and the scientific method, which influence researchers’ attitudes and prejudices, and have given shape to the field of parapsychology and its paradigms. Storm sees these factors as being of direct concern to the field, as they ultimately influence its social status as a discipline. As to the relevance of these factors to scepticism, perhaps one of the best pieces of advice found in the essay and the book as a whole, is a paraphrase from John Beloff (1990), arguing that a common-sense approach means doing “justice to the evidence while, at the same time, seeking to do the least violence to our reason and general knowledge”.

In his chapter on language and the study of parapsychological phenomena, Robin Wolffitt points out that just like any other research area, parapsychological research relies to a large extent on language for investigating and communicating its subject-matter, and suggests that it would be beneficial to parapsychology to study the communication processes involved in language exchanges found both in the laboratory and in more naturalistic settings. He demonstrates this by giving examples of using conversation analysis (CA) to analyse language patterns found in accounts of spontaneous anomalous experiences, in psychic-sitter interactions, and in experimenter-subject interactions in the mentation review period in ganzfeld experiments. In all these cases, the patterned regularities identified by CA often yield insights about these interactions and the underlying processes that would have otherwise gone unnoticed. As the editors point out, the type of information offered by CA make it seem particularly well suited for studying experimenter effects, while Wolffitt envisions a wider role by proposing that the sociological study of language offered by CA “might develop as a method for parapsychological inquiry” in its own right, and invites the broader interaction between parapsychology and sociology, welcoming the blurring of their disciplinary and methodological boundaries.

James McClenon next presents his Ritual Healing theory, whose central premise is that the physiological bases giving rise to anomalous experiences developed within an evolutionary context. Pointing out the relationship between dissociation and anomalous experiences, he argues that “genes related to dissociation increased in frequency among ancient hominids since dissociation alleviates the negative impacts of trauma”, also remarking on the relationship of dissociation with the ability to perform multiple tasks simultaneously, another highly desirable evolutionary trait. He notes the use of dissociative techniques by shamans, mediums and other psychic practitioners, as well as the intercorrelation between the propensity for anomalous experiences, hypnotic susceptibility, dissociation, fantasy proneness, temporal lobe lability and thinness of cognitive boundaries, to put forward a model of ritual healing where these qualities are utilised by a shaman or psychic to manipulate the beliefs of their patients in order to induce healing by hypnotic suggestion. McClenon suggests that the theory can provide empirically testable hypotheses regarding anomalous experiences within the framework of evolutionary psychology, which can be relevant to fields such as psychical research, anthropology, social psychol-

ogy of religion, history and medicine.

The last chapter in the book, by Pamela Rae Heath, describes a phenomenological approach to psi research and urges investigators to explore the subjective components of psi experiences. She first gives a brief review of phenomenology as a technique, describing the several stages involved in the analysis, from formulating a question, such as “what is the experience of...”, to collecting a number of descriptions from people who have had the experience, to following a set of explicit steps to analyse the descriptions in order to identify the constituents of the experience, to finally synthesising and integrating these into a consistent description of the experience which captures its universal essence.

She then proceeds to give examples of how phenomenological techniques can be applied in several types of anomalous experiences, such as synchronicity, channelling, near-death experiences, ESP, PK and exceptional human functioning in sport. These examples are both engaging and informative, not only in their elucidation of phenomenology as a technique, but also in what they reveal about the constituent themes of these experiences and their commonalities and differences. For example, in her analysis she found that all of the constituent themes of the PK experience also appeared in the ESP experiences as well. In her conclusions, she states that because of the long-standing emphasis on proof-oriented research, qualitative research has been neglected, something which she sees as representing a gap in our knowledge, as “to truly understand a phenomenon, like psi, it is important to learn its fundamental nature... By giving us an overview of psi, phenomenological research can provide important clues about its functioning that cannot only enhance our general understanding of the field, but also allow us to design better future research”.

Overall, the greatest strength of this book is in its diversity, and while it is unlikely that anyone will find all chapters in the book to be of interest, most everyone is likely to find some that are. Among the weaknesses of the book is the occasional tendency for some essays to drift off topic and for some to be far longer than they need to, which negatively affects their readability.

An interesting general impression from the book as a whole is that a feeling of disciplinary self-consciousness seems to pervade the field, as it still struggles to become accepted as a legitimate scientific discipline. This seems to underlie the many calls for parapsychology to reach out

and establish links with other disciplines, so as not to become increasingly isolated and marginalised. Such efforts are certainly to be encouraged, as science cannot be done in isolation, especially in an increasingly multidisciplinary research climate. One observation however, is that this collection considerably overemphasises potential interactions of parapsychology with phenomenological and sociological fields of study, and largely ignores potential interactions with the physical and biological sciences. It would be regrettable if this were representative of a tendency of the field to move increasingly away from empirical laboratory research in favour of sociological and phenomenological studies, only because the latter are more palatable to the mainstream, as Neppe and Palmer seem to suggest when they advocate the study of subjective paranormal experiences partly because “the subjective approach can diminish the threatening elements of parapsychology substantially”. The study of the subjective and sociological characteristics of anomalous experiences is certainly a worthwhile goal in itself, but this does not need a separate discipline for its accomplishment, and could easily be part of the activity of already established disciplines such as psychology and sociology. What parapsychology seems uniquely able to contribute, through its continued development of research methodologies specifically dedicated to this purpose, is in the empirical investigation of the psi hypothesis, and one would hope that just as much effort is expended in recruiting the help and participation of other disciplines towards this goal.

**Publication Details:** *Thalbourne, M. A., & Storm, L. (2004). Parapsychology in the 21st century: Essays on the Future of Psychic Research. McFarland & Company. 392 pp. ISBN: 0786419385. Publication Price: £32.00 approx. (paper).*

### References

Beloff, J. (1990). *The relentless question: Reflections on the paranormal*. London: McFarland.

## Book Review

Ian R. Hume and Tony R. Lawrence  
Department of Psychology, Coventry University, Coventry

### **A Review of “From Shaman to Scientist: Essays on Humanity’s Search for Spirits”, by Edited by James Houran (2004)**

This book, edited by James Houran, formerly of Southern Illinois University School of Medicine, follows on from his previous coedited volume *Hauntings and Poltergeists: Multidisciplinary Perspectives* in that it provides a collection of chapters from a diverse range of author expertise focusing on the issue, as the subtitle demonstrates, of ‘...humanity’s search for spirits’. *Hauntings and Poltergeists* provided an excellent portfolio of chapters addressing its title matter, and one came to this book with the hope that these new offerings would live up to the same standard.

The opening chapter, ‘Science in Search of Spirit’, by Annekatrin Puhle and Adrian Parker, seems closely allied to the title of the book. It begins with a brief consideration of shamanism, before moving on to the gradual development from the animistic mode of thought into the age of enlightenment and its turn towards rationality and empiricism, culminating in the science of parapsychology. Interestingly, Puhle and Parker ultimately turn this potted history on its head, by arguing quite reasonably that the spiritual and ritualistic aspects so central to the shamanic pursuit of spirits may yet prove to be a critical factor in enabling experimental success in the more scientific and empirical field of parapsychology.

In chapter two, James McClenon gives a rather nice account of how shamanism began. He asserts that shamanism was the foundation of

all religious forms and that the capacity for dissociation and trance has been selected for due to the benefits linked to therapeutic suggestion (hypnosis). Thus, fitness was enhanced for those that were more hypnotisable and therefore the frequency of genes responsible for this ability increased in the gene pool. McClenon goes on to argue that the increasing frequency of genes (and consequent changes in physiology) for hypnosis/dissociation affected the frequency and nature of anomalous experiences such as apparitions, ESP, PK and OBEs. This, in turn resulted in an increase in belief in such phenomena due to the possible benefits to be gained from holding such beliefs.

As stated previously, this is a nice account of how shamanism began but unfortunately McClenon does not consider opposing views to certain aspects of his arguments. For example, can we be certain shamanism is the foundation of all religions given the complexity of the various forms of religion? We cannot. McClenon also seems to skim over the difficulties of defining hypnosis rather too quickly and spends little time considering alternate views to his own. A discussion of what hypnosis is thought to be would have made this essay more useful to the reader and allow them to evaluate McClenon's claims more readily. Indeed, the reader would have been able to recognize that McClenon's account of how shamanism began could not be accepted with as much confidence as he would like them to.

In chapter three, on 'Spirits as Human Nature and the Fundamental Structure of Consciousness', Michael Winkelman begins by pointing out the universality of the experience of spirits across cultures, and from that wishes to propose that spirits are '... a necessary consequence of the operation of the brain in making sense of the world through the inevitable models of human nature that filter our perceptions.' (p.60). Winkelman argues that his theory is a biopsychosocial one, as opposed to a merely anthropological one. The argument that spirits represent universal and foundational features of the human psyche is not new; Jungian psychology would see the spirit experience as the expression of the archetypes, but Winkelman, clearly, seeks to put some flesh on the psychical bones of the spirit experience by accounting for spirit experiences in terms of the notion of 'neurognostic structures'. So far so good, Winkelman's opening arguments are reasonable in principle, and resonate as Winkelman notes with Jung's notions of the archetypes (particularly the archetype of spirit).

However, the rest of Winkelman's paper proves disappointing. He

ranges far and wide in his search for biopsychosocial account of the spirit experience, as must necessarily be the case, but is never capable of presenting a clear, coherent statement of his theory. At times his paragraphs are quite garbled and overly wordy, and the sum total amounts to an equally garbled and wordy proto-theory that one suspects has some merit, providing some clarity and incisiveness could be brought to bear on proceedings. We shall illustrate with an example.

Winkelman makes three basic propositions, and takes them to be foundational in any notion of what it is to be human (summarised below):

1. Humans have self-consciousness.
2. Humans have other-consciousness.
3. Humans have a sense of embodiment (Winkelman shows his characteristic wordiness when dealing with embodiment — we leave it to the reader to decide whether my summary accurately reflects what Winkelman does write).

Winkelman believes that our normative possession of these three properties means that we may, under certain conditions (never quite clearly spelled out but having something to do with brain ‘modules’ that form ‘neurognostic structures’) project our sense of self, other and embodiment out in to the environment in the form of the spirit experience. It seems rather like the notion that spirits are ‘brain farts’ that occur when the self/other/body perceptual systems go awry!

Interestingly, Winkelman does not adequately spell out one obvious logical corollary of his claims. Our sense of self, other and embodiment are *developmental* features of human experience. The newly born child has only a very rudimentary consciousness in which the sense of self and other is largely non-existent. Indeed the neonate’s sense of self is in the early months largely one of bodily-self. But the developmental nature of self, other and body directly suggest that, in keeping with Winkelman’s argument, the experience of spirits should *itself* be a developmental feature of human experience. This argument is not given by Winkelman though, even though it is a clearly testable prediction. Instead Winkelman moves on to rather broad, and once again turgidly wordy claims based on the notion of modularity of cognitive processes and their likely development from an evolutionary perspective. He also

tends to issue statements as though they were categorical fact, but which are on deeper analysis likely to be none other than as-yet unsubstantiated conjecture (his comments about the relevance of modularity, a concept from cognitive neuroscience, reveal an acquaintance with the basic literature but little more and completely ignore the contentions surrounding the notion of modular psychological functions). Although the chapter becomes a little more coherent as it moves on to topics which are more anthropological in tone, the overall sense from this chapter is that Winkelman's basic premise is interesting, but that he overstates his case, which seems built upon a far too impressionistic reading of a literature that is inevitably as deep and complex as it is broad.

In chapter four Christa Tuczay deals with ancient and medieval interactions with the dead. This is an interesting read which shows how many of the ancient beliefs have filtered through to modern day popular culture.

In chapter five, Peter Mulacz provides an overview of poltergeist researchers from the opening years of the twentieth century to present. Mulacz focuses on the reasoning, theory building, and research findings of these researchers rather than their outward lives. This chapter is an extremely valuable source of reference as many of the researchers are from mainland Europe and thus their work is unlikely to have been studied by the majority of us who can only read English.

Chapter six is by anthropologist Sylvia Grider and deals with American Children's Ghost Stories: Manipulation and Mastery of a Belief System. It is an interesting and well written chapter, but one suspects that for readers of this journal it is of extremely supplementary value.

The final chapter is by John Potts who has looked at a number of websites devoted to ghost hunting. He is eager to point out that many of the ghost hunting groups hold pseudoscientific ideas that often gain the attention of the media and thus the subject is not taken as seriously as it should be. He states that parapsychological researchers should publicly oppose ill-informed ghost-hunters. Surely it would be more sensible to work with these people rather than against them. Potts doesn't state whether he has tried this approach or whether he has ever been on a ghost hunt. Parapsychology struggles to find funding for such research and here we have groups with the enthusiasm and contacts to allow us to do research. Personally, I (Hume) have found that many groups are more than welcoming and quite open to various explanations for certain



phenomena. They have also shown considerable patience when I have been setting up equipment. They don't have to. It seems that they are willing to play ball. Further, they deserve explanations for phenomena as much as the rest of the public do. That is, if we can provide them.

The book includes a light and breezy Foreword by Professor Richard Wiseman and concludes with an Afterword by Lloyd Auerbach. Surveying the book as a whole, we feel that the book does not match up to the standards set by *Hauntings and Poltergeists*. This volume seems rather cobbled together, and suffers from a distinct lack of editorial direction. It has some good chapters, and some relatively bad chapters, but one can't escape the feeling that this is a book of limited utility to the furtherance of knowledge of spirits or to those who seek to study them by whatever means. With little by way of an overarching aim under which its chapters could be said to cohere, it ultimately seems like a book for its own sake, and very poor value for money. Priced extortionately at £31 for the paperback version, there seems little reason to rush out and add this book to one's collection.

*Publication Details:* Houran, J. (2004). *From Shaman to Scientist: Essays on Humanity's Search for Spirits*. The Scarecrow Press, Inc., Lanham, Maryland. xvii + 268 pp. ISBN: 0810850540. Publication Price: £31.00 (paper).

## Obituary for Piero Cassoli (1918–2005)

Carlos S. Alvarado  
Parapsychology Foundation, New York



Parapsychology has lost a highly energetic and dedicated member of its community with the death of Italian physician and parapsychologist Piero Cassoli on August 29 of the current year in Bologna. Dr. Cassoli had been a Full Member of the Parapsychological Association since 1965 and was awarded the PA's "Outstanding Career Award" in 2000. His work was instrumental in the development of parapsychology in his country in ways that few others can claim.

Dr. Cassoli was born on July 25, 1918. He obtained an MD degree in 1943, and in 1948 he married Brunilde Mignani, his life companion and colleague. Dr. Cassoli became interested and active in parapsychology in the late 1940s. Once associated with the Associazione Italiana Scientifica di Metapsichica (Italian Scientific Association of Metapsychics),

---

Correspondence details: Carlos Alvarado, Parapsychology Foundation, P. O. Box 1562, New York, NY, 10021-0043, USA. Email: [alvarado@parapsychology.org](mailto:alvarado@parapsychology.org).

founded in 1937, he presented a paper at an important meeting that the association held at the University of Bologna in October of 1953.

Some years later Dr. Cassoli's prominence led to two invitations to participate in conferences sponsored by the Parapsychology Foundation. The conferences, held at St. Paul de Vence, were the "International Research Coordination Meeting" (1957) and "Parapsychology Today: A Geographic View" (1971).

During his career Dr. Cassoli met many important figures in Italian parapsychology, such as Ferdinando Cazzamalli and William Mackenzie. Both he and Mrs. Cassoli enjoyed many years of friendship with Emilio Servadio.

With PA members Brunilde Cassoli and Dr. Enrico Marabini, and others, Dr. Cassoli founded the Centro Studi Parapsicologici (CSP, Center of Parapsychological Studies) at Bologna in 1954. The CSP, still active today (<http://digilander.libero.it/cspbologna/>), has been very important in the development of modern Italian parapsychology. Its story was recorded in *Un Sole Nascosto: Cinquant'Anni di Vita del Centro Studi Parapsicologici di Bologna* [A Hidden Sun: Fifty Years of Life of the Center of Parapsychological Studies of Bologna], a book authored by Brunilde Cassoli and Paola Righettini (Rome: Phoenix, 1999). This book clearly shows the important role that Dr. Cassoli played in the CSP. Not only was he president of the organization for many years, as well as a member of its board of directors at the time of his death, but he was also actively involved in its publications and other activities. These activities included a variety of one day conventions and discussion forums. Ten of these, held between 1983 and 2000, were called "Giornata Parapsicologica Bolognese" (Bolognese Parapsychological Day). As Cassoli stated in the 1986 issue of the CSP's journal, *Quaderni di Parapsicologia*, the "Giornata" was a means by which those interested in parapsychology in Italy could be brought together to exchange ideas irrespective of their differences. Such meetings did much to create and maintain a sense of identity for Italian parapsychologists, which was one of Dr. Cassoli's most important contributions to the field.

In addition to work on organization and development of parapsychology in Italy, Dr. Cassoli made many substantial contributions to the field. He is particularly well known for his interest in and work with healing phenomena. His book *Il Guaritore* (The Healer, 1979, new ed. 1985) is a modern classic that, unfortunately, is virtually unknown by those who do not read Italian. The book has sections on the history

and legal aspects of healing, and on laboratory research in PK over biological targets. But Dr. Cassoli's unique contribution consisted of the presentation of several clinical cases classified according to the quality and detail of their medical documentation. The topic was discussed further in another book, *Ricerche sulla Pranoterapia e sui Guaritori* (with Giovanni Iannuzzo, 1983), and in an English language paper published in the *European Journal of Parapsychology* in 1981.

From early on Dr. Cassoli was involved in experimental work. For example in 1954 he published, with Dr. Enrico Marabini as second author, a paper in the journal *Metapsichica* about ESP experiments conducted in Bologna between 1952 and 1953 which used multiple agents. Following Warcollier, they evaluated the data qualitatively, considering not only direct hits, but partial information such as the notion of color, movement, or part of a shape. They attempted to create a dynamic situation in an experimental context that was similar to the contexts of spontaneous cases by having the agents getting emotionally involved with the targets. In one test the target was a drawing of medicine containers and the words "with the pharmacist," typed on the target. Dr. Cassoli pretended he was the pharmacist, while another agent pretended he was the cashier. Other persons acting as additional agents came in pretending they were customers. The whole point was getting the agents involved in ways that went beyond mere concentration on a target. The results were not overly impressive, but there were several suggestive effects.

In later studies Dr. Cassoli and others studied particular phenomena and gifted individuals. For example, there was the case of traveling-clairvoyance-like studies with Pasqualina Pezzola, palm reading studies with Maria Gardini, psychometry studies with Luisa Godicini, and the study of the mediumistic paintings of Giuseppe Lanzillo. Dr. Cassoli also observed the controversial performances of Uri Geller and Gustavo Adolfo Rol. Furthermore, he was interested in firewalking.

The CSP also conducted some studies of ESP and LSD that did not obtain significant results and were never published. In the above mentioned book about the CSP by B. Cassoli and P. Righettini, the authors noted that Cassoli was the first one to ingest LSD, which took place in November of 1960. Also present in those tests were Emilio Servadio and Roberto Cavanna, among others.

Spontaneous phenomena also interested Dr. Cassoli greatly. In addition to poltergeists, he studied "miraculous" phenomena. In one of

his most fascinating papers, published in 1955 in *Luce e Ombra*, Dr. Cassoli analyzed the case of a young Italian woman who had a “crying” crucifix as well as visions of Mary. He argued that the phenomena took place in a propitious psychosocial context. First, the family in question had strong beliefs in the virgin. Second, the manifestations followed a well-publicized Italian case of a Marian apparition with which the family of the young woman had contact. Cassoli, who seemed to believe the phenomena were not faked and could not discover a conventional explanation for the “crying” crucifix, proposed that the events were related to specific social and cultural aspects capable of modeling, inducing and maintaining the phenomena in question. Another case was an investigation, published with Mrs. Cassoli in 1959, of the “weeping Madonna” of Rocca Corneta.

Dr. Cassoli was also known to be very supportive of newcomers to parapsychology. In addition, he did much to bring parapsychology to the general public. An important event was a course of ten lectures given by CSP members from January to May of 1975 to around 300 students at the Institute of Pharmacology of the University of Bologna. Dr. Cassoli presented three of the ten lectures. He gave a general introduction that opened the course, and spoke later about telepathy and theory. This is only one of the numerous educational activities conducted by Dr. Cassoli and the CSP.

Dr. Cassoli also did much work in popular forums. He wrote for such popular Italian magazines as *ESP* and *Giornali dei Misteri*, and had a section in the latter where he answered correspondence sent by readers of the magazine. Fifty-four of these letters form the text of the book *Lettere a un Parapsicologo* (Letters to a Parapsychologist, 1973), edited by Brunilde Cassoli. The correspondence was mainly concerned with the readers’ spontaneous psychic experiences. Dr. Cassoli offered advice and information about relevant parapsychological studies.

The field was popularized in unprecedented ways when Dr. Cassoli and collaborators conducted an informal radio ESP test in 1972. Over 7500 replies were received. More recently Dr. and Mrs. Cassoli published a small volume to serve as an introduction to parapsychology for the general public, *La Parapsicologia* (Parapsychology, 2000).

Furthermore, Dr. Cassoli communicated with the Italian medical world through *Minerva Medica*, one of Italy’s main medical journals. Between 1956 and 1961 the journal carried an occasional section in charge of the CSP devoted to parapsychology. In 1983 the *Enciclopedia Medica*

*Italiana* (Italian Medical Encyclopedia) carried an article about parapsychology written by Dr. Cassoli.

Dr. Cassoli was very interested in PA affairs. He wrote to me in my capacity as PA President about the Association, as well as to other PA officers, with ideas to improve the PA. He was very appreciative of the significance of the PA. In an email he sent me on May 11, 2003 he said: "For us 'from the periphery'... [the] PA is a beacon, a point of reference of which we are proud."

In an editorial in the 1990 *Quaderni di Parapsicologia* Dr. Cassoli described himself as an "old 'torch' " near extinction. While his life may be over, his "torch" is very much with us in terms of the example he gave us. I always saw Dr. Cassoli's work as an example of great tenacity, energy and positive disposition. Here was one man who, together with a small group of collaborators, actually made a difference for parapsychology in Italy. Although I had intermittent correspondence with him over the years, and he sent me copies of most of his books, I only met him once. The meeting took place at Riccione, Italy, where Nancy Zingrone and I were attending a three-day popular conference in March of 1995. He arrived, accompanied by Brunilde Cassoli and Dr. Enrico Marabini. If I remember correctly, they had just been studying some poltergeist manifestations. Dr. Cassoli, deeply interested in international and political aspects of parapsychology, engaged in conversation and at one point, in a good-natured way, he literally "grilled" us with specific questions about some political issues involving a particular figure and organization in the field. His interest, however, transcended gossip. He was truly interested in parapsychology from a global perspective and liked to be informed about its dynamics and events. To this day this encounter remains one of my best memories of my visit to Italy.

To remember Dr. Cassoli's optimism and tenacity we cannot do better than read his words of hope. Referring to the "Giornata" meetings, but relevant to the field at large, in 1986 Dr. Cassoli wrote in the *Quaderni di Parapsicologia* that the work must go on: "The obstacles are many... It does not matter. We will continue in this road. Time and our constancy will demonstrate our good faith."

## Acknowledgements

I wish to thank Dr. Massimo Biondi for providing me with important information about Dr. Piero Cassoli's life and work.

## Selected Bibliography of Piero Cassoli

### *Books*

- Cassoli, P. (1973). *Lettere a un Parapsicologo* [Letters to a Parapsychologist] (B. Cassoli, ed.). Firenze: Corrado Tedeschi.
- Cassoli, P. (1979). *Il Guaritore* [The Healer]. Milan: Armenia.
- Cassoli, P. (1985). *Il Guaritore* [The Healer] (new ed.). Milan: Arnoldo Mondadori.
- Cassoli, P., & Cassoli, B. (2000). *La Parapsicologia* [Parapsychology]. Milan: Xenia.
- Cassoli, P., & Iannuzzo, G. (1983). *Ricerche sulla Pranoterapia e Sui Guaritori: La Pratica e i Risultati Valutati dalla Scienza* [Research on Pranotherapy and Its Healers: Practice and Results Evaluated by Science]. Como: Red.

### *Papers*

- Cassoli, P. (1954). Esiste la bilocazione? Studio di una sensitiva marchigiana. [Does bilocation exist? A study of a sensitive from Marche]. *Metapsichica*, 9, 72-84.
- Cassoli, P. (1954). Indagine clinico-metapsichica sul caso della miracolata Cristina B. di Bologna [Clinical-metapsychical investigation of the case of the miraculously cured Cristina B. of Bologna]. *Metapsichica*, 9, 109-123.
- Cassoli, P. (1955). Di un caso di medianità a estriseccazione religiosa: Il "contagio psichico" in metapsichica e Angelina Ronza da Frignano D'Aversa [On a case of mediumship with religious expression: "Psychic contagion" in metapsychics and Angelina Ronza da Frignano D'Aversa]. *Luce e Ombra*, 55, 129159.
- Cassoli, P. (1955). Il movimento parapsicologico nel mondo: Panorama storico ed orientamenti attuali: Conferenza d'apertura dell'anno accademico 1955-56 al Centro Studi Parapsicologici [The parapsychological movement in the world: Historical panorama and current approaches: Opening lecture of the academic year 1955-56 to the Center of Parapsychological Studies]. *Minerva Medica*, 46, 1050-1060.
- Cassoli, P. (1958). La pirobazia in Grecia (le Anastenaria) [Firewalking in Greece (the Anasteria)]. *Minerva Medica*, 49, 3677-3686.
- Cassoli, P. (1960). II serie di esperienze di telepatia collettiva (eseguite nella città di Bologna) [Second series of experiments of collective telepathy (at the city of Bologna)]. *Minerva Medica*, 51, I-XII.
- Cassoli, P. (1973). Parapsychology in Italy today. In A. Angoff & B. Shapin (Eds.), *Parapsychology Today: A Geographic View* (pp. 187-197). New York: Parapsychology Foundation.
- Cassoli, P. (1976). Guaritore e fenomeni fisici (note preliminare) [Healers and physical phenomena (preliminary note)]. *Metapsichica*, 30, 55-61.
- Cassoli, P. (1977). "Il parapsicologo" [The parapsychologist]. *Metapsichica*, 32, 41-46.
- Cassoli, P. (1981). The healer: Problems, methods and results. *European Journal of Parapsychology*, 4, 71-80.

- Cassoli, P. (1983). Parapsicologia [Parapsychology]. In *Enciclopedia Medica Italiana* (Vol. 11). Torino: UTET.
- Cassoli, P. (1984). Una "fascia inesplorata" della psicologia: La pittura medianica [An "unexplored aspect" of psychology: Mediumistic painting]. *Bolletino del C.S.P.*, No. 9, 21-25.
- Cassoli, P. (1986). Camminare sul fuoco: Un antico rito oggi di moda: Interpretazioni e commenti [Firewalking: An ancient rite that is fashionable today: Interpretation and comment]. *"Bolletino:" Organo di Informazioni del C.S.P.*, No. 12, 26-30.
- Cassoli, P. (1986). Contributo allo studio della pranoterapia (contributo alla casistica, al laboratorio, alla interpretazione del fenomeno) [Contribution to the study of pranotherapy (contribution to case and laboratory studies, and to explanations of the phenomena)]. *Quaderni di Parapsicologia*, 17, 69-89.
- Cassoli, P. (1986). Immagini e voci indeterminate [Indeterminate images and voices]. *"Bolletino:" Organo di Informazioni del C.S.P.*, No. 13, 25-29.
- Cassoli, P. (1986). Introduzione – Saluto del Presidente [Introduction – Greetings from the President]. *Quaderni di Parapsicologia*, 17, 5.
- Cassoli, P. (1988). La "medianità" di G.A. Rol: Fatti e commenti da un libro di Renzo Allegri [The "mediumship" of G.A. Rol: facts and comments on a book by Renzo Allegri]. *Quaderni di Parapsicologia*, 19, 9-20.
- Cassoli, P. (1989). Parapsicologia e scienza [Parapsychology and science]. *"Bolletino:" Organo di Informazioni del C.S.P.*, No. 19, 34-42.
- Cassoli, P. (1990). Editoriale: I "Quaderni di Parapsicologia" rivista periodica del C.S.P. [Editorial: The "Quaderni di Parapsicologia" periodical journal of the C.S.P.]. *Quaderni di Parapsicologia*, 21, 5-7.
- Cassoli, P. (1991). La parapsicologia nella Repubblica Popolare Cinese: Da una relazione di L. Zha e T. McConnell [Parapsychology in the Popular Republic of China: On an account by L. Zha and T. McConnell]. *Quaderni di Parapsicologia*, 22, 13-22.
- Cassoli, P. (2001). Circuito chiuso: Il futuro della parapsicologia secondo Rhea White [Closed circuit: The future of parapsychology according to Rhea White]. *Quaderni di Parapsicologia*, 32, 94-105.
- Cassoli, P. (2002). La metanalisi in parapsicologia [Meta-analysis in parapsychology]. *Quaderni di Parapsicologia*, 33, 10-14.
- Cassoli, P. (n.d.). Quando i fenomeni si definiscono paranormali [When the phenomena are defined as paranormal] (<http://digilander.libero.it/cspbologna/cspnew1/speciali/cassoli-quando.htm#inizio>).
- Cassoli, P., & Iannuzzo, G. (1988). La medianità: Introduzione al tema [Mediumship: Introduction to the topic]. *Quaderni di Parapsicologia*, 19, 76-79.
- Cassoli, P., & Marabini, E. (1954). Esperienze di telepatia collettiva nella città di Bologna [Experiments of collective telepathy in the city of Bologna]. *Metapsichica*, 9, 3954.
- Cassoli, P., & Marabini, P. (1957). Esperienze triennali di lettura della mano con una sensitiva Bolognese, Maria Gardini: Proposta di nuovo metodo [Triannual experiments of palm-reading with a Bolognese sensitive, Maria Gardini: Proposal of a new method]. *Minerva Medica*, 48, 2130-2134.





