

ELEONORE ZUGUN – THE RE-EVALUATION OF AN HISTORIC RSPK CASE¹

Peter Mulacz

Austrian Society for Parapsychology

SUMMARY

The Zugun case, world famous – and much disputed – in its day, is unique in several aspects, primarily

- in the sheer number of phenomena recorded (more than 3000, out of which 844 are extremely well established)
- in the fact that the focus person lived close together with the prime researcher, Zoë, Countess Wassilko, sharing even their room for a period in excess of one year
- and in the methodical approaches implemented in the course of its investigation:
- the attempt to communicate with the unconscious of the focus person by various means, not least in order to provoke phenomena
 - the attempt to transform the character of the phenomena systematically from spontaneous ones to séance phenomena
 - a psychoanalysis of the focus person (this case appears to be the first one where psychoanalysis has been applied on a RSPK focus person)
 - confrontation of the focus person with other mediums or psychics
 - cinematographic documentation of part of the phenomena.

In other aspects, however, the Zugun case fits well into the general character of RSPK cases, e.g. in respect of the age of the focus person at the onset of the poltergeist phenomena, the bandwidth of categories of phenomena observed, etc.

With regard to the sociology of science (i.e. of Psychological Research), the impact of this case may be seen as dialectic as the above: on one hand – one is tempted to say: ‘as usual’ – an eventual exposure of the focus person with an ensuing endless discussion between the two camps, and on the other hand the unique fact that this case ultimately was to become the cradle of the later Austrian Society for Psychological Research (now Austrian Society for Parapsychology and Border Areas of Science).

The events in short: the case commenced mid-February, 1925 in the Bukowina (Romania), three months before Eleonore Zugun’s 12th birthday, and it lasted for some two years. (In comparison to the material of some RSPK databases, the focus person may be called rather young when the phenomena started, and the duration of the phenomena which terminated shortly after her first

¹ The author would like to gratefully acknowledge the support of this project by a grant provided by the “Institut für Grenzgebiete der Psychologie und Psychohygiene” (Institute for Border Areas of Psychology and Psychic Hygiene), Freiburg i. Br., Germany.

menstruation, may be called rather long.)

The poltergeist phenomena started with inexplicable movements of various objects, throwing of rocks as well as locomotion of household items. These were attributed by the superstitious peasant populace to the devil, in Romanian "Dracu", based on an alleged remark made by Eleonore Zugun's grandmother early in this case who might have induced a devil complex in the girl by that malediction.

When the case came to the knowledge of the regional media, the then eminent German Psychic Researcher, Fritz Grunewald was dispatched to the scene where he was able to establish the occurrence of "genuine paranormal phenomena". As Grunewald happened to die but short time later, the Countess Wassilko whose family used to live for centuries in the Bukowina (the easternmost province of the Austro-Hungarian Monarchy before WW I) firstly visited the girl in her native environment and subsequently took her to Vienna into her own household some time afterwards (end-January, 1926). After a period of some eight months there, she took her for another five months on an extended voyage through several European countries, thus enabling many Psychic Researchers as well as interested laypersons (several hundred people altogether) to witness the Zugun phenomena which had at that time actually changed from the locomotion of small objects, mainly interpreted as apports, to dermatographic phenomena (scratches and bites all over her face and her arms). Thus, the Zugun case shows several distinct phases. The case ended practically with the girl's first menstruation, after which there was a quick decline in the number of phenomena. Following that – the "case" being no more a case, only a biography – the girl spent a few more months in Vienna, finishing her training as a hairdresser before eventually returning home in 1928, where she started working, later became married (with no children), widowed, etc., living a "normal" life.

My own research in the Zugun case – supported by a grant provided from the *Institut für Grenzgebiete der Psychologie und Psychohygiene* – concentrated on two phases, the period in Vienna (= Phase I) and the journey (= Phase II), with emphasis on the former, and focussed on addressing *inter alia* the following problems:

- establishing the total number of phenomena (Phase I: 1754, Phase II: 1306, totalling 3060, as opposed to Countess Wassilko's estimate of 5000 approx)
- evaluation of the original reports in order to identify indicators for possible fraud (or otherwise pointing towards the "genuineness" of the phenomena):
 - the original researchers, i.e. Countess Wassilko and her colleagues, had already pointed out that Eleonore Zugun sometimes, if not controlled properly, would resort to trickery, i.e. they viewed this case as a "mixed" one
 - the analysis of the documentary film did not reveal any indicators of fraud
 - in particular, the film shows Eleonore wearing no rings on her fingers, hence the allegation by Dessoir of her fraudulently producing the scratches by pointed fingernails or the sharp-edged setting of her ring, as still perpetuated in Kurtz's "Skeptic's Handbook", must strongly be refuted
- scrutiny of Rosenbusch's alleged exposure, the analysis of which showed that this claim is not tenable; based on his skeptical belief system, Rosenbusch mistook harmless touches evoking reflex dermal phenomena for fraudulent scratches, etc.
- (translation and) evaluation of Eleonore Zugun's psychoanalysis:

Re-evaluation of an historic RSPK case

- This analysis, however, has been carried out inexpertly, and I have reasons to presume that a few of the underlying complexes (i.e. an alleged rape and an incestuous episode) attributed to Eleonore by the Countess might have been connected to her own unconscious. This supposed countertransference is likely to have been instrumental in the transformation of the Zugun phenomena from apports or locomotion of objects to dermatographic phenomena
 - though being very likely, it could, due to the lack of sources, still not positively be established whether the Zugun case was brought to the attention of Sigmund Freud himself
- investigation of the frequency distribution of the phenomena, the question being whether the distribution of maxima and minima showed an internal periodicity or any correlation with external variables (or is entirely random):
- Countess Wassilko, after observing the phenomena for two months, hypothesised that the frequency distribution, showing one distinct maximum per month, was an anticipation of the female cycle
 - Schrenck-Notzing, who, by the way, supported the Countess' research by a substantial grant, hypothesised that the maxima of phenomena appeared to coincide with the full moon

Neither of these contemporary and somewhat premature hypotheses, formulated after too short a period of observation, could be supported. The distribution of phenomena during the entire time of more than one year does not show any periodicity that comes even close to an equivalent of the menstrual cycle nor could any correlation be found when probing modern hypotheses (correlation with maxima/minima of the geomagnetic field [Persinger] the values of which were supplied by the World Data Center C1 for Geomagnetism, Copenhagen, Denmark; or LST [Spottiswood]). The correlation coefficient for the geomagnetic hypothesis is $r = 0.085$ and 0.021 respectively (for the two phases ea.), whereas for the hypothetical dependency upon the phases of the moon the respective values are $r = 0.028$ and $r = 0.015$. Hence, it has been established that these external variables had no influence on the frequency of the Zugun phenomena whatsoever. The circadian distribution is significantly different b/n Phase I and Phase II, due to the different circumstances of everyday life, obviously dependent upon the Countess' daily rhythms, i.e. a social variable. Moreover, it could be demonstrated that there is an apparently strong albeit hard-to-quantify correlation with purely psychological variables, such as suggestions, or skin contact, etc. that had impact on the phenomena, on both their frequency and their character, e.g. reflex reactions on being touched.

After and above all, this extraordinary case offers the feasibility to discuss all the above on two (or even more) levels, i.e. the case as such, and the case as a prime methodological example for applying quantitative methods on a hitherto apparent qualitative case, thus combining proof-oriented and process-oriented aspects. This is an approach to historic cases, which is different from mere reception studies. It demonstrates that the existing abundance of historic cases in parapsychology is a most valuable "treasure" that needs to be re-evaluated periodically – a permanent process –, and that the gap between idiographic and nomothetic approaches can indeed be bridged.

THE JACKSONVILLE WATER POLTERGEIST: ELECTROMAGNETIC AND NEUROPSYCHOLOGICAL ASPECTS

Andrew Nichols
City College, Gainesville, Florida
&
William George Roll
State University of West Georgia, Carrollton, Georgia

ABSTRACT

Apparent poltergeist-type occurrences were reported in a private residence in Florida from November 1996 to January 1997. In addition to reported movement of objects and other typical RSPK manifestations, unexplained appearance of water, sometimes in copious amounts, were described. The family believed the house to be "haunted."

Physical measures and psychometric techniques were employed to study the phenomena and to attempt to understand the psychological, neurophysiological, and physical processes that might underlie them.

The qualitative part of the investigation involved testing the premises with a magnetometer and non-contact thermometer. The location was photographed with Polaroid, 35mm, and infrared film. Detailed psychological examinations were administered to the three primary family members involved. It was predicted that one or more of the family members would reveal the type of personality profile--i.e. low tolerance for frustration and denied feelings of hostility--which have been associated with poltergeist agents. It was also predicted that one or more of the primary witnesses would show neuropsychological signs of unusual temporal lobe activity. Possible correlations between geomagnetic/electromagnetic activity and the alleged paranormal events were studied.

Responses to the projective psychological tests indicated that one of the witnesses conformed to the profile of a "poltergeist personality". Bender, Owen, Palmer, Rogo and Roll have described psychological profiles of suspected poltergeist agents; typically the agent is an adolescent with a low tolerance for frustration, repressing feelings of aggression and hostility. In the present study, the neuropsychological questionnaires suggested possible temporal lobe lability in two of the witnesses, including the apparent poltergeist agent. Magnetic field strengths at several locations where unexplained phenomena had been reported differed significantly from magnetic field strengths in locations where no unusual phenomena had been reported.

INTRODUCTION

Unexplained influxes of water have occasionally been reported in conjunction with poltergeist-type occurrences, although the phenomenon appears to be relatively rare. The amount of liquid may range from only a few drops, to a near-deluge.

An early case, described in some detail in the British spiritualist journal, *The Annals of Psychic Science* (April, 1906) included the usual unexplained sounds, displacement of furniture, as well as abnormal sprinklings of water. Fort (1974) chronicled sixty cases involving appearances of water

The Jacksonville water poltergeist

and other phenomena. Thurston (1959) described paranormal sprinklings of water in conjunction with object movements in the presence of poltergeist-medium Frieda Weissal. A San Francisco (USA) poltergeist case which included inexplicable water manifestations was reported by Auerbach (1996).

Gauld and Cornell (1979) found 26 cases of poltergeist and haunting (out of 500) that involved water and gave a detailed description of a 1963 case near Bath, England (pp. 301-307). In most instances, including the bath case, water poltergeists incorporate typical RSPK effects, such as object-movements. Two cases, however, were restricted to water, the German Scherfedede case (Bender, 1974, pp. 138-141) and one mentioned by Cox (1961, pp. 68-69).

Bayless (1967) and Rogo (1986) described a 1963 poltergeist infestation in which a family in Methuen, Massachusetts, were driven from their home when jets of water began springing out from the walls of the house. When they relocated to the home of a relative, the watery manifestations also began occurring there.

The Methuen case contained several similarities to the present study. In both cases the focus of the phenomena appeared to be an eleven-year-old female living in a multi-generational household. In both cases plumbers, contractors, and other experts were consulted to preclude mundane causes for the water manifestations. In contrast, the water jets reported in the Methuen disturbances were reportedly preceded by explosive sounds, while no such auditory phenomena were reported in the present case.

Description of the Case

In November, 1996, AN was contacted by Mr. James Holland, Senior loss analyst for a major insurance company. Mr. Holland had been investigating a claim for water damaged furniture and carpet submitted to his insurance company, and had been unable to determine the origin of the damage. Since the family who submitted the claim believed the water damage to have been paranormally produced, Mr. Holland contacted AN and requested that we investigate the case from a parapsychological perspective. After conducting a preliminary telephone interview, we decided to conduct an on-site investigation of the house, which was located in a suburb of Jacksonville, Florida.

After lengthy telephone interviews with several of the witnesses, AN and a colleague, Russell McCarty, visited the house on December 17, 1996. During the investigation, which lasted until January 1997, AN kept in constant contact with the family by phone. Since the PK was episodic and unpredictable, it was felt that frequent visits to the home would not be particularly fruitful. The house was a modern, seven year old single-story home in an affluent suburban neighborhood. The house included five bedrooms, a living room, dining room, kitchen, three bathrooms, and a large attached garage and swimming pool. The occupants were Mary Barton, a 62 year old professional woman; her mother Lillian, age 87; her son Kieth, age 28; and her granddaughter Krista, age 11 (pseudonyms). We interviewed these four witnesses and took photographs and magnetometer readings of the areas where incidents had been reported.

Interviews with Witnesses

AN and RM interviewed four witnesses: Mrs. Mary Barton, the head of the household; her

mother Lillian; her son Keith; and granddaughter Krista. Mary Barton is a 62 year old, self-employed professional financial consultant. She also is the owner of a pet grooming business, and conducts her business away from the home. Lillian Barton is retired and spends considerable amounts of time in the house. The interviews were conducted on December 17, 1996, except as otherwise indicated. We also conducted telephone interviews with Mr. James Holloway, owner of Sani-Pro plumbing contractors, who inspected the house for plumbing problems as a possible cause of the mysterious water appearances (no plumbing problems were found). Mr. Holloway's two assistants, Mike Thigpin and "Butch" Jenkins, were also interviewed. Both of these experienced plumbers had inspected the house and found no source of leakage to account for the copious amounts of water they both witnessed throughout the house. A Duct contractor, Mike Cavaluzzi (pseudonym) was also hired to inspect the air conditioning system for faulty duct work. None was found but all three of the contractors were wetted by the poltergeist while performing their inspections. Mr. Thigpin stated that his shirt was thoroughly soaked from mid-air from the front while standing in the living room of the house. He was so alarmed by this occurrence that he refused to return to the residence, believing that a supernatural force was present in the house.

Narrative

Mary Barton reports that unusual occurrences began in the house on Thursday, November 7, 1996, when a series of unexplained rapping noises could be heard apparently emanating from several interior walls within her home. The sounds were heard by all family members and they were assumed to have a natural cause. But in spite of the family's efforts to locate the source of the noises, none was ever found.

During this period small objects were occasionally reported to be missing by various family members, only to reappear later in unlikely places. Lillian Barton, Mary's mother, began to report that she had experienced light sprinkles of liquid, apparently water, during the night while in bed and occasionally during the day as she walked through the house. She attributed these sprinklings to faulty air-conditioning/central heating ducts, and recommended that her daughter hire someone to inspect the duct-work. Over the next several days, the sprinklings of water became more frequent, and the quantities of water increased considerably, until puddles of water were seen in several areas throughout the house. In addition, other family members including Mary, her son Kieth, and her granddaughter Krista-- began experiencing the water sprinklings as well. Unexplained cold spots were noticed in several locations within the home, and the instances of apparent object-movement increased. During one occurrence Mary was standing in the kitchen preparing dinner when a small paper cup filled with water which had been sitting on the counter behind her, at a distance of approximately eight feet, suddenly became airborne and lodged itself in the handle of one of the lower cabinets, next to where she was standing. Mary turned around quickly, but could see no one present in the room who could have thrown the cup. When she examined the cup, which was still lodged in the cabinet handle in an upright position, she was astounded to discover that the cup still contained most of the water which it had originally held. Only a slight wet spot above the cup on the cabinet door revealed that a small quantity of water had spilled as the cup had flown across the room.

Over the next several days the frequency and severity of the water influxes continued to increase, with all family members being wetted periodically, and soaked occasionally. Lillian Barton seemed to particularly attract the watery showers and was repeatedly soaked. No one could determine from where the water was originating. A heating contractor called in to inspect the ducts found them completely dry, with no sign of water damage. He concluded there was nothing in the

The Jacksonville water poltergeist

heating/air conditioning system which could account for the mysterious water influxes. A plumbing contractor was also employed to check the pipes, but again no sign of leakage or other plumbing problems were found. During the time that the contractors were present in the house, both men were themselves the target of the water-sprinkler.

By this time the family members were convinced a supernatural agency was responsible for the unusual phenomena in their home, and the furniture and carpets were suffering extensive water damage. Other phenomena occurred during this period, including a potentially dangerous event in which a large, heavy, gilt-framed mirror-- which had been securely mounted over the couch for years-- suddenly fell and narrowly missed Lillian, who was lying on the couch resting at the time. She was able to roll off the couch and avoid the falling mirror, which would undoubtedly have caused serious injury had it struck her. Upon examining the mirror it was found that the wire which secured it to the wall was unbroken, and the heavy mounting screws which protruded from the wall upon which the mirror had been hung were also intact. There appeared to be no reason why the mirror should have fallen from the wall.

The quantities of water were occasionally quite extensive. In one instance Mary, Krista, and Lillian were all drenched so extensively that they fled from the house into the adjacent driveway and climbed into the family van. No sooner had they shut the door of the vehicle than a large amount of water-- described as 'several gallons"-- splashed against the side of the van coming from the direction of the house. Since all family members who were home at the time were present in the van, no one was left in or around the home who could have hoaxed the occurrence even if such a large quantity of water could have been concealed and thrown by a potential prankster.

Convinced now that a "ghost" was inhabiting their home, the family began to speculate about the origin of the occurrences. Lillian Barton appeared to be the one most commonly afflicted by the periodic water-drenchings, and believed she was being "punished" for some past transgression or sin. Eleven-year-old Krista was convinced that the haunting began when she and three of her young friends had conducted a "seance" at her birthday party on October 18, several weeks prior to the onset of the phenomena. During the seance the girls had attempted to contact various deceased historical figures and media personalities. Krista reports that during the seance the girls had invited the spirits present to "haunt us." Krista reported that approximately one-half hour after the conclusion of the seance, several caged birds nearby began flapping their wings frantically and making sounds as if they were suddenly alarmed; simultaneously the lights in the house went out, and the television set came on spontaneously. This frightened the girls, who decided to attempt no more occult activities.

The poltergeist occurrences continued for over one month, with the peak of the disturbances occurring on or about December 6, 1997 (the family was uncertain of the exact chronology and dates of certain events). During this period water appeared in copious amounts throughout the house, with all the family members being subjected to periodic water sprinklings and occasional drenching. Lillian Barton appeared to be the primary target of these occurrences and was repeatedly soaked, according to her testimony. These water attacks seemed to occur at any time of day, although most commonly during daylight hours. Any location within the home or on the grounds nearby were susceptible to water influxes. Lillian Barton stated that she had been sprinkled or drenched while alone in the bathroom with the door locked, while arranging clothes in a walk-in closet located in her bedroom, and in other locations.

Mary Barton had noticed that these disturbances seemed to occur only when her eleven-year-old granddaughter Krista was present in the house, and initially suspected that Krista was physically creating the disturbances. She discounted this possibility when she and other members had experienced water sprinklings and other phenomena during times when Krista was under observation or known to be in another part of the house.

Hypotheses

On the basis of previous investigations of poltergeist-type occurrences and our interviews with witnesses, we focused our investigation on attempting to test the following hypotheses.

- (1) We hypothesized that the house or property where the majority of the occurrences had been reported could be a source of unusual electromagnetic or geomagnetic energies which might trigger such experiences. Previous investigations of poltergeist-type disturbances (Roll and Gearhart, 1974), (Gearhart and Persinger, 1986) had found a positive correlation between the onset of poltergeist disturbances (RSPK) and geomagnetic perturbations. Persinger (1985) suggested that transient geomagnetic anomalies associated with tectonic activity as well as electromagnetic fields produced by artificial environmental sources such as household electrical wiring and proximity to high-voltage transmission towers may trigger RSPK and haunting occurrences.
- (2) We hypothesized that one or more of the primary witnesses to the occurrences would display evidence of temporal lobe lability.
- (3) We hypothesized that the poltergeist-type phenomena expressed the psychological and emotional needs of one or more of the primary witnesses. In previous investigations of hauntings and poltergeist disturbances, the occurrences seemed to be associated with the emotional and psychological needs of the persons experiencing the phenomena.

METHOD

Investigators

AN, the principle investigator, interviewed witnesses, conducted psychological and experimental tests, and conducted tests with a magnetometer, infra-red radiometer, and Polaroid and infra-red photography. RM assisted in the preparation of experimental materials, assisted with witnesses' interviews, and conducted magnetometry and temperature/barometric pressure tests with a hygrothermometer.

Witnesses

Six adults and one adolescent reported unusual occurrences which they had experienced in the house or at other locations in the presence of one of the principle witnesses. Witnesses included the principle witnesses and two independent contractors hired by the owner of the house for the purpose of discovering the source of the mysterious influxes of water.

Materials and Equipment

Magnetometer. Two frequency weighted Tri-Field meters with a magnetic setting calibrated for

The Jacksonville water poltergeist

60 Hz sine waves were used to determine magnetic field strengths at the site.

Infrared Photography. A 35mm SLR camera with 50mm F1.8 lens was used with a Kodak Wratten Filter, No. 25 and Kodak High Speed Infrared Film (HIE 135-36) to make a total of 36 exposures at the site. The film was developed and contact printed at a commercial photo lab.

Electrostatic Voltmeter. An analog electrostatic voltmeter was employed to check surface and ambient electrostatic charge at the site.

Non-Contact Thermometer. A Raytek Raynger non-contact infrared radiometer was used to detect temperature variations at the site.

Hygrothermometer. A digital hygrothermometer was used in conjunction with the non-contact thermometer to obtain temperature and barometric pressure/ relative humidity data at the site.

Radiation Survey Meter. An ion detection radiation survey meter was used to test for ionizing radiation at the site.

Tape Recording. A portable, voice-activated tape recorder was used with audio cassette tapes to record witnesses' reports.

Video recording. A General electric VHS Camcorder was used to videotape the site-- including areas of water damage-- and for recording portions of witness' interviews.

Psychological Tests. Psychological tests were administered to four witnesses at the site. The tests were the Psychic Experiences Questionnaire (Roll and Braun), the Inventory of Childhood Memories and Imaginings Questionnaire (Wilson and Barber), the Neuropsychology Questionnaire (Roberts), and the Kiersey Sorter. In addition, a version of the House-Tree-Person projective test and the Rosenzweig Picture-Frustration Test were administered to two witnesses. AN constructed psychological profiles of the various family members based on their responses to these tests. In order to get an unbiased evaluation of these tests from a psychologist not otherwise associated with the investigation, copies of these tests and responses were sent to Dr. David Bortnick, a clinical psychologist. Dr. Bortnick was asked to evaluate these forms and provide independent psychological profiles of the witnesses on the basis of these evaluations.

PROCEDURE

Investigative site visit. AN and RM drove to the site on December 17, 1996. Upon arrival we were greeted by Mary Barton, who guided us on an orientation tour of the home. A floor plan had been previously constructed of the site based on measurements provided by the witnesses, and the locations of water accumulation and other reported disturbances were plotted on the floor plan. Following the tour of the premises, the magnetometers were used to establish a baseline reading for ambient magnetic/electromagnetic (EM) fields. These ambient EM field measurements were also plotted on the floor plan, along with areas of apparent anomalous EM field measurements.

The witnesses who were present included Mary Barton, her mother Lillian, and her granddaughter Krista. Several hours later her son Kieth also arrived home, but seemed reluctant to participate in the investigation and secluded himself in his room after greeting the investigators briefly. The principal witnesses were taken individually to a rear bedroom for separate interviews which were

recorded on audiotape. After the completion of the individual interviews, the family was assembled in the living room for group discussion, which was videotaped.

SUMMARY OF RESULTS

Magnetometry. Tests of magnetic field strengths at locations cited by witnesses as sites of water appearances or other inexplicable phenomena revealed several areas of unusually high magnetic fields. Mean field strength within the house was 2.7 milligauss, which is consistent with usual background radiation levels within a house of 1-2 milligauss. Readings of 8-10 milligauss are not uncommon and can be obtained, for instance, near a florescent light. In the present case several magnetic field transients were recorded in excess of 60 milligauss. These field readings could not be traced to household electrical wiring, plumbing, or appliances. These magnetometry data are consistent with Hypothesis 1, that sites where poltergeist and haunting activity are reported will exhibit unusual magnetic/electromagnetic properties. It may be significant that, in the present case, high-voltage transmission towers are located within one quarter mile of the afflicted house. In addition, the Jacksonville Naval Air Station is less than five miles away with its array of radar transmitters and EM producing equipment.

Fields generated by such sources may have precipitated the RSPK activity via interaction with the mind of the agent within the home. It may be of interest that water volatilizes at 500 MHZ in the UHF microwave range. At this frequency, the hydrogen bonds break down and water changes to a gaseous state, which may then disperse (1990, National Meteorological Library).

Magnetometry. A t-test of the differences in magnetic field strengths between the locations cited by witnesses as sites of water appearances or other inexplicable phenomena and sites where nothing out of the ordinary had been reported gave $t=4.397$; $p=.003$. Table 1 presents these data.

	N	Mean	Std. Dev.	t (diff) 23df	p
Target Sites	11	9.826	6.143		
				4.397	0.003
Control Sites	15	2.15	2.112		

Table 1: Results for Magnetometry Test

Neuropsychological Aspects. Krista Barton’s responses to the neuropsychological questionnaire included several signs of temporal lobe lability (CPES=32; M=20.4), such as experiencing exactly the same repetitive dream, hearing unexplained sounds of buzzing or sizzling, and a strong sense of deja vue. Krista’s neuropsychological profile was consistent with Hypothesis 2, that persons who experience poltergeist activity would exhibit signs of temporal lobe lability.

Lillian Barton, the other main witness, also completed the neuropsychology questionnaire. Her responses also indicated probable temporal lobe lability (CPES=28; M=20.4). Like Krista she reported repetitive dreams, unusual sensations of cold, and the sense of a “presence” in the house. Unlike Krista, she reported experiencing unusual smells, described as “burning rubber,” when no source for such smells could be found.

Both respondents reported that these sensations were more common when family stress levels

were higher. Stress is likely to increase temporal lobe effects in susceptible individuals. (Roll, et al. 1991)

No signs of temporal lobe lability were evident in the questionnaire completed by Mary Barton, who spent the least amount of time in the house due to her work schedule (CPES=18).

Psychological Tests. In this case, it seemed as though the RSPK was occurring only in the presence of Krista Barton, which indicated that she was the likely focus person within the family. Although Mary Barton already suspected that her granddaughter was somehow at the center of the mysterious occurrences, it was felt that undue attention should not be focused on the young girl-- potentially adding to the psychological stresses contributing to the poltergeist occurrences. For this reason, as well as to understand the family dynamics of this case, I decided to administer psychological tests to the entire family. Each family member agreed to complete these tests with the exception of Kieth Barton, who refused to participate. I will now turn to the results of these evaluations.

The psychological profiles of the three primary witnesses revealed a high degree of aggression in their interpersonal relationships. As would be expected, each family member exhibited different personality traits, but a large number of similar themes are evident in their profiles.

All show unhappiness, insecurity, and recourse to an unrealistic fantasy life. The profile of eleven-year-old Krista suggested each of these traits in addition to aggression directed against others, as well as an inability to handle even minor frustrations and an incapacity to engage in helping relationships with others. Her responses indicated an extraordinarily high aggression level mostly stemming from feelings of personal insecurity. She also tends to blame others for her problems, repressing feelings of anxiety and inadequacy. Krista was also typified as harboring a potential for physical violence, which is kept in check by her reliance on repression. Krista's relationship with her great-grandmother Lillian seemed particularly strained. It is significant that Lillian Barton was most frequently the target of the RSPK activity.

These profiles indicate a family fraught with unresolved conflicts and hostilities. These inter-family relationships were quite apparent to AN when investigating the case. The hostility between Lillian and Krista was particularly evident. Krista seemed unhappy and withdrawn, and Lillian seemed constantly vexed at her great-granddaughter. The results of the psychological profiles reveal that Krista displayed the typical profile of the "poltergeist personality" as described by Bender (1969), Fodor (1958), Rogo (1986) and Roll (1972). These findings support Hypothesis 3, that the poltergeist phenomena reflect the psychological and emotional needs of one or more of the witnesses.

Alvarado (1984) has criticized the psychological interpretation of RSPK because the projective tests often used have low reliability and were usually evaluated by psychologists who knew the identity of the presumed agent and may therefore have biased their interpretation accordingly. This criticism is valid. when most of this work was done, in the 1950's, 60's and 70's, RSPK researchers were groping for an understanding rather than systematically stating and testing hypotheses. After determining that a case included real RSPK occurrences and that these were contingent on the presence of a particular individual, the next question concerned the relation of this person to the people (usually parents or parental figures) whose belongings were being destroyed by RSPK. At that time, projective tests were routinely used by the clinical psychologists who assisted in RSPK research so these tests became tools for exploration. Beyond suggesting

aggressive tendencies, which the RSPK agents shared with numerous others, the psychological work did not reveal a great deal. It was during the search for factors more specific to RSPK agents that we discovered that some of them were subject to episodic eruptions in the central nervous system of the type that may result in symptoms of epilepsy. It seemed possible that the involuntary and recurrent bursts of energy might sometimes be related to the uncontrolled and recurrent manifestation of energy in the surroundings of the agent. But again there are thousands of people with epileptic symptoms who show no sign of RSPK (though there may be some closet RSPK subjects in this group). Fortunately it has been possible to again narrow the search as a result of the discovery of heightened electromagnetic activity at the time of RSPK outbreaks. Other factors that may contribute to RSPK are discussed elsewhere (Roll and Persinger, 1998). Martinez-Taboas (1984), who objected to the theory that RSPK is associated with neuronal discharges on the ground that the theory is too broad, will hopefully agree that we are moving in the right direction.

The Bartons, while outwardly displaying the facade of a happy home, were a deeply troubled family. Much of this disturbance was generated by repressed hostility between Krista and her great-grandmother. I suggested the family seek professional counseling. In December, Mary Barton found a suicide note written by Krista and concealed in her bedroom. Alarmed, she scheduled a psychiatric consultation for the troubled adolescent. Interestingly, there has been no activity in the home since Krista entered psychotherapy in January 1997. It is uncertain if this is the result of therapy or because the phenomena had run their course.

DISCUSSION

As with many poltergeist and haunting disturbances, this case seems to indicate a convergence between physical, psychological, social, and parapsychological factors. In considering the possibility of a deliberate hoax perpetrated by Krista Barton on her family, we attempted to ascertain her physical location during as many of the reported occurrences as possible. In a number of these situations she would have been in a position to create spurious paranormal effects by concealing and throwing small amounts of water or by spitting, but there were also numerous occurrences in which Krista was either under direct observation by other family members, or was known to be in another part of the house at the time of the occurrence.

During our investigative visit, RM was examining a water damaged area of the carpet in the living room when he felt a light sprinkling of moisture on the back of his neck and head. We examined the minute amount of fluid recovered from RM's hair and neck and, although it was of insufficient quantity for chemical analysis, it appeared to be viscous in consistency, more closely resembling saliva than water. During this occurrence, Krista was present in the room standing several feet behind RM and was not under direct observation. Thus, deliberate fraud cannot be ruled out in this instance. But such crude methods could not account for the copious amounts of water produced during previous occurrences, sometimes drenching the entire family (including Krista) simultaneously. Fraud could also not account for the experiences of the three contractors, who were wetted while working in areas which were inaccessible to Krista at the time of the occurrences.

In a number of poltergeist cases, the agent has been caught fabricating phenomena by pounding on walls, throwing objects, etc. In several cases this has occurred *after* occurrences were witnessed by investigators which could not possibly have been normally caused by the agent. Does this behavior on the part of the focus person indicate a desire to mystify, or is it due to the

same psychological factors that seem to underlie the RSPK?

Although deliberate fraud, or acts committed by the agent in a dissociated state cannot be completely ruled out in many poltergeist cases, in the Barton case the corroborative statements from witnesses outside the family and other verification make us relatively sure we were dealing with genuine RSPK.

The source of the water which appeared during the peak of these occurrences was never identified, since the phenomena reached its peak prior to our entering the case, and was largely in abeyance by the time we arrived on the scene. The large amounts of water which had caused the extensive and obvious damage to carpets, furniture, etc. throughout the house had been removed or had dried prior to our arrival. It had not occurred to the family to keep samples for later analysis. We instructed the family to retain a sample of the liquid for analysis if it should appear again, but no further water manifestations occurred after our visit to the site-- although movement of small objects and other more typical poltergeist occurrences continued to be reported for several weeks.

With the exception of the unusual influxes of water (or whatever liquid it actually was), the phenomenology of this case resembles a typical poltergeist outbreak. It is unclear why the water manifestations were exhibited in addition to the more usual sounds and object displacements. In fact the RSPK activity was relatively minor in the Barton home compared to the typical poltergeist case. The water manifestations were by far the most striking feature of this investigation.

Unknown psychological factors may have played a role in the poltergeist's "selection" of water as a weapon. In the case reported by Auerbach (1996), the water manifestations appeared to represent a symbolic retaliation stemming from the adolescent agent's resentment of being forced to participate in swimming activities against his will. Although no specific symbolic correlations were apparent in the Barton case, such metaphorical expressions of unconscious processes-- so often seen in dreams-- cannot be discounted.

Typically, this case seemed to involve no external intelligence but was an expression of the family dynamics within the household. Cases of this kind emphasize the need for researchers to be aware that poltergeists, like hauntings, may take a variety of forms reflecting the inter-personal dynamics, psychopathologies, and emotional needs of the witnesses.

REFERENCES

- Alvarado, C.S. (1984). Psychological evaluations of poltergeist agents: A critique. In R.A. White and R.S. Broughton (Eds.), Research in Parapsychology 1982. Metuchen, N.J.: Scarecrow. pp. 117-118.
- Auerbach, L. (1996). Mind Over Matter, New York. Kensington Books. (pp. 174-184).
- Bayless, R. (1967). The Enigma of the Poltergeist. West Nyack, New York. Parker Publishing. 85-96
- Bender, H. (1969). new developments in poltergeist research: Presidential Address. Proceedings of the Parapsychological Association. 6, 81-102.
- Bender, H. (1974). Modern Poltergeist Research: A plea for an unprejudiced approach. In J. Beloff (Ed.), New Directions in Parapsychology (pp. 122-143). London: Elek.
- Fodor, N. (1958). On the Trail of the Poltergeist. New York. Citadel

- Cox, W.E. (1961). Introductory comparative analysis of some poltergeist cases. Journal of the American Society for Psychical Research, 55, 44-57.
- Fort, C. (1974). The Complete Works of Charles Fort. New York. Dover
- Gauld, A. & Cornell, A.D. (1979). Poltergeists. London: Routledge.
- Gearhart, L. And persinger, M.A. (1986). Geophysical Variables and Behavior: XXXIII. Onsets of historical and contemporary poltergeist episodes occurred with sudden increases in geomagnetic activity. Perceptual and Motor Skills, 62, 463-466.
- Martinez-Taboas, A. (1984). An appraisal of the role of aggression and the central nervous system in RSPK agents. Journal of the American Society for Psychical Research, 78, 56-69.
- Mischo, J. (1968). personality structure of psychokinetic mediums. In Roll, W.G. (Ed). Proceedings of the Parapsychological Association, 5, 35-37.
- Palmer, J. (1974). A Case of RSPK involving a ten year old boy; The Powhatten Poltergeist. Journal of the American Society for Psychical Research, 68, 1-33.
- Persinger, M.A. (1985). Geophysical variables and behavior; XXII. The tectonic strain continuum of unusual events. Perceptual and Motor Skills, 60, 59-65.
- Rogo, D. S. (1986). On the Track of the Poltergeist. Doubleday. New York.
- Roll, W.G. (1972). The Poltergeist, New York, Doubleday.
- Roll, W.G. and Gearhart, L. (1974). Geomagnetic perturbations and RSPK. In W.G. Roll, R.L. Morris, and J.D. Morris (Eds) Research in Parapsychology 1973, pp.44-46.
- Roll, W.G., Maher, M.C. and Brown, B. (1991). An Investigation of Reported Haunting Occurrences in a Japanese Restaurant in Georgia. Proceedings of Presented Papers, The Parapsychological Association 35th Annual Convention, 151-168
- Roll, W.G. & Persinger, M.A. (1998). Poltergeist and nonlocality: Energetic aspects of RSPK. Unpublished manuscript.
- Thurston, H. (1953). Ghosts and Poltergeists. London. Burns Oates.
- The Annals of Psychic Science. (April, 1906.)

CORRELATES OF ESP MAGNITUDE AND DIRECTION IN THE FRNM MANUAL GANZFELD DATABASE

John Palmer
Institute for Parapsychology

ABSTRACT

Retrospective analyses were performed on data from the first sessions of 206 subjects who participated in manual ESP ganzfeld experiments designed and supervised by Dr. H. Kanthamani at the Foundation for Research on the Nature Man from 1986 through 1992. Sample sizes varied for particular analyses due to missing data. ESP scores were derived from both percipient and experimenter ratings. The percentage of direct hits (MCE = 25%) was 27.2% for percipient ratings and 36.4% for experimenter ratings. The latter percentage is highly significant, $p = 5.2 \times 10^{-4}$. The magnitude (variance) of ESP scores based on percipient ratings was significantly influenced by the identity of the experimenter assigned to supervise the agent (sender) when subjects brought their own agent to the session. The magnitude and direction (hitting or missing) of ESP scores were compared to scores on the Myers-Briggs Type Inventory (MBTI) and a Personal Information Form (PIF) that asked about various personal experiences and attitudes relevant to psi. The S-N, T-F, and J-P scales of the MBTI were combined to form a single scale (NFP) interpreted as a measure of spontaneity. Separate analyses were performed on ESP ranks and z -scores (unsquared and squared). As predicted from theory and previous research, there was a significant positive relationship between NFP and ESP magnitude, but only with experimenter ratings. With percipient ratings, there was a significant positive correlation between NFP and ESP direction, which was blocked when one particular experimenter supervised the percipients. A prediction that MBTI extraversion should correlate positively with ESP direction for females was not confirmed for either percipient or experimenter ratings. However, relationships between extraversion and both the magnitude and direction of ESP scores proved to differ significantly as a function of whether or not a particular experimenter was involved in supervising the participants. A previously reported finding that the expected positive relationship between ESP and extraversion among FP subjects on the MBTI reversed when the agents were lab staff members was not confirmed. On the PIF, percipient ratings yielded a negative relationship between subjects' level of education and ESP direction. More extreme ESP scores were provided by females, and by subjects who reported spontaneous PK experiences, a large number of coincidences in their lives, and an internal locus of control. Only the education and coincidence effects were replicated with experimenter ratings, which also yielded a significant tendency for left-handed and ambidextrous subjects to obtain more extreme ESP scores than right-handed subjects. It was suggested that percipient ratings are likely confounded by differences in subjects' skills at judging and that experimenter ratings may provide a purer measure of psi.

INTRODUCTION

At last year's PA Convention, I reported retrospective analyses of psychological correlates of ESP scores of "first-timers" in large samples of autoganzfeld data from the Psychophysical Research Laboratories (PRL) and the Rhine Research Center (RRC) (Palmer, 1997). In the PRL database, subjects completed the Myers-Briggs Type Inventory (MBTI) (Myers & McCaulley, 1985), whereas in the RRC database subjects completed a comparable but more psychometrically sophisticated global measure of personality called the NEO-PI (Costa & McCrae, 1985). In both studies, subjects completed homemade Participant Information Forms (PIFs) that asked about

their experiences of and attitudes toward psi and related mental phenomena.

An important element of the above analyses is that ESP scores were defined with respect both to their direction (hitting or missing) and their magnitude (variance). A model called the M-D model was introduced which hypothesized that the magnitude of ESP scoring should be positively associated with the spontaneity of cognitive processes and the direction of ESP scoring should be positively associated with extraversion.

Results from both databases provided qualified support for the model. In the PRL database, extraverts on the MBTI scored significantly more positively on the ESP task than did introverts; this finding was found to be attributable entirely to females. In the RRC database, there was no significant relationship between ESP and NEO-PI extraversion overall, but the relationship was significantly more positive for females than for males. In the PRL database, composite scores on the S-N, T-F, and J-P scales were treated as a measure of spontaneity, based on face-validity of the scales and Q-Sort data published in the MBTI manual (Myers & McCaulley, 1985). Extreme scorers on the ESP task were significantly more likely to fit the NFP profile than were mid-range scorers. For the RRC database, a similar "spontaneity" scale, computed from NEO-PI scales that were correlated with the respective MBTI scales, provided results in the same direction as did NFP in the PRL database, but the results were not significant. Neither database provided much of interest with regard to the PIFs.

In the present report, these analyses are extended to data from a series of experiments designed and supervised by Dr. H. Kanthamani at the Foundation for Research on the Nature of Man (FRNM) using a manual ganzfeld procedure (Kanthamani & Broughton, 1994). (FRNM is the previous name of the RRC). The database consists of data from 206 subjects who were tested from 1986 through 1992, and it is restricted to the first ESP ganzfeld session completed by each subject. The number of subjects to be reported in these analyses will not always coincide with those reported by Kanthamani and Broughton (1994) for two reasons. First, these authors restricted their analyses concerning MBTI and PIF correlates of ESP to subjects "who had never been in ganzfeld before and participated for the first time in our experiments" (p. 187). To get a larger sample of appropriate subjects, I chose to analyze data from subjects irrespective of prior testing history. This was also the policy adopted in the analyses of the autoganzfeld databases (Palmer, 1997). Second, MBTI and PIF data from several subjects were not available at the time I did my analyses.

The procedure used in these experiments is described fully by Kanthamani and Broughton (1994), and these details will not be reproduced in full here. However, a few key points will be mentioned. The targets consisted of paper pictures, such as prints of paintings. Most trials involved a sender (agent), but 46 were conducted in clairvoyance mode. In 61 sessions subjects brought along their own agents; in the other GESP sessions agents were FRNM staff or persons assigned by them. Separate experimenters were assigned to the agent (AE) and percipient (PE). When staff members served as agents, they doubled as AEs. During judging, percipients blind-rated a duplicate of the target plus three decoys on a scale of 0 to 99 for degree of correspondence to their mentation. In 173 of the 206 sessions, PE, who had recorded the percipient's mentation, blind-rated the pictures at the same time as the percipient. With a few exceptions, those subjects who completed the MBTI and PIF did so before their first ESP ganzfeld session.

The two principal findings from the retrospective analyses of the RRC and PRL databases, a positive relationship between spontaneity and ESP magnitude and a positive relationship between

extraversion and ESP hitting, are treated as hypotheses for this study, whether based on percipient ratings or experimenter ratings. One-tailed statistical tests are used for the specific tests of these relationships. All the other analyses, although mostly planned in advance, can best be considered exploratory and use two-tailed tests. Because the total number of analyses is fairly large, it is fair to say that none of the significant effects below could withstand Bonferroni corrections based on this total number. In my view, in the final analysis their validity rests in how consistently they are confirmed in other databases, which is true whether they pass the Bonferroni test or not.¹ To help guide the reader through the maze to follow, I will occasionally insert an [m] or [d] in brackets to indicate whether a particular analysis refers to a magnitude or direction effect.

OVERALL ESP SCORES

Percipient Ratings

It has become the norm in analyses of ganzfeld data to treat the proportion of direct hits as the primary measure of overall ESP and that convention will be followed here. For the 206 subjects, there were 56 direct hits (27.2%). This value does not differ significantly from the MCE of 25%, $z = +0.55$, $p = .582$.

I also analyzed the ESP scores with respect to magnitude and direction, following the same procedures used in the analyses of the other databases (Palmer, 1997). Direction effects are defined as (a) ranks 1-2 (hits) versus 3-4 (misses) and (b) unsquared z -scores. Magnitude effects are defined as (c) ranks 1-4 (extreme) versus 2-3 (middle) and (d) squared z -scores. Using ranks, neither the magnitude effect, $z = 1.12$, $p = .265$, nor the direction effect, $z = .138$, $p = .889$, are significant. The mean z -score of $-.012$ is also nonsignificant, $t(205) = 0.19$, $p = .720$. The correlation between the unsquared and squared z -scores is negligible, $r_s(204) = .026$.

Experimenter Effects. Because I was informed by a colleague that there might be nontrivial experimenter effects in the database, I decided to analyze for them. The roles of AE and PE were most frequently filled by two female FRNM staff members who will be referred to in this report as E1 and E2. A number of other FRNM staff filled these roles at various times, but none of these did so in more than ten sessions. For this reason, it was decided to lump them all together in a single category called "Other". Thus, for both AE and PE there were three categories: E1, E2, and Other.

A chi-square test comparing direct hits for the three AE categories approached significance, $\chi^2(2) = 5.78$, $p = .056$. This is attributable to a direct hit rate of 33.8% ($N = 68$) when E2 was AE, $z = +1.68$, $p = .093$, balanced by a direct hit rate of 14.3% ($N = 49$) when someone besides E1 and E2 served in this role, $z = -1.72$, $p = .085$. The corresponding chi-square for PEs was not significant, $\chi^2(2) = 2.74$, $p = .255$. The almost significant difference among the three AE categories was not confirmed for unsquared z -scores, $F(2,159) = 0.48$, $p = .548$. There also was no effect for the PE categories, $F(2,203) = 0.57$, $p = .565$.

The AE analyses of ESP scores were repeated for binary hits and for extreme ranks, to see if the

¹ As the reader may sense, I am not a big fan of Bonferroni corrections. It strikes me as extremely counter-intuitive (if not downright absurd) to suggest that the reality of a given effect (i.e., whether it has a non-zero population value) depends on how many other effects an analyst happens to look for. I believe our conservative instincts are more rationally served by an old-fashioned appeal to replication.

almost significant effect above is best interpreted as a direction effect or a magnitude effect. As the magnitude analysis was significant, $\chi^2(2) = 6.44$, $p = .040$, and the direction analysis nonsignificant, $\chi^2(2) = 1.43$, $p = .489$, we can conclude that AEs may have affected the magnitude of scoring.

It was decided to repeat the magnitude analyses on the subsample of percipients who brought their own agents, as this is the group for which the AE role seems most relevant. In line with this reasoning, the effect is significant for this subsample, $\chi^2(2) = 8.76$, $p = .013$, but not for percipients tested with assigned agents, $\chi^2(2) = 1.74$, $p = .419$. The effect for percipients who brought their own agents is confirmed by the analysis of transformed squared z -scores², $F(2,58) = 5.58$, $p = .006$, and is once again negligible for those with assigned agents, $F(2,82) = 0.16$, $p = .849$.

Experimenter Ratings

For experimenter ratings, there were 63 direct hits in 173 trials (36.4%), which is highly significant, $z = 3.47$, $p = 5.2 \times 10^{-4}$. (When the analysis is applied to first timers as defined by Honorton (1997) and by Kanthamani and Broughton (1994), there are 57 hits in 155 trials (36.8%), $z = 3.39$, $p = 7.0 \times 10^{-4}$.)

The magnitude effect proved to be significant by the ranks analysis, with 61.3% of the percipients obtaining a rank of either 1 or 4, $z = 2.95$, $p = .003$. The direction effect, however, is not quite significant, with 57.2% of the percipients getting binary hits, $z = 1.91$, $p = .056$. The mean unsquared z -score is $+1.108$, which does not reach significance, $t(172) = 1.54$, $p = .125$, demonstrating once again that direct hits are the most sensitive index of overall psi-hitting, even though in principle a continuous measure like z -scores should be more powerful. The correlation between the squared and unsquared z -scores is once again negligible, $r_s(171) = -.036$. The correlation between the unsquared z -scores based on percipient and experimenter ratings is considerable, $r_s(171) = .655$; the corresponding correlation for squared z -scores is smaller, $r_s(171) = .341$.

Experimenter Effects. The same analyses performed for percipient ratings were repeated for experimenter ratings. There is no significant evidence of influence by either AE or PE on direct hits, or the direction or magnitude of ESP scores, whether for the entire sample or for percipients who brought their own agents.

DISCUSSION

The most noteworthy finding from these analyses is a 36.4% direct hit rate derived from experimenter ratings. This is much better than the 29.6% hit rate reported by Honorton (1997) for 226 first-timers (percipient ratings) in the PRL database, the 26.5% hit rate from 155 first-timers (percipient ratings) in the RRC database (Broughton & Alexander, 1997), and the 27.2% hit rate (percipient ratings) from the FRNM database. Unfortunately, experimenter judging was not undertaken for the two autoganzfeld databases. The result from experimenter ratings in the FRNM database cannot be attributed to the fact that experimenter ratings were not available from a number of subjects. When sessions with no experimenter ratings are dropped from the analysis

² To eliminate skewness of the squared z -scores for parametric analyses, they were transformed by taking the square roots of the absolute values of the unsquared z -scores.

Correlates of ESP magnitude & direction

of percipient-based data, the direct hit rate increases only to 28.9%.

The identity of AE made a difference in the magnitude of ESP scores for percipients who brought their own agents, but only for percipient ratings. The fact that the finding holds only when percipients brought their own agents suggests that how AEs interacted with the agents socially affected the pairs' ESP scores, with E2 being particularly effective in producing hits and staff members other than E2 and E1 being particularly ineffective. The latter is not surprising in that the other staff were in effect “filling in” for E2 and E1; they may not have felt the experiment was really theirs or were not entirely comfortable with their roles through lack of practice.

MBTI “SPONTANEITY”

It is customary with the MBTI to place subjects in categories according to whether they score above or below the theoretical midpoint of its four component scales: Extraversion-Introversion (E-I), Sensing-Intuition (S-N), Thinking-Feeling (T-F), and Judgment-Perception (J-P). There are 16 possible categories, each identified by a sequence of four letters signifying the poles of the four scales which the subject's scores are closest to. For example, ENFJ labels someone who scores in the extraversion direction on E-I, the intuition direction on S-N, the feeling direction on T-F, and the judgment direction on J-P. Each of the scales can also be scored continuously.

In the PRL database, there was a clear tendency for the S-N, T-F and J-P scales to correlate highly among themselves (+.304 to +.429) but not at all with the E-I scale. As illustrated in Table 1, this trend was not nearly as pronounced in the FRNM database. It still was identifiable, however, so I chose once again to pool S-N, T-F, and J-P to form a single “spontaneity” scale so as to make the following analyses directly comparable to those applied to the PRL database (Palmer, 1997). The composite scale will hereafter be called NFP, indicating the high-scoring pole on each subscale.

	E-I	S-N	T-F
S-N	-.077		
T-F	-.195*	.179*	
J-P	-.105	.413**	.219**

* $p < .05$; ** $p < .01$

Table 1: Pearson correlation matrix for the Myers-Briggs Type Inventory in the FRNM Database (N = 175)

Percipient Ratings

The M-D model predicts that ESP ranks should be significantly more extreme for subjects classified as NFP (either ENFP or INFP) than for the other subjects. It is clear from the left side of Table 2 that such is not the case, although the weak trend is in the predicted direction, $\chi^2(1) = 0.26$, $p = .305$, 1-tailed. The right side of Table 2 illustrates that there is no direction effect, $\chi^2(1) = 0.47$, $p = .829$.

For the z -scores, however, a different pattern emerges. Although the correlation between NFP and the squared z -scores [m] is negligible $r_s(173) = +.020$, $p = .398$, 1-tailed, the correlation between NFP and the unsquared z -scores [d] is significantly positive, $r_s(173) = +.203$, $p = .007$. In other words, ESP hitters scored significantly higher on NFP than did ESP missers. The

correlation with ESP was positive for all the component MBTI scales but significant only for S-N, $r_s(173) = +.244, p = .001$.

Experimenter Effects. To assess possible experimenter mediation of the NFP-ESP relationship, I conducted linear regression analyses with z -scores (transformed squared and unsquared respectively) as the criterion variable, and NFP plus either AE or PE as the predictor variables. If there was a significant NFP by experimenter interaction, the interaction was broken down to determine its precise nature.

Frequencies (Magnitude)				Frequencies (Direction)			
	NFP	OTHER	Total		NFP	OTHER	Total
EXTREME	41	56	97	HITS	36	51	87
MIDDLE	30	48	78	MISSES	35	53	88
Total	71	104	175	Total	71	104	175
Column Percents				Column Percents			
	NFP	OTHER	Total		NFP	OTHER	Total
EXTREME	57.7	53.8	55.4	HITS	50.7	49.0	49.7
MIDDLE	42.3	46.2	44.6	MISSES	49.3	51.0	50.3
Total	100.0	100.0	100.0	Total	100.0	100.0	100.0

Table 2: ESP Scores Classified by Magnitude and Direction as a Function of MBTI NFP Categories, Percipient Rankings

An almost significant PE by NFP interaction was uncovered for the unsquared z -scores [d], $F(2,169) = 2.99, p = .053$, and this was considered close enough to continue the analysis. Further dissection revealed that the significant positive NFP-ESP correlation described above is contributed entirely by subjects for whom E1 was *not* the PE, $r_s(76) = +.439, p = 5.2 \times 10^{-5}$; for E1's subjects, $r_s(95) = -.011$. I next decided to break down the corresponding ranks analysis (which had been negligible) to look for experimenter differences there. As illustrated in Table 3, there is a cancellation, with NFP subjects tending to score below chance if tested by E1 and above chance if tested by someone else.

Experimenter Ratings

For the ranks analysis, we find in the left side of Table 4 a somewhat stronger trend in support of the M-D model than we found with percipient ratings. There were more NFP subjects among the extreme than among the middle ESP scorers to an almost significant degree, $\chi^2(1) = 2.39, p = .061, 1$ -tailed. Conversely, the right side of Table 4 reveals almost no direction trend, $\chi^2(1) =$

Correlates of ESP magnitude & direction

0.04, $p = .852$.

These trends are supported by the z -score analyses. The correlation between NFP and squared z -scores [m] is significant, $r_s(147) = +.158$, $p = .027$, 1-tailed, whereas (in contrast to what was found with percipient ratings) the correlation between NFP and unsquared z -scores [d] is negligible, $r_s(147) = +.038$, $p = .646$. For the squared z -scores, the correlation was attributable primarily to S-N, $r_s(147) = +.146$, $p = .038$, 1-tailed, and J-P, $r_s(147) = +.164$, $p = .023$, 1-tailed.

PE = E1				PE = Not E1			
Frequencies (Direction)				Frequencies (Direction)			
	NFP	OTHER	Total		NFP	OTHER	Total
HITS	17	33	50	HITS	19	18	37
MISSES	19	28	47	MISSES	16	25	41
Total	36	61	97	Total	35	43	78
Column Percents				Column Percents			
	NFP	OTHER	Total		NFP	OTHER	Total
HITS	47.2	54.1	51.5	HITS	54.3	41.9	47.4
MISSES	52.8	45.9	48.5	MISSES	45.7	58.1	52.6
Total	100.0	100.0	100.0	Total	100.0	100.0	100.0

Table 3: ESP Direction Scores as a Function of MBTI NFP Categories, E1 vs. Others as PE, Percipient Rankings

Frequencies (Magnitude)				Frequencies (Direction)			
	NFP	OTHER	Total		NFP	OTHER	Total
EXTREME	42	48	90	HITS	38	52	90
MIDDLE	20	39	59	MISSES	24	35	59
Total	62	87	149	Total	62	87	149
Column Percents				Column Percents			

	NFP	OTHER	Total		NFP	OTHER	Total
EXTREME	67.7	55.2	60.4	HITS	61.3	59.8	60.4
MIDDLE	32.3	44.8	39.6	MISSES	38.7	40.2	39.6
Total	100.0	100.0	100.0	Total	100.0	100.0	100.0

Table 4: ESP Scores Classified by Magnitude and Direction as a Function of MBTI NFP Categories, Experimenter Rankings

Experimenter Effects. None of the linear regression analyses yielded significant interaction effects involving the experimenters. Nonetheless, I decided to check directly if the PE effect found with unsquared z -scores [d] for percipient ratings would also show up with squared z -scores [m] for experimenter ratings. Sure enough, the positive correlation between NFP and squared z -scores is significant for percipients not supervised by E1, $r_s(61) = +.283$, $p = .012$, 1-tailed, but trivial for subjects who were supervised by E1, $r_s(86) = +.052$, $p = .317$, 1-tailed. However, this pattern reversed slightly for the ranks analysis.

DISCUSSION

The most noteworthy finding to emerge from the NFP analyses is the dramatic difference in the nature of the NFP-ESP relationship depending on whether the percipient or PE does the judging. With percipient ratings, we find a direction effect — higher ESP scores among high NFP subjects. With experimenter ratings we find a magnitude effect — more extreme ESP scores among high NFP subjects. This difference cannot be attributed to sample differences between the two analyses. When subjects for whom there were no experimenter ratings are removed from the analyses of percipient ratings, the correlation between NFP and unsquared z -scores [d] remains significant, $r_s(147) = +.192$, $p = .019$, and the correlation between NFP and squared z -scores [m] remains trivial, $r_s(147) = +.077$, $p = .175$, 1-tailed, although the two correlations do come slightly closer together.

It is puzzling that it was experimenter ratings which replicated the PRL finding, as the latter was based on percipient ratings. Or was it? Honorton et al. (1996) reported that for about half of the first-timers, the experimenter “point[ed] out potential correspondences [the percipient] may have overlooked” (p. 109). Such interventions could have had a major impact on how subjects made their ratings. The relative success rates of those first-timers who did and did not receive prompting by the experimenter cannot be determined precisely from the data presented in Honorton’s report, but it would appear that it was about 5% higher for those who did receive prompting. Among all trials (novices plus experienced subjects), Wiseman, Smith, and Kornbrot (1996) report that prompted subjects scored significantly higher than non-prompted ones.

For both types of ratings, although more clearly with percipient ratings, the respective NFP-ESP effects were blocked when E1 was PE. Thus, E1’s involvement in this crucial role somehow kept the characteristics of the subject that NFP measures (presumably spontaneity) from influencing their ESP scores.

Correlates of ESP magnitude & direction

The M-D model predicts that extraverts should score more positively on the ESP task than introverts, and the PRL data indicate that this relationship is attributable to females (Palmer, 1997). As shown in the right side of Table 5, although the trend in the FRNM database is in the predicted direction, it is nowhere near significant, $\chi^2(1) = 0.05, p = .413$, 1-tailed. Moreover, the trend is in the wrong direction for females, $\chi^2(1) = 0.22, p = .679$, 1-tailed. As illustrated in the left side of Table 5, even less of a relationship was found between extraversion and ESP magnitude, $\chi^2(1) = 0.01, p = .920$.

These null findings are confirmed by the z -score analyses. There is a negligible correlation between E-I and the unsquared ESP z -scores [d], $r_s(173) = +.017, p = .413$, 1-tailed³, with a reversal for females, $r_s(95) = -.070, p = .751$, 1-tailed. The correlation between E-I and squared z -scores [m] is negligible, $r_s(173) = -.007, p = .928$.

Frequencies (Magnitude)				Frequencies (Direction)			
	EXT	INT	Total		EXT	INT	Total
EXTREME	43	54	97	HITS	39	48	87
MIDDLE	34	44	78	MISSES	38	50	88
Total	77	98	175	Total	77	98	175
Column Percents				Column Percents			
	EXT	INT	Total		EXT	INT	Total
EXTREME	55.8	55.1	55.4	HITS	50.6	49.0	49.7
MIDDLE	44.2	44.9	44.6	MISSES	49.4	51.0	50.3
Total	100.0	100.0	100.0	Total	100.0	100.0	100.0

Table 5: ESP Scores Classified by Direction and Magnitude as a Function of MBTI E-I Categories, Percipient Rankings

In an interim report on the PRL database, Honorton and Schechter (1986) stated that among FP's on the MBTI, there were more direct hits among extraverts (EFPs) when percipients brought their own agents and more direct hits among introverts (IFPs) when lab agents were used, the difference being significant. No such trend was found in the FRNM data. Ignoring clairvoyance sessions, when percipients brought their own sender there were 6 hits for EFPs and 5 hits for IFPs; when lab agents were used, there were 5 hits for EFPs and 3 hits for IFPs.

Experimenter Effects. The linear regression analyses for unsquared z -scores produced significant

³ The MBTI E-I Scale is scored in the introversion direction, which can be rather confusing. To make the situation clearer, I have reversed the direction of all reported correlations involving E-I. Thus, a positive correlation means higher ESP scores by extraverts.

interactions between E-I and both AE, $F(2,130) = 3.63$, $p = .029$ and PE, $F(2,169) = 3.47$, $p = .033$. Further dissection revealed that it made a difference whether E2 was involved in the session as either AE or PE. When she was so involved, extraverts scored more positively than introverts, $r_s(107) = +.231$, $p = .008$, 1-tailed. When she was not involved, the introverts scored higher, $r_s(64) = -.281$, $p = .989$, 1-tailed. The difference between these two values is highly significant, $z = 3.29$, $p = .001$. Comparable results are obtained with the ranks analysis, although the significance levels are not as extreme. There are more hitters than missers among subjects classified as extraverts on the MBTI when E2 was involved, $\chi^2(1) = 3.58$, $p = .029$, 1-tailed, whereas the reverse is true when she was not involved, $\chi^2(1) = 2.36$, $p = .938$, 1-tailed.

Because there was an experimenter effect involving AE, I decided to see if this effect is attributable to subjects who brought their own agents. When the linear regression analyses for unsquared z -scores [d] was repeated for this subsample, there once again was a significant AE by E-I interaction, $F(2, 44) = 4.72$, $p = .014$. Further inspection of the data revealed that when E2 was AE, extraverts scored more positively than introverts, $r_s(22) = +.322$, $p = .113$, 1-tailed, but when she was not AE, the introverts scored higher, $r_s(24) = -.292$, $p = .949$, 1-tailed. The two correlations are significantly different, $z = 2.14$, $p = .032$. The effect is much less pronounced when percipients with assigned agents are analyzed, although still in the same direction. When E2 was AE, the correlation between E-I and the unsquared z -scores [d] is $r_s(30) = +.210$, $p = .124$, 1-tailed; when she was not AE, $r_s(39) = -.072$, $p = .328$, 1-tailed. This time, the correlations are not significantly different, $z = 1.16$, $p = .246$. Thus, only when subjects brought their own agents did the identity of AE make a significant difference.

When ranks are substituted for z -scores the same trends emerge, although somewhat more weakly. For percipients who brought their own agents, extraverts scored higher when E2 was AE, exact $p = .052$, 1-tailed, whereas introverts scored higher when she was not AE, exact $p = .659$, 1-tailed. For percipients with assigned agents, the trend for extraverts to score higher than introverts is weakened, $\chi^2(1) = 0.13$, $p = .360$, 1-tailed, and there is virtually no trend when E2 was not AE, $\chi^2(1) = 0.01$, $p = .469$, 1-tailed.

Neither regression analysis approached significance when the transformed squared z -scores were the criterion variable.

Experimenter Ratings

As shown in the right side of Table 6, the extraversion-ESP results from the ranks analysis are more supportive of the M-D model for experimenter ratings than for percipient ratings, just as we found with NFP. There are more hitters among extraverts than introverts to a suggestive degree, $\chi^2(1) = 1.79$, $p = .091$, 1-tailed. However, in contrast to the PRL data but consistent with the percipient ratings in the FRNM data, the effect is stronger for males, $\chi^2(1) = 2.79$, $p = .095$, than for females, $\chi^2(1) = 0.08$, $p = .391$ 1-tailed, although at least there was no reversal for females. As shown in the left side of Table 6, there is no magnitude effect for extraversion, $\chi^2(1) = 0.44$, $p = .509$.

The predicted effect is weaker for the z -score analysis than for the ranks analysis. There is only a slight positive correlation between E-I and the unsquared z -scores [d], $r_s(147) = +.043$, $p = .302$, 1-tailed. The correlation is once again stronger for males, $r_s(63) = +.199$, $p = .112$, and it reverses for females, $r_s(82) = -.081$, $p = .770$, 1-tailed. The overall correlation between E-I and the squared z -scores is trivial, $r_s(147) = +.038$, $p = .646$.

Correlates of ESP magnitude & direction

Experimenter Effects. The regression analyses yielded a significant E-I by AE interaction for the unsquared z -scores, $F(2,104) = 4.25$, $p = .017$, and the same interaction is almost significant for the squared z -scores, $F(2,104) = 2.98$, $p = .055$. Further inspection revealed that whether or not E2 was AE is the crucial factor. For sessions when E2 was AE, the expected positive relationship between E-I and ESP is significant, $r_s(47) = +.253$, $p = .040$, 1-tailed, whereas when she was not the AE we find a slight reversal, $r_s(59) = -.098$, $p = .550$, 1-tailed. However, these correlations are not significantly different, $z = 1.81$, $p = .070$. Likewise, when E2 was AE, extraverts obtained more extreme squared z -scores [m] than introverts, $r_s(47) = +.294$, $p = .04$, and this reversed when she was not the AE, $r_s(59) = -.150$, $p = .247$. These correlations do differ significantly, $z = 2.30$, $p = .021$. This magnitude effect was not predicted by the M-D model.

Frequencies (Magnitude)				Frequencies (Direction)			
	EXT	INT	Total		EXT	INT	Total
EXTREME	40	50	90	HITS	42	48	90
MIDDLE	23	36	59	MISSES	21	38	59
Total	63	86	149	Total	63	86	149
Column Percents				Column Percents			
	EXT	INT	Total		EXT	INT	Total
EXTREME	63.5	58.1	60.4	HITS	66.7	55.8	60.4
MIDDLE	36.5	41.9	39.6	MISSES	33.3	44.2	39.6
Total	100.0	100.0	100.0	Total	100.0	100.0	100.0

Table 6: ESP Scores Classified by Direction and Magnitude as a Function of MBTI E-I Categories, Experimenter Rankings

Using ESP ranks, we find that when E2 was AE, the expected tendency for extraverts to obtain more hits than introverts is significantly confirmed, $\chi^2(1) = 5.12$, $p = .012$, 1-tailed. The trend is considerably weaker for other AEs, $\chi^2(1) = 0.18$, $p = .334$, 1-tailed. As for the magnitude effect, extraverts are somewhat more likely to have extreme ESP scores than introverts when E2 is the AE, $\chi^2(1) = 1.64$, $p = .200$, whereas introverts score more extremely when someone else is the AE, $\chi^2(1) = 2.86$, $p = .091$.

When the regression analyses were restricted to percipients who brought their own agents, the interactions between E-I and AE are not significant for either unsquared or squared z -scores. Further inspection of the data confirmed that none of the results reported above differ as a function of whether percipients brought their own agents.

DISCUSSION

The expected tendency for extraverts to score more positively than introverts on the ESP task was not confirmed to a statistically significant degree by either percipient or experimenter ratings. Moreover, in contrast to the PRL database, the predicted trend is stronger for males than for females, and even reverses for females by some analyses.

A possible clue to help explain these discrepant findings comes from the fact that the nature of the relationship between extraversion and ESP seems to have been influenced in the FRNM experiments by who the experimenters were. For both percipient and experimenter ratings, the predicted positive relationship between extraversion and ESP is found only when E2 served as an experimenter. This is the case with percipient ratings whether she was AE or PE, but for experimenter ratings only her role as AE is relevant. In the latter case, there was an unexpected magnitude effect; more extreme ESP scores among extraverts when E2 served as AE.

For percipient ratings, we find once again that E2's role as AE only makes a difference for percipients who bring their own agents. For experimenter ratings, the effect holds about equally for all percipients.

A secondary finding of possible importance is that subjects tested in sessions in which E2 was either AE or PE described themselves as much more extraverted on the MBTI, $M = 100.3$, than those tested in sessions when she was not, $M = 111.8$; $t(173) = 2.86$, $p = .005$. It is unlikely that this strange occurrence is due to any experimenter influence on how subjects filled out the MBTI, because the great majority did so before the ESP session. The proportion of sessions in which E2 was involved as AE or PE differed markedly from series to series, so perhaps extraverted subjects simply volunteered more frequently for series in which E2 was a frequent experimenter. This also seems to be an unlikely explanation, because those subjects not tested by E2 were little more extraverted if they volunteered for series in which E2 was involved in more than half the sessions, $M = 112.3$, than they were if E2 was involved in less than half the sessions, $M = 111.6$.

Status of the M-D Model

At this point, it might be helpful to summarize how the M-D model has fared in the face of results from the FRNM database. In a nutshell, the model predicts that spontaneity (indicated by NFP on the MBTI) should be positively related to the magnitude of ESP scores in the ganzfeld. In contrast, extraversion should be positively related to the direction of ESP scores, especially for females.

When scores are based on percipient ratings (those used in the previously analyzed PRL and RRC autoganzfeld databases), support for the model was not found. Although for both spontaneity and extraversion the observed trends were in the predicted direction, they were nowhere near significant. Moreover, the predicted extraversion effect (such as it was) is attributable to males and actually reversed for females. Finally, using z -scores, a significant positive relationship was found between spontaneity and ESP *direction*, which was contributed entirely by subjects for whom E1 was not PE. This effect was *not* predicted by the model.

The model fares somewhat better when ESP scores are based on experimenter ratings. The predicted positive relationship between spontaneity and ESP magnitude approaches significance with the ranks analysis and reaches significance with the z -score analysis. The predicted trend for extraverts to score more positively on the ESP task than introverts was stronger than it was when

Correlates of ESP magnitude & direction

perceptant ESP ratings were used, but it was not significant by either statistical test, and once again the effect was solely due to males. However, the overall extraversion - ESP relationship was found to be significantly influenced by the identity of the AE. Further breakdown of the data revealed that when E2 served as AE, the predicted relationship was significant by both the rank and z -score measures.

In conclusion, the model received overall support only if we assume that the experimenter ratings in the FRNM database are comparable to the perceptant ratings in the PRL and RRC databases. Although there may be some basis for this assumption (as noted above), it is nonetheless rather tenuous. Thus, support for the M-D model from the present analyses may best be described as ambiguous (except for the restriction of the extraversion - ESP relation to females, which was clearly undermined).

Personal Information Form (PIF)

Analysis of PIF data from the PRL and RRC databases yielded only one effect of borderline significance: a positive correlation between self-rated competitiveness and squared z -scores in the RRC data. Twenty-two items from the FRNM PIF were evaluated: sex, date of birth (age), handedness, education (number of years completed), belief in psi, PK experiences, experiences of coincidences, use of medications, adequacy of sleep, practice of a mental discipline (e.g., meditation), practice of a physical regimen (e.g., yoga), participation in a self-improvement program (e.g., psychotherapy), dream recall, lucid dreams, liking of fantasy, losing oneself in random thoughts, belief in luck, personal luckiness, willingness to take risks, internal locus of control, outgoingness, and competitiveness.⁴

In keeping with the procedure used with the other databases (Palmer, 1997), I first performed one way analyses of variance with PIF items as the dependent variable and ESP ranks (1 vs. 2-3 vs. 4) as the independent variable. If the resulting F gave $p < .10$, further analyses were undertaken to determine if there was a significant ($p < .05$) direction or magnitude effect. These latter analyses generally consisted of correlations between ESP z -scores and PIF items on the one hand, and Mann-Whitney U -tests between ESP ranks (hits/misses or extreme/middle) and PIF items on the other. When a PIF item was scored dichotomously, chi-square tests were substituted for the U -tests.

Perceptant Ratings

The only significant direction effect was a negative relationship between ESP and level of education, but the significance was obtained in only one of the analyses, $r_s(174) = -.171$, $p = .023$; $U = 3418$, $p = .174$.

Several PIF items yielded magnitude effects that were significant by both analyses. Extreme ESP scorers were more likely than mid-range scorers to: (a) be female, $r_s(187) = .161$, $p = .027$; $\chi^2(1) = 4.07$, $p = .044$; (b) report having had a spontaneous PK experience, $r_s(183) = +.155$, $p = .035$; $\chi^2(1) = 4.92$, $p = .027$; (c) report a high frequency of coincidences, $r_s(176) = +.163$, $p = .030$; $U = 4787$, $p = .011$; and (d) have an internal locus of control, $r_s(172) = .270$, $p = 3.1 \times 10^{-4}$; $U = 4757$, $p = .002$. Correlational results for both perceptant and experimenter ratings are presented more completely in Table 7.

⁴ A few items, most notably frequency of ESP experiences, were not analyzed because the questions differed on different versions of the PIF and these could not be reconciled.

PIF Item	P-Judging		E-Judging	
	<i>z</i>	<i>z</i> ²	<i>z</i>	<i>z</i> ²
Education	-.171	-.033	-.171*	.021
Sex (F>M)	-.072	.161*	-.015	.053
PK Experience	.052	.155*	.006	.063
Coincidences	-.028	.163*	-.068	.176
Internal Control	-.077	.270*	.049	.103
Handedness (R>L/A)	.056	-.140	.075	-.194*

Significant values in boldface.

*Significant also by analysis of ESP ranks

Table 7: Spearman Correlations Between ESP *z* and *z*²-Scores and Significant (*p* < .05) PIF Items

Experimenter Ratings

Only two of the above relationships held up with experimenter ratings. Once again, there was a negative relationship between level of education and the direction of ESP scoring, $r_s(149) = -.171$, $p = .036$; $U = 2137.5$, $p = .023$. There was also a positive relationship between frequency of reported coincidences and ESP magnitude, $r_s(152) = +.176$, $p = .029$; $U = 3382$, $p = .053$.

The only new relationships involve handedness. Subjects had been classified in the database as either right-handed, left-handed, or ambidextrous. As illustrated in the left side of Table 8, there was a tendency for left-handed and ambidextrous subjects to obtain more extreme ESP scores than right-handed subjects, $r_s(148) = .194$, $p = .017$; $c^2(1) = 7.20$, $p = .007$. Discounting right-handed subjects, left-handed subjects tended to score above chance and ambidextrous subjects below chance, but with such a small *N* this trend was significant by only one of the analyses, $r_s(23) = .306$, $p = .137$; Fisher exact $p = .049$. This is illustrated in the right side of Table 8.

DISCUSSION

The most solid of the above findings are probably those involving level of education and coincidences, because they were significant using both percipient and experimenter ratings. However, the strongest single effect was that between internal locus of control and ESP magnitude by percipient ratings. This item was apparently inspired by the well-known distinction introduced by Rotter (1966) and for which a separate personality scale exists. The concept is intended to distinguish people who believe they are responsible for what happens to them from those who believe their fate is determined more by events outside their control. The extreme alternatives on the PIF were worded as “things just happen” and “people make things happen”. The Rotter scale to the best of my knowledge has been used in only one other ganzfeld experiment, where it failed to correlate significantly with psi (Kanthamani & Palmer, 1993).

	Frequencies (Magnitude)			Frequencies (Direction)			
	L/A	R	Total	L	A	Total	
EXTREME	21	69	90	HITS	11	3	14
MIDDLE	4	56	60	MISSES	4	7	11

Correlates of ESP magnitude & direction

Total	25	125	150	Total	15	10	25
Column Percents				Column Percents			
	L/A	R	Total		L	A	Total
EXTREME	84.0	55.2	60.0	HITS	73.3	30.0	56.0
MIDDLE	16.0	44.8	40.0	MISSES	26.7	70.0	44.0
Total	100.0	100.0	100.0	Total	100.0	100.0	100.0

Table 8: ESP Scores Classified by Magnitude and Direction as a Function of Handedness, Experimenter Rankings

The relationships of ESP scores to experiences of PK and coincidences are similar to findings reported by Honorton (1997) for the PRL database and by Kanthamani and Broughton (1994) from their analyses of the FRNM database. However, whereas these authors found directional effects (more hitting among persons with more personal psi experiences) we find a magnitude effect (more extreme scoring with more personal psi or psi-like experiences). The specific wording of the items could, of course, be the reason for the discrepant findings, but there are other possibilities as well. Two other findings reported by Kanthamani and Broughton (1994) are not replicated in the present analyses, namely significant psi-hitting on the part of practitioners of mental discipline and those classified as F and P on the MBTI. Potential explanations of all these differences between the present analyses and theirs are (1) incomplete overlap of the subject samples analyzed and (2) the fact that Kanthamani and Broughton (1994) based their statistics on the scoring differences between isolated subsamples of subjects and MCE, whereas statistics in the present paper are comparisons between groups of subjects (for example, reported versus did not report practicing a mental discipline).

Finally, a brief word about the handedness effect with experimenter ratings is in order. Psi is often considered a predominately right-hemisphere phenomena (Braud, 1975), so it is not surprising that the greatest magnitude of psi was found among left-handed and ambidextrous subjects, who have more right-hemisphere dominance than do right-handed subjects. Of course, it is disappointing that this trend was not duplicated with percipient ratings. On the other hand, the fact that overall significance of psi scores was obtained only with experimenter ratings suggests that these ratings may be the most accurate representation of the level of psi that actually occurred. If that is the case, it may also be that the significant relationships with percipient ratings, especially when the outcomes were unique to percipient ratings, might represent differences in judging skill. In other words, this skill may be part of what the “ESP” scores from percipient ratings are measuring, in which case the relationships with experimenter ratings should be taken more seriously as correlates of psi.

REFERENCES

- Braud, W. G. (1975). Psi-conducive states. *Journal of Communication*, **25**, 142-152.
- Broughton, R. S., & Alexander, C. H. (1997). Autoganzfeld II: An attempted replication of the PRL ganzfeld research. *Journal of Parapsychology*, **61**,

- Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Honorton, C. (1997). The ganzfeld novice: Four predictors of initial ESP success. *Journal of Parapsychology*, **61**, 143-158.
- Honorton, C., Berger, R. E., Varvoglis, M. P., Quant, M., Derr, P., Schechter, E. I., and Ferrari, D. C. (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., & Schechter, E. I. (1986). Ganzfeld target retrieval with an automated testing system: A model for initial ganzfeld success. *Proceedings of the Parapsychological Association 29th Annual Convention*, Rohnert Park, CA, 401-414.
- Kanthamani, H., & Broughton, R. S. (1994). Institute for Parapsychology ganzfeld-ESP experiments: The manual series. *Proceedings of the Parapsychological Association 37th Annual Convention*, Amsterdam, 182-189.
- Kanthamani, H., & Palmer, J. (1993). A ganzfeld experiment with "subliminal sending". *Journal of Parapsychology*, **57**, 241-257.
- Myers, I. B., & McCaulley, M. H. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Palmer, J. (1997). Correlates of ESP magnitude and direction in the PRL and RRC autoganzfeld data bases. *Proceedings of the Parapsychological Association 40th Annual Convention*, Brighton, UK, 283-298.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, **80**, 1-28.
- Wiseman, R., Smith, M., & Kornbrot, D. (1996). Exploring possible sender-to-experimenter acoustic leakage in the PRL autoganzfeld experiments. *Journal of Parapsychology*, **60**, 97-128.

ESP AND REG PK WITH SEAN HARRIBANCE: THREE NEW STUDIES

John Palmer
Institute for Parapsychology

ABSTRACT

Results are reported from three experiments with the gifted subject Sean Harribance (S.H.) during a visit to the Institute for Parapsychology in April 1997. S.H. had been tested extensively at the Psychical Research Foundation and the Institute for Parapsychology during the late 1960s and early 1970s. The most common test procedure was to have S.H. guess the sex of concealed photographs. He scored significantly above chance in 9 of 14 series, and there was evidence of better scoring on cards placed to his left. He also obtained significant positive results guessing down-through concealed decks of ESP cards and in “psychic shuffle” experiments. Finally, he obtained significant positive results in one free-response ESP experiment in which he gave blind psychic readings. The first of the three studies reported here attempted to replicate a card-guessing experiment reported by Child and Kelly (1973) involving the use of systematically unbalanced decks. We were unable to replicate their chief finding, a significant tendency for S.H. to call symbols that appeared 7 times in the deck more frequently than symbols that appeared 3 times in the deck. Although there were fewer trials ($N = 1500$) than in the Child and Kelly experiment, lack of statistical power cannot explain the failure to replicate. The one significant post-hoc finding was psi-missing the day after which S.H. learned a friend had suffered a serious illness. We were also unable to replicate two previous findings, that S.H. had called more doubles (the same call twice in a row) on runs in which the number of hits were well above chance, and that he had obtained a relatively high proportion of hits on symbols that appeared infrequently in the deck. In the second experiment, S.H. completed 76 100-trial runs on an REG PK game called *Volition* in which subjects attempt to affect the upward or downward drift of a line moving horizontally across a computer screen. In contrast to the normal protocol, S.H. kept his eyes closed during each run and thus received no ongoing feedback. Overall results were in the opposite direction from conscious intent (psi-missing) to an almost significant degree ($p = .053$). Neither S.H.’s choice of aim (upward or downward) nor the hand he used to initiate each run with the computer paddle were significantly associated with the results. In the third experiment, S.H. attempted to influence the performance of an Institute staff member (C.A.) on the computer ESP game *ESPerciser*. In one condition, S.H. attempted to influence C.A.’s scores while present in the room with her. In the other two conditions, S.H. was not present with C.A.: in one of these he attempted remote influence, whereas in the other he did a personal meditation unrelated to the psi task. Because of scheduling difficulties, only 34 of the specified 60 10-trial runs were completed. Overall ESP scores were slightly below chance, but there was significant psi-missing in the meditation condition ($p = .017$). This result supports similar findings from an earlier study with a different psychic.

INTRODUCTION

In the history of parapsychology only a handful of practicing psychics have succeeded in controlled laboratory tests of ESP and PK. The most extensively tested of these is Lalsingh (Sean) Harribance. S.H. was born in Trinidad, where he spent many of his early years giving personal readings at charitable events without pay (Stump, Roll, & Roll, 1970). He subsequently moved to the United States and is currently living in Houston, Texas. In his later years, he has earned his living giving paid readings to clients, many of whom are businessmen, and he has set up

his own foundation. S.H. and his wife have kept an extensive record of predictions he has made concerning a wide range of newsworthy events, along with reports of their corroboration in newspapers and magazines.

S.H. was brought to the attention of American parapsychologists by Mr. Hamlyn Dukhan of Trinidad in the late 1960s (Stump et al. 1970). This introduction led to an extensive series of tests conducted at the Psychical Research Foundation and the Institute for Parapsychology, both located in Durham, North Carolina, during the late 1960s and early 1970s.

The most frequent of these tests used a procedure in which S.H. attempted to guess the sex of pictured in concealed photographs (Klein, 1971; Morris, Roll, Klein, & Wheeler, 1972, Series 10; W. Roll, 1972; Roll & Klein, 1972; Stump et al., 1970). The tests consisted generally of ten-trial runs consisting of five male and five female photos lined up in random order. In the two best controlled experiments with S.H. present in the room with the photos, the latter were covered with a blanket and a wooden frame to prevent visual and tactile clues (W. Roll, 1972; Stump et al., 1970, Series 7). In the other experiments, S.H. and the photos were in different rooms. In one of these, S.H. succeeded in both GESP and clairvoyance trials (Roll & Klein, 1972), but in another he succeeded only in GESP trials (Klein, 1972). All the other series were conducted under clairvoyance conditions. Of the 14 total series, nine yielded significantly positive results. Four of the five failures were suboptimally controlled exploratory series reported by Stump et al. (1970).

Stump et al. (1970) noticed that in their well-controlled Series 7, S.H.'s success was concentrated on the first five trials of the run. The authors attributed this finding to the fact that S.H. attempted to balance his calls on the last five trials so that each sex would be called an equal number of times in the run as a whole. As these latter calls were thus rationally determined to a large extent, one would not expect them to be psi-mediated. Roll and Klein (1972), however, found just the opposite result: significantly better scoring on the last five trials. Call balancing was still present, but it was not as pronounced as in Stump et al.'s (1970) Series 7. Roll and Klein (1972) noted that the seemingly disparate findings could be reconciled by taking into account S.H.'s physical orientation to the photos; in both experiments, S.H. scored best on those photos placed to his left side.

In other experiments, S. H. was asked to guess down through decks of standard ESP cards enclosed in sealed boxes (Morris, 1972; Morris et al., 1972, Series 11; Morris, 1972). Scores were significantly positive in both studies. Morris (1972), who analyzed his data for response bias effects, found that S.H., who had a propensity to call the same symbol twice in a row (doubles) more frequently than expected by chance, called them significantly more often on high scoring runs than on chance runs. S.H. was also most successful on the symbols he called least frequently, a confirmation of Stanford's (1975) response bias hypothesis.

Two other experiments with ESP cards used the "psychic shuffle" procedure (Kanthamani, 1974; Morris, 1973). Either the experimenter or S.H. wrote down an arbitrary, supposedly random sequence of 50 ESP symbols, except in one series where the sequence was derived from a random number table. Then a deck of 50 cards was shuffled with the intention of stopping at the number of shuffles that would cause the resulting sequence of cards to best match the sequence written down previously. Sometimes S.H. shuffled the cards, and sometimes he told the experimenter when he or she should stop shuffling. In all but one of the ten series of trials reported, the scores were significantly positive. Morris (1973) found that S.H. was most successful when the cards

were shuffled a relatively large number of times.

Another successful ESP experiment used a free-response procedure in which S.H. gave personal readings to ten male and ten female target persons (Roll, Morris, Damgaard, Klein, & Roll, 1973). S.H. was told the person's sex and allowed to hold an undeveloped Polaroid photograph of the person, plus strands of hair and a note card with a personal question written on it, both of which were sealed in an opaque envelope. The target persons then blind ranked for personal accuracy all the readings intended for their sex. Results were significantly positive (relatively high rankings for the person's own reading) overall, and separately for females.

Testing for PK was much less extensive than for ESP. S.H. was cited as a particularly successful subject in a study where subjects were asked to resuscitate anesthetized mice more quickly than identically treated and matched control mice (Watkins & Watkins, 1971). S.H. also scored significantly above chance in two of four dice series in which dice were juggled inside an electrically powered rotating canister (M. Roll, 1972). Unfortunately, controls against bias in the target selection procedure were inadequate in the two successful series.

A number of additional psi experiments with S.H. that employed a variety of procedures are briefly summarized by Damgaard (1972). They included five card-guessing studies, one free-response ESP study, and one attempt to influence the behavior of goldfish by PK. Except for two of the card-guessing studies, all results were nonsignificant. Two short clairvoyance series with an electronic random event generator (REG) also yielded nonsignificant total scores (Kelly, Hartwell, & Artley, 1978; Kelly & Lenz, 1976). These nonsignificant findings are not sufficient to cancel out the significant findings reported above.

S.H. visited the Institute for Parapsychology from April 2 to April 6, 1997, to participate in further testing. The visit was originally S.H.'s idea, as he felt a strong need to provide additional scientific evidence, not so much for his own psi abilities, but for psi generally. He also offered to pay his own travel expenses. As soon as it became logistically possible for us to arrange the visit, we were pleased to do so.

S.H. participated in a number of different formal tests during his visit. Some involved the exploration of neuropsychological factors, and these will be discussed in a separate paper. In this paper, I will focus on the exclusively behavioral research.

CARD GUESSING WITH UNBALANCED DECKS

The most theoretically interesting experiment conducted during the testing of S.H. in the 1970s was an ESP clairvoyance experiment in which the decks consisted of fixed but unequal numbers of the five standard ESP symbols: star, circle, cross, square, and waves (Child & Kelly, 1973). Specifically, the frequencies in each deck of 25 cards were 9, 7, 5, 3, and 1, with the symbols appearing a given number of times counterbalanced across decks. The authors were primarily interested to learn if S.H. would pick up on the relative frequencies of the different symbols independently of correct matches on individual trials. For example, would he call the card appearing 9 times in the deck more frequently than the card appearing just once? If so, this would provide evidence for a global or gestalt-like perception of the target sequence. I was particularly interested in this process because I had found possible evidence for it in the data of some of Helmut Schmidt's high-scoring subjects tested with REGs (Palmer, 1996a).

Overall, S.H. obtained 21.1% hits over 9,000 trials (360 runs), which gave a z of 2.69, $p < .01$. His hit rates on the five target frequencies were about the same. However, in line with the global perception hypothesis, he called symbols appearing 7 times in the decks significantly more frequently than those appearing 3 times, and this was true even when only misses were analyzed. However, the 9-1 split yielded no such response preference.

I decided to conduct an essentially strict replication of this experiment with the hope of finding additional evidence for the global perception hypothesis.

Method

The time available for the study allowed for only 60 runs, which was quite a bit less than the 360 runs administered by Child and Kelly (1973). However, some of this loss was recouped by eliminating the 9-1 splits, which were not successful in their experiment; by converting these to 7-3 splits, the number of such splits was doubled. Thus, the frequencies of symbols in the decks were 7-7-5-3-3.

The decks were assembled as follows. First, 60 sequences consisting of different random permutations of the numbers 1 to 25 were created by using a random permutation program in Turbo Basic seeded by the time interval in milliseconds between two keyboard button presses. This source has proved to be satisfactorily random in numerous previous experiments in which it has been used. A simple code was then devised to convert the numbers to ESP symbols. For instance, in one run the numbers 1-7 might become (7) circles, 8-14 (7) squares, 15-19 (5) crosses, 20-22 (3) waves, and 23-25 (3) stars. It turned out that 60 runs allowed each possible combination to be used once but only once, so perfect counterbalancing was achieved. A second random permutation of the numbers 1-60 from the same source was then used to randomize the order of the decks. The data sheets were then handed over to an administrative assistant, Lisa Woods, who assembled the actual decks of cards, placing each in a cardboard packet with the backs of the cards facing the top, that is, the side without the opening. Each packet then was sealed and the run number written on it. The packets were kept in Ms. Wood's office until ready for use.

Eleven test sessions were dispersed over the five days of S.H.'s visit. The plan was to have no more than six runs per session, with the precise number of runs per session based on S.H.'s preference. Nine of the sessions consisted of 6 runs and two consisted of 3 runs. For each session, S.H. and I were seated on opposite sides of a table facing each other. Before each run, I placed the packet containing the appropriate deck of cards face up on the table equidistant between S.H. and myself. S.H. made his calls as I recorded them on a standard record sheet. The calls were tape recorded, which proved to be necessary because S.H. sometimes made his calls very rapidly, making it hard for me to keep up. After the run, I opened the pack and recorded the card sequence adjacent to the call sequence on the record sheet and circled the hits. This was not the case in runs 37-48 and 50-60, when at S.H.'s request, two decks were called in a row, with feedback occurring after the second run. From S.H.'s point of view, this was like calling a single 50-trial run.

The written call sequences were transferred to computer for analysis, and the call sequences checked against the tape recording. This was done by a student assistant not otherwise involved with the project.¹

¹ I wish to thank Mr. Spencer Shephard for his contributions to the research.

During the checking of the recorded calls and targets it was found that decks 9, 10, and 11 had been improperly assembled. Deck 9 had 8 squares but only 2 stars; deck 10 had 4 squares and 2 stars; in deck 11, cross, waves, and square all appeared five times, leaving only one 7-3 split for this run. These errors reduced by 4 the number of 7-3 splits available for analysis. They also degraded the theoretically perfect counterbalancing, but the harm is trivial. Finally, in run 58 S.H. made only 24 calls, as I lost track of the trial number and stopped S.H. one trial too soon.

Results and Discussion

Over the 1499 scoreable trials, S.H. obtained 286 hits, for a hit rate of 19.1%, where MCE is 20%. This is associated with a z of -0.89 , which is not close to being statistically significant. Post-hoc analysis revealed significant psi-missing on runs 19-30, which were conducted, at S.H.'s request, in two back-to-back sessions on the morning of April 4. (This constituted 20% of the total trials.) S.H. obtained 44 hits over these 300 trials (14.7%), $z = -2.31$, $p = .021$. This finding is reported only because S.H. had been informed the night before that a close friend and business associate had just suffered a serious heart attack. This news might have affected his mood, especially in the morning sessions when all but three of these runs were conducted.

Following the procedure used by Child and Kelly (1973), the global perception hypothesis was analyzed by calculating the proportion of runs in which 7-targets were called more frequently than 3-targets. Because of the errors noted above, only 59 runs were usable for this analysis. (Run 10 had no 3-targets and was eliminated; runs 9 and 11 each had one 7-target and one 3-target and thus were usable.) There were 24 runs that supported the hypothesis ($7 > 3$) and 29 that reversed it ($3 > 7$); 6 runs were ties. This outcome is associated with a nonsignificant corrected z of -0.28 .

A secondary trial-based analysis was no more encouraging. Again eliminating run 10, there were 579 calls on 7-targets and 586 calls on 3-targets. This difference, in the opposite direction from that hypothesized, is nonsignificant, $z = -0.21$.

Because the overall hit rate was so close to chance, it was not deemed useful to conduct separate analyses for miss trials.

Two sets of post-hoc analyses were computed. Child and Kelly (1973) reported no significant tendency for the hit rates to differ on targets as a function of their frequency. We found no significant pattern either. S. H. obtained 148 hits on 826 7-targets (17.9%), $z = -1.50$, $p = .134$; 70 hits on 347 3-targets (20.1%), $z = -0.03$; and 65 hits on 310 5-targets (21.0%), $z = +0.43$.

Second, I wanted to see if I could replicate the findings reported by Morris (1972) concerning S.H.'s tendency to call an excess of doubles on high-scoring runs and to score more hits on the symbols called least frequently. However, it needs to be kept in mind that there are three relevant respects in which this study differed from Morris's. First, Morris did not use unbalanced decks. Second, S.H. scored significantly above chance in Morris's study. Third, many fewer runs were collected in this study than in Morris's. In the present experiment, S.H. called an average of 21.8% doubles, which, echoing Morris (1972), is high to a suggestive degree, $z = 1.73$, $p = .084$, 1-tailed. Morris had compared S.H.'s rate of calling doubles on runs with a score of 11 (high) to that on runs with a score of 5 (chance). As S.H. had no run scores higher than 9 in the present experiment, I decided a priori to define high scoring runs as those in which S.H.'s score was 8 or higher. In 7 high-scoring runs, S.H. called doubles 17.8% of the time, $z = -0.89$, which is in the opposite direction from that found by Morris. On 9 chance runs, S.H. called doubles 22.2% of

the time, $z = 0.82$. The difference between these two percentages does not approach significance, $\chi^2(1) = 1.45$, $p = .229$. As for S.H.'s hit rate on the symbols called least frequently, it was found, contrary to Morris (1972), that he tended to score below chance on symbols that he called three times or less per run. Over 174 such trials, this hit rate was 14.9%, $z = -1.67$, $p = .095$. The hit rate jumped to 23.7% for symbols called four times, and the percentages for more frequently called symbols demonstrated no linear pattern.

Overall, the results of this experiment were disappointing, and the trends reversed what had been found by Child and Kelly (1973) and by Morris (1972). Although the sample size was low as compared to these other studies, it is doubtful that this factor can account for the failures to replicate. One possibly relevant difference is that S.H. indicated awareness that this was an unbalanced deck study even though he was not explicitly told this. He probably linked it in his mind to the earlier experiment with Child and Kelly (1973), in which he was told about the design during debriefing.

REG PK EXPERIMENT

Method

A second experiment conducted with S.H. during his visit used a computer PK game called *Volition*. Developed by the late Charles Honorton, *Volition* uses the hardware REG board developed at the Psychophysical Research Laboratories, which is connected to an Apple II microcomputer. During the task, subjects sit in front of the computer screen and try to make a line move toward the top or bottom of the screen as it advances from left to right. This line reflects the cumulative deviation of binary PK trials produced by the REG, with each segment of the line representing the integrated result of 100 binary bits. Each trial is initiated by a button press on the Apple's paddle. Following each run, the data are transferred to disk and summary statistics are presented on the screen as feedback.

In the present experiment, each run consisted of 100 trials, which S.H. generated continuously by keeping the button on the paddle pressed down. He was allowed to choose before each run whether to aim for a high or low score. During some preliminary testing designed to acquaint S.H. with *Volition*, we tried an option which kept the screen blank during the run itself. S.H. kept his eyes closed during these runs, and the computer sounded a tone when the run ended. Following each run, analysis software was activated that reproduced the completed graph on the screen and provided the summary statistics. Because S.H. preferred this non-visual method, and it seemed to be producing good results, it was decided to use it for the formal experiment.

S.H. completed 76 runs in three sessions consisting of 25, 25, and 26 runs respectively. The extra run was added to the last session so that there could be an equal number of runs in each handedness condition (see below). Each session was conducted on a separate day. Which hand S.H. used to press the button was systematically manipulated: in the first session, all presses were made with the right hand; in the second session, all presses were made with the left hand; in the third session, the right hand was used on odd-numbered trials and the left hand on even-numbered trials. The reason for this manipulation has to do with a theory applicable to this task called *decision augmentation theory* (DAT) (May, Utts, & Spottiswoode, 1995). According to DAT, so-called micro-PK effects are not created by influencing the output of the REG, but rather by initiating the sampling of the bit stream at a precise moment that captures a sequence that happens

by chance to contain a “significant” excess of 0s or 1s. Because this moment is defined by the timing of the button press, which in turn is controlled to some extent by the brain hemisphere contralateral to the hand pressing the button, I thought such a manipulation might capture brain hemisphere differences in psi mediation.

The time of S.H.’s visit overlapped with that of another experiment using the same REG for psi testing.² Randomicity tests of the REG were conducted for this latter experiment using the analysis routines included in the software package containing the *Volition* game, as well as the *ESPCriscer* game employed in the third experiment with S.H. (Psychophysical Research Laboratories, 1985). These randomicity tests, which compute chi-square and Kolmogorov-Smirnov tests on individual bytes as well as a generalized serial test assessing dependencies of successive bits up to length eight, provided no evidence of nonrandomness.

Results

The mean z-score over all 76 runs was $-.216$, which is very close to significant in the psi-missing direction, $t(75) = 1.964$, $p = .053$. This effect was strongest in the third session, during one stage of which 16 of 17 consecutive runs yielded negative zs. The mean z for this session was $-.365$, $t(25) = -2.21$, $p = .036$.

S.H. produced a mean z-score of $-.277$ on 46 high-aim runs and $-.121$ on 30 low-aim runs. This difference is not significant, $t(74) = 0.70$, $p = .488$. Likewise, the handedness manipulation had no effect on scoring. The mean z-score was $-.218$ for the 38 left-hand runs and $-.213$ for the 38 right-hand runs, $t(74) = -0.03$, $p = .979$.

EXTERNAL INFLUENCE ON ESP TASK

In October, 1994 I conducted an experiment in which a gifted subject, Malcolm Bessent (M.B.), was asked to remotely influence the performance of C.A. on a computerized ESP game (Palmer, 1996b). Evidence for such influence was found, but only in a control condition in which M.B. was engaged in either reading or meditation. I decided to conduct a related experiment with S.H. that included a condition in which he attempted to influence the performance of Cheryl Alexander (C.A.), a Research Associate at the Institute, while he was in the room with her.

Method

The psi task made use of *ESPCriscer*, a computer game developed by Charles Honorton that combines the principles of forced-choice and free-response ESP and allows for testing in both clairvoyant and precognitive modes. The software was run on the same Apple II computer, and the same REG was used as in the *Volition* experiment described above.

Subjects interact with the computer by means of a game paddle. After optional instructions appear on the screen, the subject is presented with four white rectangles. These provide the subject with the opportunity to generate spontaneous mental impressions of the target, as in a free-response ESP task. When the subject completes this process, a button press on the paddle causes the white rectangles to be replaced with four geometric designs, imagery-evoking words, or drawings of common objects or living beings. The four choices for each trial are randomly selected from among 24 sets or “packs”, and the positions of the four target alternatives on the screen are also determined randomly for each trial. The subject rotates the dial on the paddle to

² This experiment, conducted by Dr. Richard Broughton to test Dr. Edwin May’s data augmentation theory, also used the *Volition* game. The report of this research has not yet been submitted for publication.

mark which of these alternatives he or she believes is the target and then presses the button to register his or her choice. The screen then asks the subject to indicate whether this response was based on an impression (an image or verbal association), a feeling (of being drawn to one of the choices), or a guess (no specific reason for the choice).

For each run of up to 24 trials, the REG randomly determines if the mode is to be clairvoyance or precognition. The subject does not know which has been chosen. If it is clairvoyance, the REG selects the target during the impression period; if it is precognition, it selects the target after the subject indicates the basis for his or her choice (impression, etc.).

After the subject makes this choice, the three incorrect target alternatives are blacked out, thereby providing feedback of the actual target. If the choice was correct, the word HIT appears in large block letters at the bottom of the screen, the screen flashes on and off, and the computer adds reinforcing sound effects. At the end of the run, a table appears on the screen that presents the number of total hits for the run as well as for each response basis (impression, etc.), the probability of the result being due to chance, and whether the mode was clairvoyance or precognition. The data are then stored on disk for later analysis.

The original design called for C.A. to complete 20 10-trial games of ESPerciser in each of three conditions. In Condition 1, S.H. attempted to influence her psychically to get a good score while seated in the room with her. He did not tell her what guesses to make. In Condition 2, S.H. attempted the same kind of influence from a room on a different floor of the building. In Condition 3, S.H. was in the same room as in Condition 2, but he did a personal meditation unrelated to the ESP task. During Conditions 2 and 3, C.A. was kept blind as to which of these conditions was assigned for the session. There were to be two 10-run sessions in each condition, with the order of conditions being: 2,3,1,1,3,2. Unfortunately, because of scheduling difficulties, we were not able to complete the pre-assigned number of trials. The second sessions for Conditions 2 and 3 were not undertaken. C.A. broke off the second session of Condition 1 because she was feeling very tired, and we were unable to schedule the remaining trials. Thus, there were 14 total trials in Condition 1 and 10 trials in each of the other two conditions.

Results and Discussion

Over all 34 runs, C.A. obtained a mean of 2.21 hits per run, where 2.50 is expected by chance. This value is not significant, $t(33) = -1.30$, $p = .203$. Consistent with her past performance (Palmer, 1996b), C.A. scored more positively in the clairvoyance mode (mean = 2.50) than in the precognition mode (mean = 1.88). The precognition value is significantly below chance, $t(15) = 2.18$, $p = .046$. However, the two means do not differ significantly from each other, $t(32) = 1.40$, $p = .172$.

In Condition 1 (S.H. present in room), C.A. scored an average of 2.71 hits per run, which is slightly above chance. In Condition 2 (remote influence), she scored slightly below chance (mean = 2.20). In Condition 3 (meditation), she averaged 1.50 hits per run, which is significantly below chance $t(9) = 2.92$, $p = .017$. The difference among the three means approached significance in a one-way analysis of variance, $F(2,31) = 2.72$, $p = .081$.

The scheduling difficulties introduced the possibility of an optional stopping artifact, although scoring levels had nothing to do with our decisions. The stopping was most "optional" in the second session of Condition 1, in which C.A. chose to stop after four runs. The average score on

these runs was 3.00, which is slightly higher than the mean of 2.6 for the first session of Condition 1. However, the most impressive finding was not in this condition, but rather in Condition 3.

The significant psi-missing in Condition 3, the one in which S.H. meditated, is noteworthy because it is most comparable to the condition in which the psychic M.B. influenced C.A.'s scoring in the earlier experiment of this type (Palmer, 1996b). In that experiment, the influence occurred on an aspect of C.A.'s performance that she was not attending to (differential scoring on targets with different colored backgrounds), whereas in the present experiment the parameter of influence was simply the overall ESP score. Also, whereas the results in the experiment with M.B. were in the direction he intended during the influence period, S.H.'s results were in the opposite direction, psi-missing instead of psi-hitting.

GENERAL DISCUSSION

The results obtained with S.H. in this series of experiments did not reach the levels found in the research conducted with him in the 1970s. This is perhaps attributable in part to the relatively small number of trials we were able to generate, but this cannot account for the very close to chance findings obtained in the replication of the unbalanced deck study.

It is noteworthy that the marginal evidence of psi that did appear in this series of experiments was all in the psi-missing direction. This evidence includes the post-hoc discovery of significant psi-missing on the morning after S. H. learned of his friend's heart attack in the unbalanced deck study, the close to significant psi-missing in the Volition PK study, and the significant psi-missing in the meditation condition of the external influence study.

News of his friend's heart attack is not the only psychological factor that may have contributed to these negative findings. It was at S.H.'s request that this research project was undertaken. He was eager to do it because of a strong need he felt to prove to other scientists that psi is real. He expressed extreme frustration at the slow pace at which this acceptance has been forthcoming and seemed to feel a personal responsibility to remedy the situation. What this means from a testing standpoint is that S.H. put a great deal of pressure on himself to demonstrate a convincing level of psi in the present project. In my experience with psi testing, such an attitude is counter-productive, leading either to chance scoring, tight variance, or psi-missing. Perhaps that is what happened in this case.

To me, the most encouraging finding was the psi-missing in the meditation condition of the external influence experiment. Except for the direction of scoring, it provides a conceptual replication of the finding I reported earlier with Malcolm Bessent (Palmer, 1996b), and the change of direction is understandable in light of the mood factors discussed above. It adds support to the hypothesis that PK or active-agent telepathy effects might be most likely to occur when someone is not actively trying to produce them but simply hopes for success. It also provides indirect evidence for experimenter psi, in that this hopeful but passive frame of mind is presumably the one adopted by most psi experimenters or principal investigators during the course of testing.

REFERENCES

- Child, I. L., & Kelly, E. F. (1973). ESP with unbalanced decks: A study of the process in an exceptional subject. *Journal of Parapsychology*, **37**, 278-297.

- Damgaard, J. (1972). Recent research with Lalsingh Harribance: A summary of additional research with Lalsingh Harribance [Abstract]. In W. G. Roll, R. L. Morris, & J. D. Morris (Eds.), *Proceedings of the Parapsychological Association, Number 8, 1971* (pp. 74-76). Durham, NC: Parapsychological Association.
- Kanthamani, H. (1974). Psi in relation to task complexity. *Journal of Parapsychology*, **38**, 154-162.
- Kelly, E. F., Hartwell, J. W., & Artley, J. L. (1978). A second two-choice clairvoyance EEG study with Lalsingh Harribance [Abstract]. *Journal of Parapsychology*, **42**, 54-55.
- Kelly, E. F., & Lenz, J. E., Jr. (1975). EEG correlates of trial-by-trial performance in a two-choice clairvoyance task [Abstract]. In J. D. Morris, W. G. Roll, and R. L. Morris (Eds.), *Research in parapsychology 1975* (pp. 22-25). Metuchen, NJ: Scarecrow Press.
- Klein, J. (1972). Recent research with Lalsingh Harribance: A comparison of clairvoyance and telepathy [Abstract]. In W. G. Roll, R. L. Morris, & J. D. Morris (Eds.), *Proceedings of the Parapsychological Association, Number 8, 1971* (pp. 71-72). Durham, NC: Parapsychological Association.
- May, E. C., Utts, J. M., & Spottiswoode, S. J. P. (1995). Decision Augmentation Theory: Toward a model of anomalous mental phenomena. *Journal of Parapsychology*, **59**, 195-220.
- Morris, R. L. (1972). Recent research with Lalsingh Harribance: Guessing habits and ESP [Abstract]. In W. G. Roll, R. L. Morris, & J. D. Morris (Eds.), *Proceedings of the Parapsychological Association, Number 8, 1971* (pp. 72-74). Durham, NC: Parapsychological Association.
- Morris, R. L. (1973). Complex psi and the concept of precognition [Abstract]. In W. G. Roll, R. L. Morris, and J. D. Morris (Eds.), *Research in parapsychology 1972* (pp. 95-97). Metuchen, NJ: Scarecrow Press.
- Morris, R. L., Roll, W. G., Klein, J., & Wheeler, G. (1972). EEG patterns and ESP results in forced choice experiments with Lalsingh Harribance. *Journal of the American Society for Psychical Research*, **66**, 253-268.
- Palmer, J. (1996a). Evaluation of a conventional interpretation of Helmut Schmidt's automated precognition experiments. *Journal of Parapsychology*, **60**, 149-170.
- Palmer, J. (1996b). External psi influence on ESP task performance. *Journal of Parapsychology*, **60**, 193-210.
- Psychophysical Research Laboratories (1985). *PsiLab][user's manual*. Princeton, NJ: Psychophysical Research Laboratories.
- Roll, M. (1972). Some recent dice-throwing studies with Lalsingh Harribance [Abstract]. In W. G. Roll, R. L. Morris, & J. D. Morris (Eds.), *Proceedings of the Parapsychological Association, Number 8, 1971* (pp. 19-21). Durham, NC: Parapsychological Association.
- Roll, W. G. (1972). Recent research with Lalsingh Harribance: Physical aspects of the target [Abstract]. In W. G. Roll, R. L. Morris, & J. D. Morris (Eds.), *Proceedings of the Parapsychological Association, Number 8, 1971* (pp. 67-69). Durham, NC: Parapsychological Association.
- Roll, W. G., & Klein, J. (1972). Further forced-choice experiments with Lalsingh Harribance. *Journal of the American Society for Psychical Research*, **66**, 103-112.
- Roll, W. G., Morris, R. L., Damgaard, J. A., Klein, J., & Roll, M. (1973). Free verbal response experiments with Lalsingh Harribance. *Journal of the American Society for Psychical Research*, **67**, 197-207.
- Stanford, R. G. (1975). Response patterns in extrasensory performance. *Journal of Communication*, **25**, 153-161.

Correlates of ESP magnitude & direction

Stump, J. P., Roll, W. G., & Roll, M. (1970). Some exploratory forced-choice ESP experiments with Lalsingh Harribance. *Journal of the American Society for Psychical Research*, **64**, 421-431.

Watkins, G. K., & Watkins, A. M. (1971). Possible PK influence on the resuscitation of anesthetized mice. *Journal of Parapsychology*, **35**, 257-272.

CURRENT RESEARCH IN GIVING THE GANZFELD AN OLD AND A NEW TWIST

Adrian Parker and Joakim Westerlund
University of Gothenburg and University of Stockholm

ABSTRACT:

An update on the Gothenburg studies using the manual ganzfeld is presented. The two most recent studies, studies IV and V, have replicated the earlier ones with a direct hit frequencies of respectively 47% and 37%. The data base for all five studies (N=150) is clearly significant with an overall mean frequency for direct hits of 35% ($z = 2.92$). The four monitored studies gave a significantly higher effect than the non-monitored study and taken together have a direct hit frequency of 39% which is statistically highly significant ($z = 3.59$). In an attempt to reduce the number of sessions required to attain statistical significance, the final study reported here departed from the standard ganzfeld by randomly viewing all four clips as targets, with ten minute periods per clip in one session. It also incorporated a control series in order to evaluate qualitative hits. Results were only encouraging for those who previously had succeeded with the standard ganzfeld. A way of combining the virtues of the two approaches in the form of a digital computerised ganzfeld is discussed. The objective of the digital ganzfeld is to improve the efficiency of ganzfeld by the use of two target film clips per session and by having access to real time recordings of the mentation reports during the target review procedure.

INTRODUCTION

While the meta-analysis of recent work with the ganzfeld (Milton and Wiseman 1997) can be criticised for its lack of definition as to what constitutes a ganzfeld study, it appears clear that the recent studies have shown the technique to have a varying degree of success. There is consequently an real need to specify the variables that determine successful outcome. While it is tempting to explain the replication difficulty in hitherto unspecified psychological variables, an attempt has also been made to postulate an expected decline from the theoretical point of the quantum physics, in this case being due to an inherent balancing effect in nature (Bierman 1994). Nevertheless, even if such effect is supposed to be present, one way of maintaining significant outcomes, is for future ganzfeld experiments continually add new variations or twists. New variations on a theme have of course a psychological interest in creating the necessary novelty in order for the experimenter to maintain the psi-conducive conditions.. Moreover, those researchers who have been involved in more than a hundred ganzfeld sessions can testify to this need for a variation on a theme.

In considering such variations in design, the results of the above meta-analysis may teach us some caution should be exercised in order not to deviate too far from the successful recipe of the basic technique. In considering such a design, it was reasoned that priority should be given to reducing the number of man-hours spent on Ganzfeld and to generally improving the efficiency of the technique. Given that one session can take two or more hours, the demands of having to carry out

Correlates of ESP magnitude & direction

a 100 or more sessions in order to first determine whether or not there is a psi effect, becomes an obvious hindrance to encouraging other researchers to enter this field.

There would then seem to be two considerations influencing future research with the ganzfeld: As well as a need for innovation in the design, the above mentioned meta-analysis reinforces the need for a conservative and systematic approach in the form of exact replications and efforts to specify as far as possible the psi-conducive conditions. These two considerations may not be entirely incompatible and in the current work at Gothenburg we have opted for a complementary approach.

The conservative approach is illustrated by the three previous series of studies that have been at Gothenburg and the two current follow up series reported below. In these we have used a manual ganzfeld, each with a pre-set N of 30 trials (Parker, Frederiksen & Johansson 1997). Analysis of the results from these studies gave support for a superiority of scores obtained with auditory monitoring of the receiver's mentation reports in place and with targets showing a content of unexpected emotional change. Additional predictors of hits were found to be Myers-Briggs scores polarising on Feeling as a personality type, as well as particularly high scores on the Sheep-Goat Scale and the Magical Ideation Scale. Further analysis of scores on Sheep-Goat Scale and Magical ideation scale found the effect to be limited not only to belief in and experience of subjective psi, but also to correlate significantly with the so-called psychotic items of the magical ideation scale (Parker, Grams, & Petterson 1998), thereby replicating a previous finding by Michael Thalbourne (1994).

The magical ideation scale is a heterogeneous and sub-clinical scale. High scores on it may not be a pure indicator of pathology but may to some extent be an expression of the common thread to the relationships that appear to go through a series of closely related processes such as creativity, fantasy proneness, absorption as well as schizotypal thinking. This has recently termed this transliminality and a scale has been developed to measure this (Thalbourne et al 1997). Accordingly, the transliminality scale is being used in our current work and the last two studies / 4 and %) have been designed to test predictions from this and other measures.

Study 4 was initially intended to replicate the previous work with music students and a further series was to be aimed at obtaining high scores by applying the above criteria for success that had been arrived at during the first three series. Unfortunately the experimenter who was to be involved in the testing of music students became ill soon after the start of the series, so in practice participants were recruited mainly from the groups (new age and personal psi-experiences) that had previously been successful during earlier series 2 and 3.

Study 5 followed a recruiting procedure similar to that used in the first three studies. We advertised amongst the general population for participants reporting spontaneous ESP experiences. As well as further replication the purpose of this study was to further examine the relationship of various measures such as the transliminality scale to psi-scores.

The more progressive approach has concentrated on improving the efficiency of the ganzfeld. For this we decided to focus on aspects of the standard ganzfeld that have not been put to use: the three other films in the series and the often striking qualitative nature of the hits. It was reasoned that if the psi-source functions in a similar way to that of other external influences on hypnagogic imagery, then its influence may not be limited to one film at one point in time. Indeed it would be of psychological interest to see if the ganzfeld imagery could change with a changing psi source. It

was therefore decided to use all four video clips as the psi-source during an extended ganzfeld. The presentation of the four video clips is randomised order and the task of the receiver is then to re-constitute this viewing order. The probability of identifying the correct order of the first film would be 1 in four, and the probability of placing the second one in correct would become 1 in 3 leaving the third and four films to be placed at odds of 1 in 2. The probability of placing all four clips in the correct order thus becomes the sum of these probabilities: 1 in 24. Clearly if the psi influence is fairly constant in strength over time this should enormously increase the efficiency of the procedure.

Nevertheless, there are some potential disadvantages of this procedure. Although far from all participants think in terms of probabilities, some might regard the task as a daunting one. An equally important aspect concerns the risk that the qualitative aspects might lose their striking subjective significance. Since correspondences in the all the mentation report are sought with all four video reports, it will naturally be harder to subjectively evaluate the role of chance in such correspondences that occur. Having obtained and begun to document impressive hits in our earlier series, it was important to include in future designs a means of evaluating these especially the synchronous aspects to the correspondences between mentation reports and the contents of the video clip that was actually being shown at the time. Qualitative aspects are a much neglected area although an impressive documentation of good hits was collected by Charles Honorton and a similar collection has been made by Kathy Dalton from her work at Edinburgh. Despite the sophistication of the autoganzfeld, the synchronicity aspect could not be evaluated since it was apparently not part of the programmed set-up in the autoganzfeld.

Collaboration with Jan Dalkvist at the University of Stockholm led to the simple innovation of including a stereo-video-player in order to make a copy of the clip as it is being shown while at the same feeding in the sound from the contemporary mentation report to the one of the video's sound channels. The result of this would be to obtain a recording of the target film clips along with the accompanying mentation reports as occurring in real time.

The problem remains of how to evaluate account for the role of chance correspondences in qualitative hits and the solution we have opted for is to incorporate a second mono video-player and stereo video-recorder in set-up in order to run a control series. This control series is chosen randomly from equivalent series and the order of the clips are also played randomly and copied by a video-recorder with a stereo input which enables the sound input from the mentation reports to simultaneously recorded along with the film being viewed.

This is a obviously a new innovation requiring new precautions against. One elementary precaution is that the two stereo-recorders and the two mono video-players should be operated simultaneously by the same remote controls. Experimental and control clips need to be also of the same length. Differences in starting and stopping points between the control and experimental tapes could otherwise give cues as to the identity of the tapes..

Another aspect is defining the time limits for the window of observation. The degree of divergence in time that is allowed from the points of correspondence between the mentation report and the video clip, is something which needs to be experimentally determined.

STUDIES IV AND V

These studies were intended to be replications of the previous auditory monitored series using

groups that were predicted to maximise the scoring rate. Additional analyses are also being performed on various tests and measures given to the participants in an attempt to identify some further predictors of psi-hitting.

Methodology

A conventional manual ganzfeld was used here with a 30 tape of seashore waves and with auditory monitoring of the receiver's mentation reports. Target clips were chosen from video sets by means of random number tables. The final entry points into the random number tables was reached at by five successive dice throws. Essentially the same methodology was used as that described in the earlier monitored studies except for that in this case a new video library had been made up which when completed consisted of 28 sets of four video tapes together with composite tapes containing a duplicate recording of all four clips, to be used in the judging process. The receiver team and the sender team were isolated until feedback. Each session complete with introduction, some psychological test administration, as well as post session feedback lasted between 2 and 3 hours.

Procedure

As has been described elsewhere (Parker, Johansson, & Frederiksen 1997), emphasis was placed on promoting a expectancy of success at the same time as being careful not to evoke achievement demands. An informal introductory talk is arranged during which the procedure is discussed over coffee. It is emphasised that we test the usefulness of technique rather the participants. Two teams are involved, the experimenter who stays with the receiver and the assistant experimenter who accompanies the sender (and can even function as an sender in those cases where the receiver was unable to bring a friend).

Subjects

The study IV began using students recruited from the School of Music in Gothenburg but for reasons mentioned earlier it was not possible to continue to recruit from this source so the study had to be completed by recruiting from our previous sources: New Age groups and advertisements for those reporting spontaneous paranormal experiences.

Study V used participants that either responded to an advertisement in a local newspaper for individuals reporting spontaneous paranormal experiences or to a later article in the same newspaper about our research.

For both studies, those contacting us and who were thought to be suitable participants, were asked to bring with them a person whom they believed could function well together with in a potentially telepathic situation.

Hypotheses

Both studies followed the essential formula for success that had evolved from the previous two studies. This is to recruit participants from groups claiming to have genuine psi experiences and as described previously to attempt to create a psi-conducive atmosphere in the experimental situation. It was therefore expected that the scoring rate would be similar to that of the psychic experiences groups occurring in studies II and III, that is approximately 40%. Only direct hits were to be used for this analysis.

It was also considered important to replicate these findings with a new generation of experimenters. For 42 of these 60 trials, AP had, if any, a very peripheral role in the experiments. The experiments were conducted by students carrying the experiments as part of their Masters thesis.

Results.

For study IV, the hit rate was 47% with a size effect of .46 ($z = 2.53, P = .006$, one-tailed) thereby replicating the previous two studies. Study V gave a direct hit rate identical with that of studies II and III, namely 37%.

The data base at this point gives a z score of 2.92 ($p = .002$ one-tailed) and an effect size of .24. Since the results for the auditory series differ significantly from the non-auditory series, some justification can be made for treating these separately. The four Auditory monitored studies produced a hit rate of 40% and an effect size of .33 ($z = 3.59, p = .0002$ one-tailed) see table 1.

Series	Trials	Hits	Frequency	z score	Size effect
Study I: (Non-monitored)	30	6	20%	- 0.84	-.08
Study II (monitored)	30	11	37%	1.26	.23
Study III (monitored)	30	11	37%	1.26	.23
Study IV (monitored)	30	14	47%	2.53	.46
Study V (monitored)	30	11	37%	1.26	.23
Data base to date	150	53	35%	2.92	.24
Data base for the monitored studies	120	47	39%	3.59	.33

Table 1: Overall Results from the Gothenburg Ganzfeld Series

SERIAL GANZFELD

The use of the serial ganzfeld here is essentially exploratory, aimed at seeing if the psi-effect can be magnified through the use of sequential targets. The use of a series of four different targets is a departure from the standard ganzfeld procedure and therefore the results are presented separately.

A second purpose of using this technique was that of providing a further “selection test” for those who had succeeded with the previous manual ganzfeld technique in order to identify members of a group who would be promising to work further with.

Methodology

Subjects

Correlates of ESP magnitude & direction

Subjects were recruited from some of those who had taken part in testing (N=7) and from those who had responded to a current newspaper article about our work. In addition, two of the experimenters used themselves as subjects. Since the technique is obviously more demanding in terms of psi-performance, the aim was recruit as many participants as possible from those who had succeeded with a standard ganzfeld. In practice time constrictions made it only possible to use eight of these.

Procedure

The same basic introduction and orientation procedure that had been used in the standard ganzfeld was followed. The series consisted of 30 trials (each with different receivers). The Ganzfeld session lasted in this case 46 minutes, and consisted of a 6 minutes of relaxation music followed from 40 minutes of sea shore waves. Eight minutes constituted a sending period followed by 2 minutes intervals to allow for the interchange of tapes.

In accordance with the arrangement described earlier the use of a stereo-video player enabled a recording of the receiver's mentation reports to be fed on to a copy of the clips as it was being recorded.. This would provide a later means of studying possible synchronous correspondences. By the latter half of the series, in accordance with the above rationale, a control series was finally incorporated into the set-up.

Hypotheses

As reasoned above a completely correct placement should occur by chance once in 24 trials. A permutations table shows that two correct placements would occur 6 times in the 24 permutations and one correct placement 8 times in 24 trials. Converting to N =30 gives an expectancy table of:

4 correct:	1.25
2 correct:	7.5
1 correct:	10
0 correct:	11.25

The hypotheses were 1) that there would a significantly higher number of all four and of two . correct placements than chance expectancy and 2) that those participants who had had previous experience with the manual ganzfeld would give significantly higher scores .

Results

The results were well in accord with those calculated above on the basis of chance expectancy.

The serial ganzfeld was originally intended as a further selection gate for those who we would finally select as special ganzfeld subjects. It was decided to make a post-hoc analysis for novices subjects compared with those who had succeeded with an earlier an manual ganzfeld. On this basis and in accordance with the hypothesis of counting correct matches of 4 and 2 as hits, the experienced group obtained 4 hits out of the 8 trials and the novice group 3 hits out of 22 trials (chi squared =.3.04, P<.08).

Correct placements	Number obtained (N=30)
--------------------	---------------------------

Hits = 4	1
Hits = 2	8
Hits = 1	7
Hits = 0	14

Table 2: Correct placements of the order of sending

CONCLUSIONS

It is very tempting to conclude that the results of these studies teach us that it is wiser to stay with old suitors than court new ones, but maybe it is possible to learn from both. Clearly the standard Ganzfeld can only be fully reliable if and when we can identify the underlying factors characterising successful ganzfeld subjects. Although the results of the five studies are very encouraging, the as yet identified predictors (psychic experiences, belief in psi or even high scores on the Myers-Briggs Feeling type) may prove too wide as variables to reliably guarantee significant outcomes in studies with only 30 sessions.

As for the serial ganzfeld, there may be several reasons why the study was not very successful. The first of these suggested, albeit weakly, by the data: It does not appear to be a suitable task for novice subjects. In trying out the technique, we also discovered some practical difficulties. The addition of a control series made the procedure burdensome. It became quite simply at times clumsy and stressful to attempt to handle the insertion and removal of 8 video tapes in one session. Moreover in order to avoid cues that might identify the experimental and control tapes, no stopping and re-playing of tapes should be allowed. We attempted to deal with this by using longer 4 minute clips which were recorded twice on the target tapes. This however gave little time to concentrate on the contents of specific scenes in the film clips and once again made the procedure somewhat stressful. Finally, these longer recordings and short (8 minute) sending periods may have deprived us of what earlier had seemed to be an important cue in making a correct matching in the standard procedure: the occurrence of repeated themes in the mentation reports.

A solution to this would be to combine the attractive aspects of both procedures. The new set up we are developing will use a digital computerised presentation of the target film and decoy film sequences as well as the synchronous auditory recording of the mentation reports.

We propose to present two target film clips per session, each chosen from its own set of 4 possible targets. The advantage of this is that it means that each individual subject making two correct identifications will have the possibility of individually obtaining a result which borders on statistical significance ($P = .06$) (whereas two correct placements in the serial ganzfeld is in practice not even equivalent to one direct hit in the standard ganzfeld).

During the target review procedure, the films will these will be on large projection screens and the target assessment can then actually take place in the receiver room thereby making use of situation specific learning.

The review of the mentation report will first follow that of the standard ganzfeld procedure in viewing each film in the set, one at a time. Finally all four films will be viewed in different sections

Correlates of ESP magnitude & direction

of the screen at the same time as segments of the recorded mentation report is played back in real time with the target film and the three other decoys. This will enable participants to actually hear their mentation reports recorded with the target clip as it was viewed and with the decoy film clips that were not viewed. Synchronous (real time recording) of the mentation reports with the target film clips will also enable an accurate study and evaluation of qualitatively good hits to be made a later date. By using this new technology and special subjects, it is hoped we can obtain level of hitting in which that something truly new can be learned about psi.

REFERENCES

- Bierman, D. (1994) The decline of the ganzfeld: Patterns in elusiveness. In A. Parker . et al symposium. *Proceedings of Presented Papers, 37th Parapsychological Association Convention*, Amsterdam.
- Milton, J. & Wiseman, R. (1997) Ganzfeld at the cross-roads. A meta-analysis of the new generation of studies. *Proceedings of Presented Papers, 40th Parapsychological Association Convention*, Brighton.
- Parker, A. & Grams, D.(1997) Some variables relating to success in the ganzfeld. *Proceedings of Presented Papers, 40th Parapsychological Association Convention*, Brighton.
- Parker, A. Grams, D. & Petterson, C. (1998) Further analysis of variables relating to success in the ganzfeld. Paper to be submitted to *Journal of Parapsychology*
- Parker, A., Frederiksen, A., & Johansson, H. (1997) Towards specifying the recipe for Success with the Ganzfeld: Replication of the Ganzfeld findings using a manual Ganzfeld with Subjects reporting paranormal experiences. *European Journal of Parapsychology* 13, 15-27.
- Thalbourne, M. A. (1994) Belief in the paranormal and its relationship to schizophrenia-relevant measures: A confirmatory study. *British Journal of Clinical Psychology*, 33, 78-80.
- Thalbourne; M. et al. (1997) Transliminality: Its nature and correlates. *Journal of the American Society for Psychical Research*, 91, 305-331.

EFFECTS OF DISTANT HEALING INTENTION THROUGH TIME AND SPACE: TWO EXPLORATORY STUDIES

Dean I. Radin
Consciousness Research Laboratory
University of Nevada, Las Vegas¹

Fátima Regina Machado and Wellington Zangari
Interpsi, São Paulo, Brazil²

ABSTRACT

Two double-blind experiments examined fluctuations in an individual's autonomic nervous system when distant healers directed their intention towards them. The target individual's respiration, heart rate, fingertip blood volume, and spontaneous electrodermal activity were continuously monitored during 20 randomly counterbalanced, one-minute "treatment" and control epochs. The first experiment examined the effects of a group's healing intention directed in real-time at volunteers isolated 200 meters away. The second experiment studied the effects of Umbanda mediums located in São Paulo, Brazil, who directed their healing intentions towards volunteers monitored two months earlier in Las Vegas, Nevada.

The first study showed that the groups' healing intention was associated with an increase in breathing rate ($p = 0.053$, two-tail) and a decrease in electrodermal activity ($p = 0.055$, two-tail) in isolated volunteers. The second study showed that despite the separation of 6,000 miles in space and two months in time, the mediums' healing intention was associated with an increase in the distant volunteers' fingertip blood volume ($p = 0.013$, two-tail) and an increase in electrodermal activity ($p = 0.031$, two-tail).

Alternative explanations for the reported effects are considered and the experimental outcomes and implications are discussed.

INTRODUCTION

Studies investigating the effects of distant intentionality on the human autonomic nervous system have shown small, statistically significant modulations of autonomic activity at distances of a few dozen meters to many kilometers (Radin, 1997). For effects that apparently transcend space, we may expect – for relativistic reasons – that they also transcend time. A few pilot studies have confirmed this suspicion (e.g., Braud, David & Wood, 1979), leading Schlitz and Braud (1997) to speculate that distant intentionality effects "... could allow direct attentional and intentional influences to be directed 'backwards in time' to influence probabilistic events involved in seed moments or initial formative conditions harmful or helpful to health and well-being" (p. 72). Given the pragmatic value of genuine retroactive healing, it behooves us to examine the possibility of time-separation effects in distant intentionality.

Another prospect about distant intentionality that merits further observation is the broad popularity of prayer circles and healing groups. Combined with evidence suggesting that "bonded pairs" produce larger effects than individuals in mind-matter interaction studies (Dunne, 1993), and possible "field effects" associated with the mental activities of groups (Bierman, 1996; Nelson

et al, 1996; Radin 1997), it is possible that distant healing intention (DHI) by groups may be more efficacious than the same intention offered by an individual.

To explore these issues, two experiments were conducted, the first examining DHI effects by groups, and the second examining time-space separations in distant intentionality.

EXPERIMENT 1: GROUP DISTANT HEALING INTENTION

This study examined the effects of a group's collective DHI on the autonomic functioning of a distant individual. In this exploratory study, no attempt was made to compare group vs. individual DHI effects. The group was asked to send non-directional healing intention towards a member of the group who was isolated at a distance. "Non-directional" means maintaining beneficial, nurturing thoughts without a specific healing outcome in mind. We speculated that DHI effects would manifest as autonomic relaxation because prayer is similar to meditative, relaxed mental states (Benson, 1992), but the hypothesis was simply that the physiological status during treatment epochs would differ from that during control epochs, thus two-tailed probabilities were employed.

Participants

The experimenter in this study (E, the first author) recruited two groups of volunteers from the Las Vegas, Nevada community for this experiment. Group 1 consisted of 21 people taking part in a six-week, spiritually-oriented self-improvement program. Group 2 consisted of 10 people who were part of an Edgar Cayce study group. This latter group had been meeting in Las Vegas for nearly 20 years, and many of the founding members were still in the group.³ Individuals in both groups maintained high degrees of belief in the efficacy of distant mental healing effects.

Equipment

Participants playing the role of distant "patient" were continuously monitored throughout experimental sessions with a psychophysiological monitor made by Coulbourn Instruments, model LabLinc V. The raw signals from the LabLinc V were fed through a Data Translation DT-2821 analog-to-digital board housed in a Packard Bell 100 MHz Pentium-based PC, and controlled by Datawave Technologies "Experimenter's Workbench" software.

The Lablinc V consisted of modules for monitoring electrodermal activity (Model V71-23, set at 1000 mV/ μ Siemens, with AC coupling and DC excitation), a photoplethysmograph for monitoring blood volume pulse in the fingertip and heart rate (Model V71-40), breathing volume measured by a bellows transduced through a strain gauge coupler (Model V72-25), and skin temperature change from a reference point set to zero at the beginning of the session (Model V71-30, set at 1 V/degree C). Data were recorded at 10 samples per second and streamed to the hard disk. In addition, an IBM Thinkpad 700 laptop provided a time-of-day signal monitored by a LabLinc bioamplifier module (Model V75-01), and an IBM Thinkpad 350 was used to automatically provide the healers with counterbalanced instructions.

The following six variables were formed from the raw physiological measures:

- respiration rate per minute, averaged over the preceding five seconds
- respiration volume (chest expansion), averaged over the preceding five seconds
- heart rate in beats per minute, averaged over the preceding three seconds
- fingertip blood volume, averaged over the preceding three seconds

- skin conductance response
- skin temperature on the pad of the left thumb

Procedure

Volunteer “patients”

E arranged for the group participants to meet in the Consciousness Research Laboratory at the University of Nevada, Las Vegas. When the group had assembled, E escorted them to a comfortable conference room on the other side of the building (see Figure 1). E explained the purpose of the experiment, described the basic procedure, and then everyone signed an informed consent form. After a short discussion to clarify the procedure, the group selected one person from the group to act as the DHI target or “patient” (P).⁴ The method of selecting P for each session was left up to the group.

Before leaving the room with P to return to the laboratory, E gave instructions for the group to wait for a nearby telephone to ring. This ring would provide a signal for the group to begin the session. Then E and P returned to the laboratory via a path of about 200 meters, and through three locked doors, as illustrated in Figure 1.

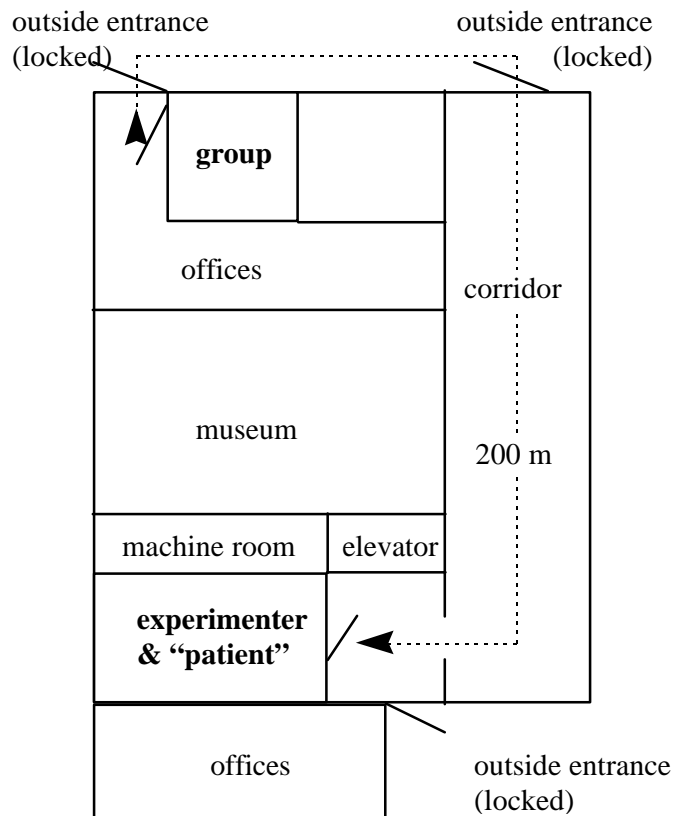


Figure 1. Schematic of the laboratory and group meeting rooms.

In the lab, E asked P to fill out a short questionnaire asking the first four questions shown in Table

Effects of distant healing

1. At the end of the session P filled out the last question in Table 1. Then E asked P to sit in a reclining chair and remove his or her wristwatch. The watch was removed because it might restrict blood flow in the hand, and it also eliminated any way for P to precisely track the passage of time.

Q1) How do you feel before the session begins? (fatigued / depressed, to agitated / anxious)
Q2) What interest or experience do you have with distant mental healing? (none at all, to extensive)
Q3) Do you believe that distant mental healing will be effective in this experiment? (not at all, to absolutely)
Q4) What is your motivation to be healed. (none at all, to extreme)
Q5) How do you feel after the session ended?

Table 1. Each participant answered these five questions using a seven-point scale.

To the left hand, E attached (1) a photoplethysmograph to the pad of the third finger with a sponge clamp, (2) Na/Na-Cl electrodermal electrodes moistened with electrolyte conductive cream to the pads of the first and second fingers with velcro fasteners, and (3) a thermister to the pad of the thumb. Then a bellows was comfortably fitted around P's chest with a velcro fastener, and P was asked to rest his or her left hand on a pillow. A colorful, randomly moving screen saver was displayed on a computer monitor about 10 feet in front of P to provide him or her with something interesting to look at. Many participants reported that they kept their eyes closed during the session.

After ensuring that the electrode signals were being detected properly, E started the physiological data recording and then dialed a telephone located just outside the conference room where the group was waiting. E let the phone ring twice, then hung up. P knew that the session would begin shortly after E placed the phone call, and that it would last about 20 minutes, but neither E or P had any indication of when the session actually began, or when the individual epochs began, nor did they know anything about the assignment of the treatment or control conditions during those epochs. In addition, the group did not receive any feedback about P's physiological state during the session. At the end of the data recording period, and after removing the electrodes, discussing P's experiences, and having him or her answer the last question on the questionnaire, E accompanied P back to the conferencing room where the group was waiting. Another person volunteered to be P, and the process was repeated.

Volunteer "healers"

The phone ringing near the group acted as a signal for one person in the group to press a button on an IBM Thinkpad 365 laptop computer in the room. This button-press began the session, and it also allowed the laptop to record session timing information and to provide instructions on when and how the group should apply its collective intention. This worked as follows: As soon as a member of the group pressed a button on the laptop, a program running on the laptop set a pseudorandom number generator seed number to the current clock time and generated 10 random bits.⁵ If a bit was 1, this would result in the sequence treatment-control (T C), and if the bit was a

0 the order would be control–treatment (C T). These 10 random bits were thus used to create a sequence of 20 randomly counterbalanced epochs. Immediately after generating this sequence, the program began the session.

Each T or C epoch was 44 seconds, followed by a 16 second inter-epoch period, making each session last approximately 20 minutes.⁶ At the beginning of each successive epoch, the laptop automatically indicated that the group should either (a) send collective healing intentions towards P upon seeing P’s photo displayed on the screen (the T condition), or (b) they should rest and return their intention back towards the group upon seeing a pleasant landscape photo on the screen (the C condition). At the end of each 44-second T or C epoch, the laptop indicated that the group should prepare for a possible change of instructions by displaying a photo of the Buddha. Data recorded during the 16-second inter-epoch periods were not analyzed.⁷

Analytical Method

To evaluate the probability of the difference observed between physiological waveforms recorded during the treatment and control conditions, random permutation analysis (RPA) was employed. Recommended by Blair & Karniski (1993) for evaluating psychophysiological measures recorded in two different conditions, this method provides the probability of the difference between any two arbitrary time-series waveforms without making any assumptions about the shape of the distributions or sequential dependencies among adjacent physiological measures. This method is especially useful because it provides an arbitrarily close estimate of the exact probability of the observed difference. The method is explained in detail in Appendix A.

Hypothesis

The hypothesis was that a group’s distant healing intention would be associated with significant changes in distant individuals’ average autonomic physiology in treatment versus control conditions. Because no direction was predicted, two-tailed probabilities were employed. Other hypotheses considered in the original formulation of this study will be discussed in future publications.

A total of 15 sessions was planned, and it was decided in advance that no data would be analyzed until after all sessions were completed. The number of sessions was selected based upon the number of trials conducted in similar DHI experiments by Braud and Schiltz (1991), and in a similar successful DHI study conducted by Radin, Taylor and Braud (1995).

RESULTS

Participants

Group 1 came to the Consciousness Research Laboratory once in the evening and contributed four sessions over the course of four hours. Group 2 came to the lab four times, each for about three hours in the evening, contributing a total of 11 sessions.

Table 2 lists information about the 15 participants and their answers to the questionnaire. The average age of the 10 women and 5 men was 52.1, ranging from 31 to 70. Their answers to the questions about how they felt before versus after the session showed that nearly all moved from various levels of discomfort (answers above or below the midpoint of 4) to comfort. This is not surprising given that the experiment required them to simply relax in a reclining chair for 20

Effects of distant healing

minutes, knowing that their friends were sending beneficial, nurturing thoughts. Their answers to questions about experience, belief and motivation to be healed were fairly high, reflecting the enthusiasm in these groups for distant healing.

ID	Session Date	Gender	Feeling Before	Feeling After	Experience	Belief	Motivation
701	5/23/97	M	5	4	6	7	5
702	5/23/97	F	4	4	7	7	7
703	5/23/97	M	4	4	7	7	7
704	5/23/97	F	3	4	4	6	4
705	5/27/97	F	2	4	7	7	7
706	5/27/97	F	1	4	7	4	4
707	5/27/97	F	4	4	5	7	7
708	6/3/97	F	2	4	7	7	7
709	6/3/97	M	2	5	7	6	5
710	6/3/97	F	5	4	6	7	6
711	6/10/97	F	3	3	5	7	5
712	6/10/97	F	1	4	7	7	5
713	6/10/97	F	4	4	7	7	7
714	6/17/97	M	4	4	5	6	6
715	6/17/97	M	3	4	5	5	4

Table 2. Participant information overview.

The numbers in the Before, After, Experience, Belief and Motivation columns refer to responses on the 7-point scales shown in Table 1.

Procedural change

In sessions 1 through 4, contributed by Group 1, the original plan was for the laptop to display a digitized photo of P or a landscape scene as instructions for the T and C epochs, respectively. Unfortunately, there were so many people in Group 1 that not everyone could clearly see the computer screen. Thus, one person was selected to watch the laptop screen and announce aloud when a treatment, control or inter-epoch period had begun. This proved to be a successful means of letting the group know what it should concentrate on, and when.

In sessions 5 through 15, with Group 2, a few participants again mentioned that they had trouble seeing the screen, and most indicated that they preferred to keep their eyes closed throughout the session. As a result, the program was altered to have the laptop play a pre-recorded digital audio file in which E spoke the phrase “Please send healing energies now” at the beginning of treatment epochs, “Please prepare for a change of instructions” at the beginning of an inter-epoch period, and “Please return your attention to the group” at the beginning of control epochs. The laptop’s audio output was amplified through an external speaker system to allow everyone in the group to easily hear these instructions.

Physiological analysis

After all sessions were completed, three data files from each session were consolidated into a single file. File 1 consisted of the raw physiological data recorded on the Packard Bell computer, File 2 consisted of time-stamp data recorded by the Thinkpad 700 laptop (used to time-synch the physiological data), and File 3 consisted of the sequence and time of counterbalanced conditions recorded on the Thinkpad 365 laptop (used by the group). To ensure proper time-

synchronization among these files, the clocks of the two laptops were both set to Universal Time as maintained by the U. S. Naval Observatory.⁸ The combined dataset was analyzed using randomized permutation analysis (see Appendix A) to compare the average physiological results recorded in the two counterbalanced conditions. Table 3 shows the results.

The effect size obtained for EDA, $r = -0.496$ ($p = .055$, two-tail), is nearly twice as large as the meta-analytic estimate of $r = .25$ for previous distant intentionality studies involving electrodermal activity (Braud & Schlitz, 1989, 1991; Radin, 1997).⁹ There is no meta-analytic effect size estimate for changes in breathing rate in DHI experiments, but that too was nearly significant in this study with $r = .499$ ($p = .053$, two-tailed). None of the other physiological measures were significantly different from chance.

Group	p	z	r	2-tail p
breathing rate	0.027	1.932	0.499	0.053
breathing volume	0.425	0.189	0.049	0.850
blood volume pulse	0.546	-0.115	-0.030	0.908
electrodermal activity	0.973	-1.922	-0.496	0.055
heart rate	0.636	-0.348	-0.090	0.728

Table 3. Results of random permutation analysis applied to the data collected in the group experiment.

The value p is the one-tailed probability of obtaining treatment-control differences as large or larger than the observed difference; z is p transformed into a standard normal deviate; r is effect size determined as z/\sqrt{N} , where $N = 15$ (Rosenthal, 1991); and the last column is the 2-tailed p.

Note that Table 3 does not list skin temperature, which was one of the physiological signals monitored. After discussing the design of this study with several psychophysicologists, we were convinced that changes in skin temperature were too slow-moving to be reliably detected in the space of the 44 second epochs employed in this study, thus this factor was dropped from further analysis.¹⁰

EXPERIMENT 2: BRAZIL

This study examined the effects of DHI through space and time. Professional healers in São Paulo, Brazil, over 6,000 miles from Las Vegas, were asked to apply their healing intention to volunteers whose physiology was monitored two months before in the Consciousness Research Laboratory in Las Vegas.

The second and third authors invited spiritual healers from the Umbanda tradition in Brazil to act as the healers in this study. Umbanda is a religion that evolved from an amalgam of religious cultures and practices, primarily Catholicism, Candomblé and Kardecist Spiritism (Pordeus, 1996). Umbanda is a spiritistic religion, in that it stresses the relationship between humans and nature, especially nature spirits, through rituals called “giras.” Giras are generally practiced in spiritual settings like temples, called “terreiros.” Each terreiro is led by an individual who is also the spiritual advisor to the Umbanda mediums. The mediums act as intermediaries between a panoply of spirits evoked by the giras and individuals who ask for their intercession. Umbanda spirits include Oxalá (associated with Jesus), Iemanjá (associated with Mary), Exus (amoral

spirits), and Pombas-giras (prostitutes). Giras rituals are commonly used for healing, giving advice, and providing magical spells.

Participants and Equipment

The same equipment, physiological measurements, and questionnaire used in the first study were also used in this study.¹¹ Volunteer “patients” were recruited by the first author in Las Vegas in May 1997, and the healers were recruited by the second and third authors in São Paulo in July and August 1997.

Procedure

Volunteer “patients”

When a volunteer arrived at the Consciousness Research Laboratory in Las Vegas, E explained the purpose of the experiment, described the procedure, and then the participant signed an informed consent form. As in the Group study, E invited P to fill out a questionnaire asking the questions shown in Table 1. Then E took four color digital photos of P, attached the electrodes and checked the physiological signals as before. When P indicated that he or she was ready to begin, E started the physiological software and for the next 24 minutes physiology was continuously recorded as P simply relaxed in the reclining chair.¹²

Volunteer healers

Software similar to that used in the first experiment was also employed in this study. The program, along with copies of all digital photos of the participants, their original signed consent forms, and a compressed version of all of the digitized physiological data (i.e., the raw data) were sent to the two experimenters in São Paulo, Brazil (E_{SP}) for the healing part of the experiment. Most sessions were conducted in a terreiro in São Paulo that had agreed to participate in this study. When a medium was ready to begin the healing session, E_{SP} pressed a button on a computer. This caused the software to generate a sequence of 20 randomly counterbalanced recording epochs, as in the Group study.

For this study, rather than displaying a picture of the Buddha to indicate the inter-epoch periods, a picture of the Umbandan spirit “Omolu” was used. The second author had selected the picture of Omolu because she liked it, learning only later that Omolu is a divinity associated with diseases, smallpox in particular (Bastide, 1989, p. 564). In this study, each recording epoch was planned to be 50 seconds, followed by a 10 second inter-epoch period.¹³

Analytical Method

The design and analysis of DHI experiments assumes synchronized timing between the healers’ intention and the recording of the participants’ physiological data. This time-matching makes it possible to assign a meaning to the various counterbalanced segments of the data. However, in this study, given that healers were asked to focus their intention “backwards in time,” we must clarify our temporal assumptions. Time-synchronization ordinarily combines two factors: Synchronization *between* healer and participant sessions, and also synchronization *within* the sessions. Another way to state this is that we must assume that our clocks start at the same

moment in time, and we assume that our clocks are running at the same speed.

In the present study, we also make these two assumptions, except that the way the clocks started, and the way they maintained the same standard clock speed, was *by agreement* among the healers, participants, and experimenters. There are other ways of interpreting what may be happening in this type of study (as discussed later), but for now we will simply define that the moment E began a physiological recording session in Las Vegas was the same as the moment that E_{SP} began a healing session in São Paulo. We further assume that both clocks from that point ran at the same speed (at least until the end of a session).

After all healing sessions were completed, E_{SP} mailed the computer files containing the healers' counterbalanced instructions, one file per session, to the first author. These files also contained date and timestamps for each session. This data was evaluated using the same randomized permutation method used to evaluate the Group study data.

Hypothesis

The hypothesis was that the mediums' healing intention in São Paulo in July and August 1997 would be associated with significant changes in the individuals' autonomic physiology recorded in Las Vegas in May 1997. Directional effects were not predicted, thus two-tailed probabilities was used.

RESULTS

Subjective reports

Twenty sessions were planned, but a total of 21 volunteers wished to participate, so 21 sessions were conducted. Table 4 summarizes information about the 21 volunteers and their answers to the questionnaire. The average age of the 13 women and 8 men was 50.3, ranging from 28 to 71. Participants' answers to questions about experience, belief and motivation showed that they were relatively enthusiastic about distant healing intention, and again they moved from various stages of discomfort before the session to fairly uniform levels of comfort afterwards.

Session details

The "patient" recording sessions in the laboratory at the University of Nevada were uniformly unremarkable; most participants enjoyed relaxing in the chair for 20 minutes. The sessions with healers were more varied.

ID	Patients Time	Patients Date	Healers Time	Healers Date	Gender	Befor e	After	Experienc e	Belief	Motivation
601	13:36	5/19/97	21:19	7/25/97	F	4	4	1	3	3
602	11:22	5/21/97	21:44	7/25/97	F	5	5	6	7	7
603	12:00	5/21/97	22:11	7/25/97	M	1	4	4	4	7
604	14:34	5/21/97	21:06	7/30/97	M	4	4	6	4	5
605	12:28	5/22/97	21:31	7/30/97	M	1	4	7	7	7
606	16:12	5/22/97	21:56	7/30/97	F	6	4	7	6	7
607	11:32	5/23/97	22:20	7/30/97	F	5	4	6	6	6
608	11:01	5/25/97	21:13	8/12/97	M	3	4	1	1	4
609	12:02	5/25/97	21:31	8/12/97	F	4	4	4	5	5

Effects of distant healing

610	11:02	5/26/97	21:55	8/12/97	M	3	6	5	5	7
611	11:56	5/26/97	22:19	8/12/97	M	4	3	4	3	7
612	12:40	5/26/97	22:43	8/12/97	F	4	4	4	4	4
613	13:39	5/29/97	23:07	8/12/97	F	1	4	7	7	7
614	14:34	5/29/97	21:04	8/14/97	F	6	4	6	7	7
615	15:41	5/29/97	21:27	8/14/97	F	5	4	6	7	6
616	17:50	5/29/97	21:51	8/14/97	F	3	4	6	7	6
617	14:50	5/30/97	22:14	8/14/97	M	4	4	5	6	7
618	16:22	5/30/97	22:38	8/14/97	F	4	4	4	7	7
619	17:06	5/30/97	23:02	8/14/97	F	4	4	7	7	4
620	18:46	5/30/97	21:54	8/19/97	M	4	4	7	7	4
621	19:53	5/30/97	22:21	8/19/97	F	4	4	5	7	7

Table 4. Participant information overview for the Brazil study.

Three of the 21 healing sessions involved people who were not Umbanda healers but held strong beliefs in distant mental intention effects. These included sessions 620 and 621, which were not conducted in the terreiro but in E_{SP}'s office in São Paulo. Session 620 involved two people acting as healers, because they preferred to work together. Session 608 was originally run with a medium in the terreiro, but E_{SP} accidentally mistyped the epoch time-length, which caused the session to last 12 minutes instead of the planned 20. As a result, the same session was run again later in E_{SP}'s office with a non-Umbandan healer. Session 608 therefore had two different, partially overlapping counterbalanced sequences, but to avoid data selection this session was included in the overall analysis using the treatment/control sequence generated when it was run the second time.

The first three sessions and session 605 were conducted in silence in the terreiro. During the other sessions in the terreiro and in E_{SP}'s office, New-Age music was played. To assist the healers' focus in a given session, they gazed at a photo of the "target" person displayed on the computer screen, and they also held the original consent form signed by that individual. It was explained to the healers that the distant "patients" were to be affected according to when the session had actually taken place, i.e. *in the past*, when their photos were taken and when they had signed the consent forms. The healers indicated that they understood this instruction.

Physiological analysis

The effect size obtained for EDA, $r = 0.470$, was nearly twice as large as the previous meta-analytic estimate of $r = 0.25$, as was the effect size for changes in blood volume at $r = 0.54$ (Radin, 1997). The other physiological measures did not significantly differ from chance.

Brazil	p	z	es	2-tail p
breathing rate	0.402	0.247	0.054	0.805
breathing volume	0.854	-1.055	-0.230	0.291
blood volume pulse	0.007	2.473	0.540	0.013
electrodermal activity	0.016	2.155	0.470	0.031
heart rate	0.229	0.743	0.162	0.457

Table 5. Results of 10,000 random permutations for the Brazil experiment.

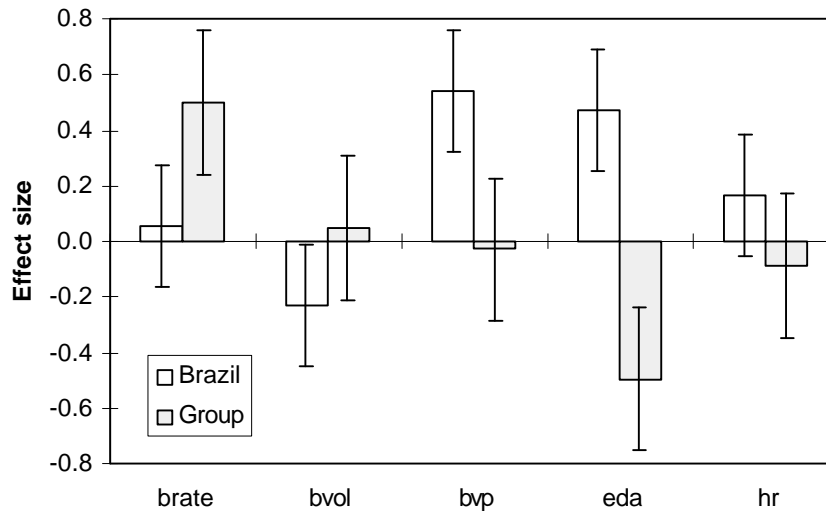


Figure 2. Effect sizes for physiological measures in the Group and Brazil studies, with one standard error bars. The values correspond to breathing rate, breathing volume, blood volume pulse in the fingertip, electrodermal activity, and heart rate.

Figure 2 shows the effect sizes (with ± 1 standard error bars) for both studies. Based on the observed changes in EDA, these results suggest that on average individuals in the Group study responded to distant healing intention by becoming more relaxed, whereas on average individuals in the Brazil study became more aroused. However, the overall picture is actually more complicated because breathing rate increased for the Group participants, which seems contrary to systemic relaxation, and fingertip blood volume increased for the Brazil participants, which seems contrary to systemic arousal.

DISCUSSION

Alternative explanations

Can these results be explained as design flaws, or artifacts? Schlitz and Braud (1997) suggest six alternative explanations for successful results in these types of experiments.

1. *The results are due to sensory or subliminal cues passing between the healers and “patient” participants.* The participants and experimenter in the Group study were isolated from the healers by 200 meters, numerous concrete walls and three locked doors. In the Brazil study, they were separated by two months and 6,000 miles. Given that there were no other forms of communication between the healers and participants in these studies, this hypothesis can be rejected.
2. *The results are due to internal physiological rhythms that may have influenced the participant’s autonomic nervous system activity.* This was ruled out through the use of a randomly counterbalanced schedule of treatment and control periods. With proper random sequencing, there should be no systematic biasing effects due to physiological drifts or rhythms. Examination of the actual counterbalanced sequences (TC vs. CT) used in the two experiments, shown in Figure 3, confirm that the counterbalancing was adequately random. That is, by chance we would expect approximately equal counts of the sequence TC as CT.

Given a total of 15 sessions for the Group study, and 21 for the Brazil study, the expected frequencies for TC are 7.5 and 10.5, respectively. The expected counts are shown as dashed lines. Chi-squared tests show that the observed frequencies were distributed uniformly at random ($\chi^2_{\text{Group}} = 2.2, p = .99, \chi^2_{\text{Brazil}} = 4.4, p = .88$).

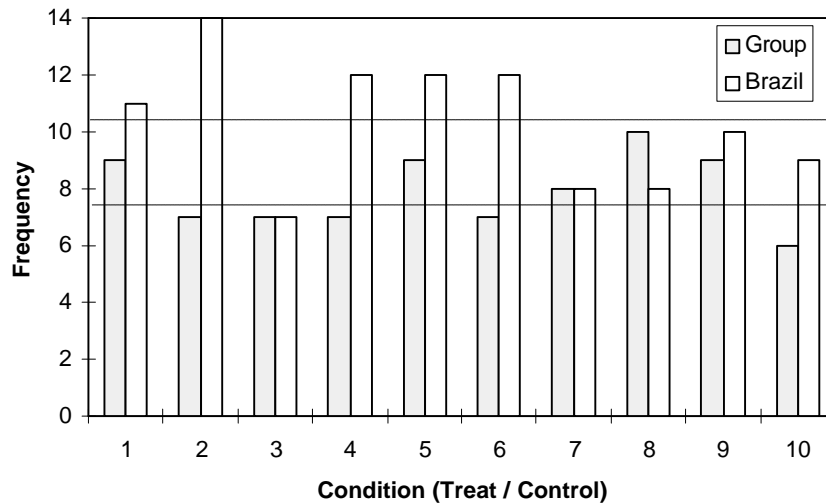


Figure 3. Count of the number of times the sequence “TC” was used in the Group and Brazil experiments, over all epochs (i.e., there were 20 epochs per session, based on 10 counterbalanced sequence-pairs).

3. *The results are due to chance correspondences between the healers’ intentions and the participants’ physiological responses.* The random permutation analysis used to estimate the probability of these differences showed that in both experiments, without making any assumptions about sample distributions or inter-sample autocorrelations, differences in two of the five physiological measures were nearly significant at $p < .05$.¹⁴
4. *The results are due to recording errors or misreadings of the data.* All data recording was fully automated. In addition, in the Brazil study, copies of all experimental materials and the raw physiological data were sent to the experimenters in São Paulo before the healers’ counterbalanced sequence data were generated. It is unlikely that this hypothesis can explain the results.
5. *Participants self-regulated their physiology to conform to the experimenter’s expectations.* The double-blind protocol prevented the experimenter and participant from knowing anything about the healers’ randomly counterbalanced sequence of treatment and control epochs, thus the experimenter had no expectations that could be conveyed to P.
6. *The results are due to arbitrary selection of data.* The number of sessions in each study was planned in advance, and all data were analyzed at the end of the study.

In sum, common methodological flaws that may generate artifactual results in these studies were taken into account by the experimental design. The obtained results are not adequately explained by known design flaws, and they are consistent with the results of previous experiments.

Experimental Differences

Why did the Group and Brazil experiments result in nearly significant changes in EDA, but in opposite directions? The difference between the EDA outcomes in the two studies is significant (see Table 6).

Condition	p (two-tail)
breathing rate	0.233
breathing volume	0.379
blood volume pulse	0.067
electrodermal activity	0.004
heart rate	0.440

Table 6. Difference between physiological results in the Group and Brazil experiments.

We speculate that the difference was due to how the participants felt about being the object of remote intention in the context of the two experiments. Members of the Group experiment knew each other well, and they periodically came together as a group expressly to focus their intention on each other. Thus, Group participants would probably associate remote group intention with their feelings about local group intention – healing intention results in a pleasant, meditative, relaxed state. By contrast, the healers and patients in the Brazil study never met face-to-face, and the participants were probably more anxious about being remotely “affected” from the future, by spiritual mediums 6,000 miles away.

The empirical evidence supports this speculation. Experiments investigating “the feeling of being stared at” show that when the starers and starees are relative strangers, EDA significantly rises when remote staring takes place (Schlitz & Braud, 1997). It is as though the body reflexively responds to remote staring as a threat, and this in turn causes autonomic arousal. However, individuals who are specifically trained in a dyadic meditative practice whereby they gaze into their partner’s eyes while meditating, show a significant drop in EDA (Braud, Shafer & Andrews, 1993a,b).

Interpretations

If these results are not due to obvious artifacts or to chance, then what is happening? Does DHI really influence a distant person’s autonomic physiology? At first blush, this is the obvious implication (let’s call it Interpretation 1), but influence is just one of several possible interpretations.

Interpretation 2

Observed effects may be due to psi perception (Bem & Honorton, 1994), in which the distant person unconsciously perceives the “nonlocal intention” of the healer and uses this suggestion to self-regulate their physiology. The distinction between influence and perception is important, because DHI as a causal influencing force might be able to catalyze or even force dramatic healing effects to take place, such as those produced by the use of biochemical interventions like antibiotics. By contrast, DHI as a remote perceptual effect might be limited to producing still-effective, but quite a different sort of effect, similar to what can be achieved through mind-body techniques like hypnosis, placebo and biofeedback.

Interpretation 3

Another possibility is that the outcome of the experiment is determined by the individual who hits the button on the laptop computer to begin the healer's session (let's call him or her the "starter"). This would require precognition, or perhaps PK. Recall that when the laptop button is hit, a program generates random counterbalanced assignments for the treatment and control epochs. The random sequence depends on when the button is hit, thus if the starter had precognition and could perceive the future results of the experiment, then he or she could simply start the session at a favorable time. Or, if the starter were able to influence the computer, this could also result in a favorable counterbalanced sequence.

Precognition as a viable explanation depends on how long the starter has to wait. If the results of a given experiment require the starter to wait for years, precognition becomes less plausible. In the present studies, the counterbalanced sequence is generated by a pseudorandom number generator, which in turn is set by a seed number linked to the computer system clock. When the button is pressed, the program reads the clock time to the nearest second, and that is used as the seed number. If we assume that the participant's future physiology was not externally influenced at all, then we could obtain a new experimental outcome once per second, as the computer generates a new counterbalanced sequence once per second.

How many seconds would the starter have to wait for a favorable time in the present studies? In the Brazil study, two of the five physiological measures were significant at the $p < .032$ level, two-tailed. Assuming that these measures were independent (they aren't, but say they are), the cumulative binomial probability of up to two measures significant at this level out of five is $p = 0.0003$, which corresponds to a standard normal deviate (z score) of 3.42. If we assume this $z = 3.42$ is akin to a single, combined z score over 21 sessions, then we would require an average score of $z = 0.75$ over each of 21 sessions to achieve this overall level of significance. This average is equivalent to a probability of $p = 0.28$. Events with $p = 0.28$ occur by chance about 1 in 4.4 times, assuming independent events. In the present experiment, this means that in principle the starter need wait only 4 to 5 seconds to get a favorable start time. Thus, in spite of the fact that none of the starters for these studies reported any conscious intent to wait for favorable times to begin the sessions, nor did they get session-by-session feedback (ever, by design), we cannot exclude the possibility that precognitive timing "caused" the present results. The precognitive timing hypothesis has been developed in detail in the "Decision Augmentation Theory" of May, Utts and Spottiswoode (1995).

Interpretation 4

A fourth interpretation of these results is that nothing unusual happened until the data were analyzed. This is associated with the psi "observational theories," which are in turn based on interpretations of quantum mechanics (e.g., Schmidt, 1987, 1993). In the present case, the first observer and first analyst of the data was the first author, who throughout these studies maintained an expectation – based on previous experimental outcomes – that these studies would be successful. Thus, we cannot dismiss the possibility that observer expectations played a role in the outcome.

CONCLUSION

This experiment adds to the empirical database supporting previous laboratory studies indicating that distant healing intention (DHI) is associated with statistically significant fluctuations in spontaneous human autonomic physiology. The study on group healing intention resulted in an effect size for changes in electrodermal activity which was about twice as large as previous meta-

analytic estimates involving individual intention. This suggests that groups may enhance DHI effects.

The second study examined time and space-displaced DHI by spiritual healers located over 6,000 miles from the participants, and time-displaced two months in the future. Again, the effect size for changes in electrodermal activity was nearly twice as large as previous meta-analytic estimates. This suggests that use of professional healers may enhance DHI effects, and that DHI is independent of both space and time.

Finally, detection of effects that apparently transcend space and time highlights a methodological paradox. These studies rely upon conventional techniques such as double-blinds, randomly counterbalanced conditions, and time-synchronized measurements, to study unconventional hypotheses. Overall, the studies provide evidence in favor of the hypotheses. But it is worth nothing that the hypotheses posit phenomena that challenge the very assumptions underlying the methods. That is, for informational, perceptual or energetic effects (however we wish to interpret these) that transcend time and space, the very concept of “double-blind” no longer makes sense. Nor does random counterbalancing, sensory shielding, or precise time-synchronization. In other words, through strict use of standard methodologies, we have apparently transcended the limits of those very methodologies.

ACKNOWLEDGMENTS

These studies were supported by generous grants from the Fundação Bial, Porto, Portugal, and the Society for Psychical Research, London, England. The Consciousness Research Laboratory is also grateful for receiving grants from the Institut Für Grenzgebiete der Psychologie Und Psychohygiene, Freiburg, Germany; the Parapsychology Foundation, New York City, USA; the Bigelow Foundation, Las Vegas, Nevada, USA; and the John Bjorkhem Foundation, Stockholm, Sweden.

APPENDIX A: RANDOMIZED PERMUTATION ANALYSIS

Random permutation analysis (RPA) is not as well known as traditional inferential statistics, but as computing power has increased over the last twenty years, this computational approach to statistics has rapidly grown in popularity. Hundreds of publications in many disciplines demonstrate the advantages of this method over traditional inferential statistics (e.g., Efron, 1983; Noreen, 1989, Manly, 1997). RPA is particularly useful for analyzing the highly skewed, autocorrelated time-series datasets generated in psychophysiological experiments.

To illustrate this method as applied in the present study, consider the average superposed epoch waveforms in Figure 4. This shows the estimates of electrodermal activity (EDA), averaged per second for the treatment and control conditions, across all 15 sessions in the Group experiment. That is, these curves were formed by averaging the 10 treatment and 10 control curves per session, and then averaging those curves over 15 sessions. This represents a total of 15×20 or 300 epochs, of which 150 were treatment epochs and 150 were controls. At 44 seconds per epoch, this means Figure 4 summarizes a grand total of 110 minutes of treatment intention and 110 minutes of control or “non-intention.”

The question is whether the average EDA in the two conditions significantly differed. If the average of the treatment curve in Figure 4 is generally smaller than the average of the control curve, then the cumulative *difference* between the two curves will progressively decrease. This is

shown in Figure 5. A value of special interest is the cumulative endpoint, which is around -110 in Figure 5. We now compare this cumulative difference to the same differences resulting from all possible curves that *could have resulted* if other counterbalancing sequences had been used.

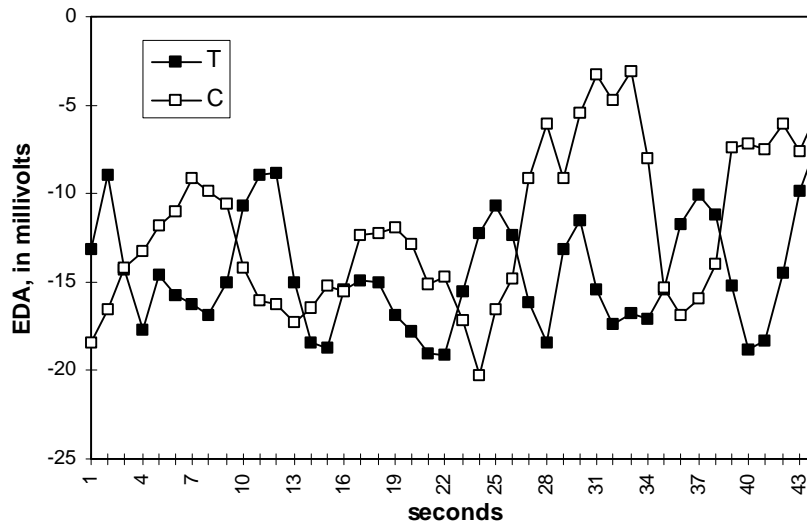


Figure 4. Superposed epoch waveforms for electrodermal activity in the treatment (T) and control (C) conditions, averaged over all epochs in all 15 sessions of the Group experiment. The y-axis is electrodermal activity (spontaneous fluctuations in skin conductance response), measured here in terms of millivolts as recorded by the physiological hardware.

To do this, reconsider the superposed epoch analysis curves shown in Figure 4. These two curves are based on the treatment and control values actually recorded in 15 individual sessions. That is, Figure 4 was formed by taking the averages of 15 *pairs* of curves, where *each* of those curves was itself the average of ten repeated, 44-second sessions. Let us code each of those 15 session-pairs such that each pair is by definition labeled a “1.” We can then code a new pair of curves as “0” by flipping the original treatment with the controls. That is, we simply call the actual treatment curve the control curve, and vice versa.

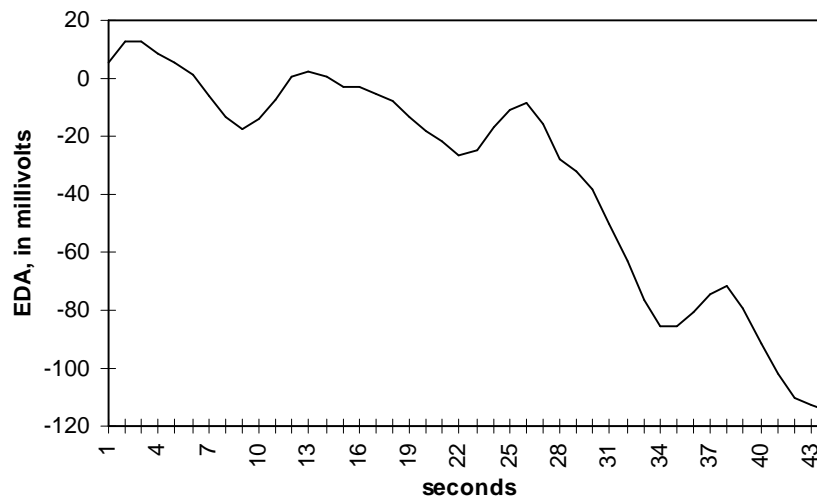


Figure 5. Cumulative difference of the two mean waveforms shown in Figure 4.

With this coding scheme in mind, the original curves shown in Figure 4 may be represented as a series of 15 1’s: [111111111111111]. We can create a new pair of superposed epoch curves by taking a different sequence, say [011111111111111]. This latter example means we reverse the

treatment and control curves originally recorded in session 1, and use the original treatment and control mean curves from the remaining sessions. With this new dataset we can create a new pair of superposed epoch means exactly as we did for Figure 4, and then form a new cumulative difference as in Figure 5. The new cumulative end point will be slightly different than the original.

Now create another new sequence, say [10111...], go through the same process, and get a new cumulative end point. Then do it again, and repeat. In this example, there are 2^{15} possible sequences, or 32,768 permutations of these two curves. If we re-calculated the superposed epoch curves and the cumulative end-points for every one of these permutations, we will have created the population of all possible cumulative differences, on a per-session basis, given this dataset. We can then rank the magnitude of our original cumulative end point within this exhaustive set of 32,768 end points.

Let's say that our end point was ranked the 102nd largest out of the entire set. We would immediately see that the probability of getting a mean difference as large as, or larger than, the actual curve we obtained, was exactly 102/32768, or $p = 0.00305$. That is, these two curves were significantly different at $p = .003$ one-tailed, or $p = .006$ two-tailed.

If we had not 15 pairs of curves, but N curves, then the number of permuted results is 2^N . For experiments with say 30 participants (or 30 sessions), this means for an exhaustive comparison we'd have to calculate 2^{30} or over one billion combinations to find the exact probability of our original results. Even for fast computers that could take a while, so instead of calculating the entire set of possible results, we can run a smaller number of random permutations and judge our results in that subset of the population. It turns out that this permutation method converges fairly quickly to a final estimate of the probability, so that a random set of 10,000 permutations is usually more than enough to form an accurate estimate of the exact probability. For example, in an analysis of the data in the second experiment reported here, Figure 6 shows that the probability estimates begin to stabilize after about 300 random permutations. Nevertheless, following the recommendations of Blair and Karniski (1993), the analyses reported here based on 10,000 random permutations.

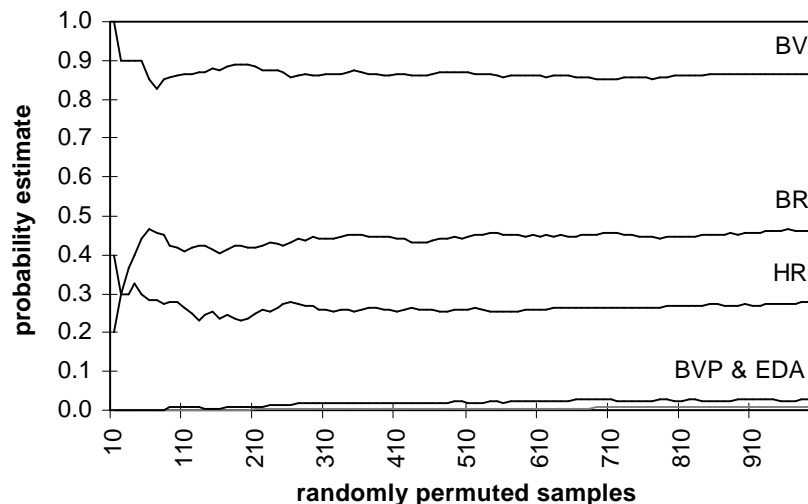


Figure 6. This graph shows how quickly the probability estimates converge towards stable values using the randomized permutation method. This is for data recorded in the Brazil experiment.

REFERENCES

- Bastide, R. (1971) *As religiões africanas no Brasil*. São Paulo: EDUSP.
- Bem, D. J. & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin*, **115** (1), 4-18.
- Benson, H. (1992). *Relaxation response*. New York, NY: Random House.
- Bierman, D. J. (1996). Exploring correlations between local emotional and global emotional events and the behavior of a random number generator. *Journal of Scientific Exploration*, **10**, 363-374.
- Blair, R. C. & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, **30**, 518-524.
- Braud, W. & Schlitz, M. A (1991). Consciousness interactions with remote biological systems: anomalous intentionality effects. *Subtle Energies*, **2**, 1-46.
- Braud, W. G. & Schlitz, M. (1989). A methodology for the objective study of transpersonal imagery. *Journal of Scientific Exploration*, **3** (1), 43-63.
- Braud, W., Davis, G. & Wood, R. (1979). Experiments with Matthew Manning. *Journal of the Society for Psychical Research*, **50** (782), 199-223.
- Braud, W., Shafer, D. & Andrews, S. (1993a). Reactions to an unseen gaze (remote attention): a review, with new data on autonomic staring detection. *Journal of Parapsychology*, **57**, 373-390.
- Braud, W., Shafer, D. & Andrews, S. (1993b). Further studies of autonomic detection of remote staring: replications, new control procedures, and personality correlates. *Journal of Parapsychology*, **57**, 391-409.
- Dunne, B. J. (1993). Co-operator experiments with an REG device. In K. R. Rao (Ed.), *Cultivating consciousness: Enhancing human potential, wellness, and healing*, Westport, CT: Praeger, 149-63.
- Efron, B. (May, 1983). Computer-intensive methods in statistics, *Scientific American*, pp. 116-130.
- Manly, B. F. J. (1997). *Randomization, bootstrap and Monte Carlo methods in biology* (2nd ed.). London, UK: Chapman & Hall.
- May, E. C., Utts, J. M. & Spottiswoode, S. J. P. (1995). Decision augmentation theory: toward a model for anomalous mental phenomena. *Journal of Parapsychology*, **59**, 195-220.
- Nelson, R. D., Bradish, G. J., Dobyns, Y. H., Dunne, B. J. & Jahn, R. G. (1996). FieldREG anomalies in group situations. *Journal of Scientific Exploration*, **10**, 111-142.
- Noreen, E. (1989). *Computer-intensive methods for testing hypotheses*, New York: Wiley.
- Pordeus Jr., I. (1996) Lisboa de caso com a Umbanda. In *Dossiê Magia*. São Paulo: USP.
- Radin, D. I. (1997). *The Conscious Universe*. San Francisco: HarperCollins.
- Radin, D. I., Taylor, R. D. & Braud, W. (1995). Remote mental influence of human electrodermal activity: A pilot replication. *European Journal of Parapsychology*, **11**, 19-34.
- Rosenthal R. (1991). *Meta-Analytic Procedures for Social Research*. Newbury Park, NJ, and London, England: Sage Publications.
- Schlitz, M. & Braud, W. (1997). Distant intentionality and healing: Assessing the evidence. *Alternative Therapies*, **3** (6), 62-73.
- Schmidt, H. (1987). The strange properties of psychokinesis. *Journal of Scientific Exploration*, **1**, 103-118.
- Schmidt, H. (1993). Observation of a psychokinetic effect under highly controlled conditions. *JP*, **57**, 351-372.

¹ Present affiliation: Interval Research Corporation, 1801 Page Mill Road, Bldg C, Palo Alto, CA 94304.

² Rua Vicente Jose de Almeida, 228 - Jd. Cupece - CEP: 04652-140 - Sao Paulo/SP - Brazil.

³ We are indebted to Dr. Sally St. John, Rev. Bonnie James, and the members of the two groups who graciously volunteered their time and efforts for this experiment.

⁴ In particular, it was explained that the purpose of this study was to examine short-term physiological changes in the “patient,” and that no claims for medically relevant healing outcomes were intended or implied by the experimental procedure.

⁵ Written by the first author in Microsoft QuickBasic™.

⁶ Physiology was generally recorded continuously for 24 minutes to provide a few minutes to calibrate and test the equipment before the session, and to start the session with the phone call.

⁷ In actual practice, it turned out that the inter-epoch period lasted 19 seconds instead of 16 seconds because the original program was developed on a faster computer than the laptop used by the healers to generate their condition and timing information. The extra three seconds were probably caused by the laptop’s slower disk access speed and screen display times. This time delay did not affect the final data analysis because the epochs were synchronized to the physiological data stream by the time-of-day timestamps within both datasets.

⁸ A freeware program called NetTime SNTP, by Kip Haas, Tonic Software, 1996, was used to access the Universal Time clock over the Internet.

⁹ EDA activity in this study was suspected to be negative, reflecting a relaxation response, but since we did not predict a specific direction in advance, we compared the absolute value of the observed effect size to the meta-analytic estimate.

¹⁰ We thank Prof. Dr. Dieter Vaitl, Prof. Dr. Alfons Hamm, Dr. Alan Smith, Dr. Michael Cantwell, and Prof. John Gruzelier for their comments.

¹¹ We are deeply indebted to Mãe Rute, founder and owner of the Casa de Caridade Pai Zé Pilintra e Caboclo Pena Branca, a Umbanda terreiro in São Paulo, and to the mediums and other Brazilian healers who participated in this study.

¹² Physiological recordings were generally 24 minutes in length to provide extra pre- and post-session time.

¹³ The actual length of each recording period in this study turned out to be 63 seconds rather than the originally planned 60 seconds. As we discovered in the Group experiment, this was most likely due to the use of a slower computer used in the experiment than the one used to develop the timing program. All analyses took this 63-second session length into account.

¹⁴ If the physiological measures were truly independent, we could use a Bonferroni adjustment to take the five tests into account. However, these measures are interdependent.

FURTHER INVESTIGATION OF UNCONSCIOUS DIFFERENTIAL ANTICIPATORY RESPONSES TO FUTURE EMOTIONS

Dean I. Radin
Consciousness Research Laboratory
University of Nevada
*Las Vegas, Nevada, USA*¹

A double-blind experiment with 50 participants was designed to replicate previous studies indicating that the autonomic nervous system unconsciously responds to future events. Heart rate, fingertip blood volume, and skin conductance level were monitored before, during and after the presentation of randomly selected calm and emotional pictures. Results showed that these combined physiological measures differentially responded in accordance with a “presentiment” (pre-feeling) hypothesis, $p = 0.017$. Analysis of those trials involving the most negative and the most positive pictures, as rated by independent judges, resulted in a larger combined effect, $p = 0.0002$. After stimulus presentation, 38 of the 50 individuals responded in accordance with the expected orienting reflex. As predicted, that group produced a larger combined presentiment effect compared to people who responded in idiosyncratic ways, $p = 0.00004$. Analysis of potential sensory cues and unconscious statistical schemes as alternatives to the presentiment hypothesis revealed no suitable normal candidates that could systematically cause these differential anticipatory responses.

INTRODUCTION

In August 1991, Sheila King sat down at a \$500 slot machine at Caesars Palace casino in Las Vegas. After playing five tokens, she won a \$250,000 jackpot. While waiting for the winning check to be drawn up by the casino, she played another \$500 machine and won \$50,000. A few minutes later, she won another \$50,000 on a third machine. Over the next three years, she continued to win, amassing nearly \$200 million in slot jackpots (Pulley, 1998). This story is so astonishing that we would dismiss it as pure fantasy had it not appeared in the staid *New York Times*.

The question is whether some cases of remarkable luck, sometimes attributed to “intuitive hunches,” may be due to gaining information about future events, specifically *emotional* responses to future events, as slot machine jackpots are likely to induce. To study this question, an experiment was designed to monitor autonomic physiology before, during and after exposure to calm and emotional stimuli, using a randomized double-blind protocol. The rationale for using emotional targets is the observation that the great majority of spontaneous psi experiences, including runs of remarkable luck, are associated with strong emotions (Rhine, 1969; Delanoy, 1989). Another reason for monitoring the autonomic nervous system is that hunches are frequently described as being viscerally “felt” rather than cognitively “known,” suggesting that, like other cognitive and perceptual processes, psi perception occurs unconsciously and can be detected through changes in nervous system functioning.

The hypothesis then is that the autonomic nervous system becomes aroused a few seconds before seeing emotional stimuli, and not as much before seeing calm stimuli. This *differential*

¹ Present affiliation: Interval Research Corporation, Palo Alto, California, USA

anticipatory response is called “presentiment,” and is associated with the well-known “orienting reflex” (Kimmel, Van Olst & Orlebeke, 1979). The orienting effect (OR) is especially suitable for studying presentiment because it reflects an organism’s unconscious “what is it?” response to a novel or meaningful stimulus, OR is associated with momentarily enhanced sensory capabilities after a stimulus (Siddle, Kuiack & Kroese 1983), and the psychophysiological literature on OR and anticipatory arousal provides a solid knowledge base on which to build.

Previous studies on presentiment found the predicted changes in electrodermal activity (EDA, measured as skin conductance level), heart rate (HR) and blood volume pulse (BVP) in the fingertip (Radin, 1997). Since then, a replication by Bierman at the University of Amsterdam, using new stimulus pictures, hardware and software, supported the original findings (Bierman & Radin, 1997, 1998).

The present study was intended to replicate the earlier research using 50 new subjects, and test the expectation that if presentiment genuinely reflects the emotionality of future targets, then by examining only those trials independently rated as the most emotional vs. the most calm, we should see a larger effect.

Previous research

Agent-receiver studies

The idea of using psychophysiological responses as detectors of psi performance is not new. Clever variations have been tried over the years, and many studies have provided evidence suggesting that psi effects may be easier to detect with unconscious measures than with conscious responses. This is possibly because unconscious psi information is not blocked by psychological defense mechanisms. Most of these studies used agent-receiver paradigms, where an agent tried to remotely affect a receiver.

For example, Tart (1963) found that the electroencephalographs (EEG), EDA and fingertip blood volume in isolated receivers were more active when he, as the agent, received electric shocks. Duane & Behrendt (1965) found that increased alpha in one identical twin’s EEG was suggestively related to increased alpha in the other, isolated twin. Targ & Puthoff (1974) examined receivers’ EEG while flashing a strobe light in a distant agents’ eyes. The average power of the receiver’s occipital EEG alpha significantly decreased, as predicted. Later, Kelly & Lenz (1976) and May, Targ & Puthoff (1979) tried replicating these effects and both groups obtained suggestive or significant evidence that aspects of the EEG frequency spectrum responded differentially to the remote stimulus and control periods. In related studies using EEG evoked potentials, Lloyd (1973) found evidence suggesting that EEG evoked potentials could be used to discriminate between stimulus and non-stimulus conditions. Millar (1976) could not replicate Lloyd’s findings, but later replication efforts by Warren, McDonough and Don (1992) and Grinberg-Zilberbaum et al (1992) were successful.

Similar agent-receiver studies, using EDA or vascular (typically blood volume) responses instead of EEG, have been reported by researchers from the 1950s to the present (cf. Radin, 1997).²

² E.g., Hettinger (1952), Matrick (1955), Figar (1959), Dean (1962, 1966), Rice (1966), Esser, Etter & Chamberlain (1967), Barron & Mordkoff (1968), Nash & Nash (1969), Sanjar (1969), Beloff, Cowles & Bate (1970), Schouten (1976), Haraldsson (1972, 1980); Braud (1981, 1993), Braud & Schlitz (1989, 1991), Braud,

Many of those studies provided evidence that remote agents affected the receivers when the target stimuli were emotionally strong. In virtually all of these studies, receivers could not systematically report when the stimuli took place with any accuracy.

Psychophysiological studies of precognition

Of more direct relevance to the present approach, Levin & Kennedy (1975) used a reaction time task to see whether contingent negative variation (CNV), a slow brainwave indicator of anticipation, could be used to differentially detect a future target. Participants were asked to press a key when a green light appeared (in the future), but not when a red light appeared. A random number generator determined which light would appear. Levin & Kennedy found significantly larger CNV's just before the RNG selected a green light as compared to just before it selected a red light. A few years later, Hartwell (1978) reported a similar study using CNV. Participants saw a picture of a person of the same sex or the opposite sex. If the picture was of the opposite sex, they pressed a button as quickly as possible, otherwise no response was made. Various forms of warning stimuli were used to alert the subject that a picture was about to be shown, from explicit warnings to prepare to respond, to no warning which served as the precognition condition. Thirteen of 19 planned statistical tests were in the predicted direction, but overall the results were not significant. In hindsight, it seems that these CNV experiments were on the right track, but perhaps green/red lights and same/opposite sex stimuli were insufficiently motivating to generate robust effects.

In a similar approach, Klintman (1983, 1984) studied a Stroop task where a person was shown a patch of color followed by the name of a color. Klintman asked the person to speak aloud the name of the patch of color as quickly as possible, and then speak aloud the word that followed, again as quickly as possible. Klintman was mainly interested in studying the reaction time to the second stimulus, but on a whim he decided to measure the time it took to speak aloud the color of the *first* stimulus, the color patch. He reasoned that he might be able to use the first reaction time as a baseline to form a more sensitive measure of the second reaction time. He was surprised to find that the first reaction times had greater variance than expected. He investigated further and discovered that the initial reaction times were faster when the color patch and color name matched, and slower when the following color name mismatched. After considering conventional explanations for this effect, Klintman decided to test the possibility of what he called "time reversed interference." By this he meant that precognitive sensing of a future stimulus somehow caused cognitive interference when the future stimulus was a mismatch. After five experiments, each study using a somewhat different design, the combined result for all experiments was $p \approx 2 \times 10^{-6}$.

More recently, Bechara, Damasio, Tranel & Damasio (1997) reported an experiment in *Science* studying decision-making aspects of intuitive hunches. Among other factors, they measured skin conductance response (SCR) in 10 normal people and in 6 patients with prefrontal brain damage while they individually played a gambling game using cards. The game involved four card decks and a task to select a deck and turn over a card, one at a time. Some cards would win money, others would lose money. Two of the decks were "bad" in that they had a higher percentage of losing cards. Selecting cards from these two decks would eventually lead to overall losses, whereas selections from the other two decks would lead to overall winnings. Without knowing which were the good or bad decks, or the distribution of losing cards in a deck, or how many

cards they had to select to end the game, Bechara et al reported that the normal people “began to generate anticipatory SCRs whenever they pondered a choice that turned out to be risky, before they knew explicitly that it was a risky choice.” In attempting to explain this effect, Bechara et al speculated that

The bias mechanism identified here is distinct from other neural mechanisms whose integrity is crucial for decision-making.... In other words, we propose an addition to mechanisms already recognized as necessary for proper reasoning rather than an alternative to those mechanisms. (p. 1295)

Bechara et al did not propose what that additional mechanism might be, but their results showed that even before a single gambling loss took place, and before the individuals were consciously aware of their decision-making strategies, both normal and brain-damaged people showed larger anticipatory SCRs when they unknowingly selected the bad decks than when they selected the good decks (Bierman, 1997). This task, and the psychophysiological results, provides an unexpected conceptual replication of the presentiment hypothesis.

METHOD

Subjects

Most participants in this study were volunteers recruited from staff, faculty and students at the University of Nevada, Las Vegas, and from visitors to the Consciousness Research Laboratory. A few volunteers were recruited from the local community through advertisements seeking people interested in a study on intuitive hunches. There were 23 men and 27 women, ages ranging from 19 to 79, with an average age of 41.

Procedure

Participants first read an informed consent form, which explained that they would sit quietly and passively observe a series of randomly selected color photos while aspects of their physiology were continuously monitored. It was explained that the photo content would range from calm to highly emotional; nothing more was asked of the participants. If a volunteer agreed to participate, after filling out the paper forms, he or she was seated in front of a computer monitor, and electrodes were attached to his or her first three fingers of the left hand. These included silver-silver chloride electrodes to the index and second finger,³ and a photoplethysmograph (PPG) to the tip of the third finger. The electrodermal electrodes measured skin conductance level (SCL), and the PPG measured both heart rate and blood volume.

The experimenter started a software program which calibrated and tested a computer-controlled physiological monitoring system (J&J Electronics, Model I-330, analog-to-digital conversion with 12-bit resolution). The program collected data at 5 samples per second, and initially displayed the output of the I-330 system on-screen so the experimenter could judge whether the monitor was operating properly.

If everything was working as expected, then at the experimenter’s prompting, the participant pressed a mouse button with the right index finger. The computer screen remained blank for five

³ As recommended by Venables & Christie (1973), p. 108.

seconds, then a colorful picture was shown for three seconds in 600 x 800 resolution, followed by a blank screen for ten seconds, as shown in Figure 1. At the end of the 10 second “after-stimulus” period, a message appeared stating that when ready, the participant could press the mouse button to begin the next trial. At this time, the experimenter explained that during these inter-trial periods, if the participant needed to stretch, scratch, or move, they should do so and then relax a few more seconds to settle back down. If the participant indicated that they understood the procedure, the experimenter retired to the other side of the lab room, behind a screen, and the participant conducted 40 trials at will and unobserved. To enhance the contrast of the stimulus pictures during the experiment, and reduce electrical interference, the overhead fluorescent lights were turned off and a small red incandescent lamp was turned on.

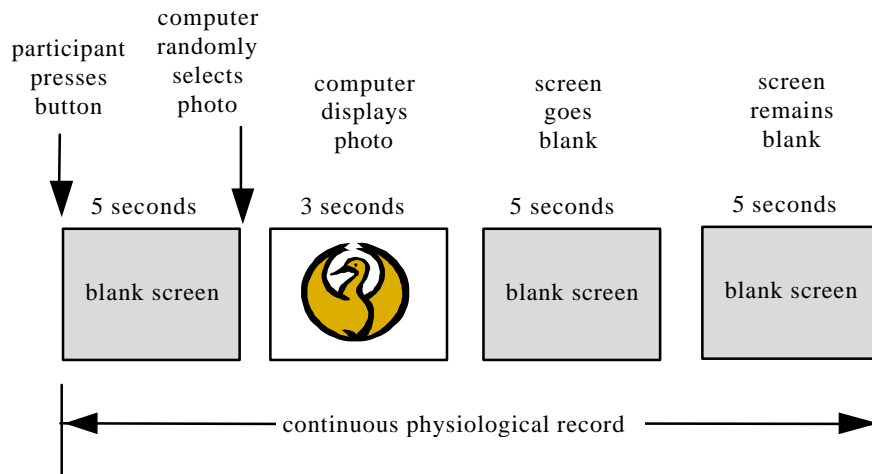


Figure 1. Illustration of experimental procedure.

Target selection

When the participant pressed the button, the software reset the seed-number of the controlling program’s pseudorandom number generator (PRNG) algorithm to the present clock time, and the physiological monitor began to continuously collect data at 5 samples per second.⁴ Immediately after the 25th sample, the software created a new seed-number by adding the outputs of the EDA and BVP values recorded on the 25th sample to the computer’s current clock time. The newly seeded PRNG was then used to randomly generate a number from 1 to 150, which determined the target stimulus to display. The selected target was immediately retrieved from the hard disk and displayed for 3 seconds. (Retrieving and displaying the picture took about 0.1 second).

Note that through this procedure, from the moment a trial was initiated with a button press to just before the stimulus was displayed, the target *was not yet determined*. There were no sounds (due say, to movements by the computer’s hard disk as it retrieved the digitized target), or electromagnetic changes in the computer monitor display, or any other hints that might have provided a sensory cue about the identity of the upcoming target.

In addition, even though participants did not know the numbers of calm and emotional targets, or the size of the target pool, the pictures were randomly selected *with replacement* to prevent statistical hints about the future targets from accumulating over the course of the experiment. To avoid repeating a stimulus, if a randomly selected target had been used before, the selected

⁴ Written in Microsoft QuickBasic 4.5 by the author.

target's *class* (i.e., emotional or calm, as described below) was retained and another picture from that same class was randomly selected and displayed. In this way, the initial probability of seeing a calm or emotional target was held constant throughout the experiment, and no targets were repeated.

Target content

The experiment reported in Radin (1997) involved a pool of 120 pictures: two-thirds calm and one-third emotional. The target pool for the present experiment was expanded to 150 pictures: 100 calm and 50 emotional. This one-third vs. two-thirds proportion was used to avoid the physiological habituation expected to occur if participants saw too many emotional stimuli in the same session.

All of the calm pictures were selected from a collection of high resolution color digital photographs,⁵ and with two exceptions, all of the emotional pictures were retrieved from publicly accessible photo archives on the Internet. The two exceptions were photos from adult magazines. The calm pictures included beautiful landscapes, seascapes, and portraits of people, fruit, trees or animals. The emotional pictures portrayed a range of erotic and violent topics, following the style and content of the *International Affective Picture System* (IAPS, Bradley, Greenwald & Hamm, 1993; Hamm & Vaitl, 1993), a standardized set of photos developed for psychophysiological studies of emotion.

To provide an independent subjective assessment of our target set, five men and five women were asked to independently view a randomly selected group of 50 target pictures from the pool of 150, one picture at a time, on two dimensions: emotionality (level of arousal) and valence (negative to positive). The rating dimensions each consisted of 100 points, and the rating method asked the person to view the picture on a computer screen and move a pointer across a continuous scale to indicate their assessment for each of the two dimensions. Each individual's assessments were normalized into z-scores, and the z-scores per target were combined across all 10 people to provide a two-dimensional distribution of target emotionality/valence in z-score space. The outcome, shown in Figure 2, resulted in a high correlation between valence and emotionality ($r = 0.89$). This picture set differs from the IAPS in that it contained no pictures rated as both highly emotional and highly positive.

To demonstrate that the targets were partitioned into two distinctly different categories, Figure 3 graphs the mean subjective ratings for emotionality and valence in the two target categories of 100 calm and 50 emotional targets. There is virtually no overlap in mean ratings for either emotionality or valence in the two target classes.

⁵ Corel Professional Photos CD-ROM.

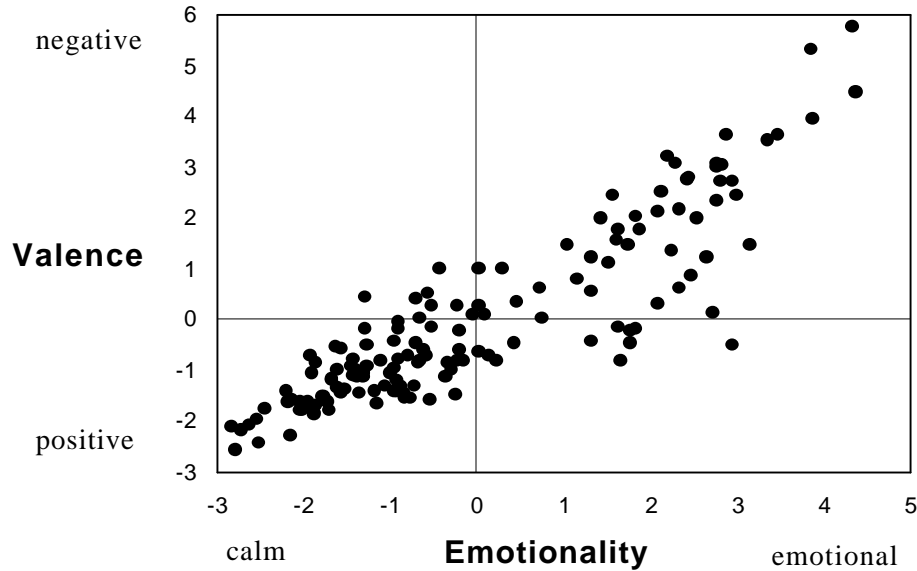


Figure 2: Subjective assessment of targets in z-score space by 10 independent judges.

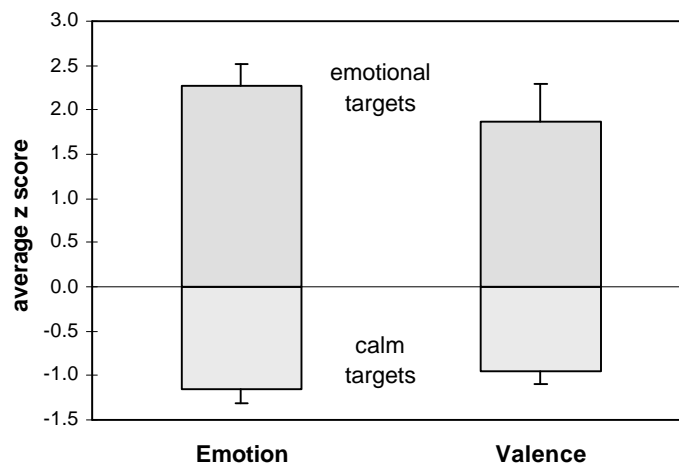


Figure 3: Means for subjective ratings of targets in the calm and emotional categories, with 95% error bars.

Hypotheses

H1: The autonomic nervous system will show a greater degree of anticipatory arousal before exposure to emotional targets than before calm targets.

To explore H1, data contributed by individuals in a single session and data pooled across all participants, were examined using a superposed epoch analysis (described later). These analyses were useful in visualizing the outcomes, but were not suitable for making formal assessments about the statistical significance of the physiological differences between the two target conditions. To determine the significance of individual and pooled trials, we used a randomized resampling or perturbation technique to determine probabilities free of distributional assumptions

and independent of autocorrelation dependencies between adjacent samples and within individuals.⁶

The probabilities produced by the resampling method were transformed for each physiological measure per individual into a standard normal deviate (z score). These z scores were combined across all participants to provide an overall statistical test for each physiological measurement. Finally, a single, “systemic” z score was created by combining the three physiological measures according to the expected anticipatory effect observed in earlier presentiment research. That is, the hypothesis predicts that the autonomic nervous system will show a systemic, differential “presponse” to future emotional stimuli, thus any one, or group of autonomic measures we care to monitor would reflect the predicted presentiment effects. Because we monitored three autonomic measures, this prediction was expressed mathematically as $Z = (Z_{EDA} - Z_{BVP} + Z_{HR})/\sqrt{3}$.⁷ This combined z score provided a single measure of presentiment per person, and a way of assessing overall statistical significance for all participants combined.

Due to high levels of heterogeneity expected among different peoples’ responses to emotional stimuli (Claridge, 1967, p. 27), we did not expect to obtain a robust effect upon pooling data from all subjects. Rather, as stated more explicitly in the following hypotheses, we expected that presentiment would be stronger under certain conditions.

H2: The greater the target emotionality and valence, the greater the presentiment effect.

This is based on the simple assumption that we are affected most by what attracts our attention. Target photos that force our attention, motivation and interest – predictably photos of sex and violence – will produce larger effects than pictures of sunsets at the beach and mountains. The first method used to test H2 examined data from those trials where the targets, subjectively rated by independent judges, were the top ten negative valences and top ten positive valences. The same analytic methods used to test H1 were applied to this “high contrast” subset with the prediction that the high contrast results would show larger effects than all data pooled.

The second method examined the correlation between subjectively rated target emotionality and the physiological measures recorded before, during and after presentation of those targets, per trial. These correlations were expected to be fairly large and in predictable directions after stimulus presentation, reflecting what is already understood about physiological responses to emotional targets. However, the presentiment hypothesis predicted that these correlations would also be non-zero, and in predictable directions *before* the targets were seen.

H3: Physiological responses after seeing the target will be related to physiological responses before seeing the target.

This differs from H2 in that here we are interested in seeing what happens when participants are segregated into “optimal” and “non-optimal” responders. That is, by taking into account idiosyncratic response styles, and by creating two groups of people according to how they respond *after* seeing stimuli, we should be able to reduce the heterogeneity among peoples’ physiological responses, and thereby increase our ability to detect presentiment effects.

⁶ As recommended by Blair & Karniski (1993) for evaluating psychophysiological waveform differences.

⁷ Z for BVP is negative because presentiment predicts that blood volume in the fingertip will drop before exposure to emotional targets.

To study this hypothesis, individuals were classified as *optimal*, meaning they responded differentially after seeing the targets in ways that are consistent with the classical orienting reflex, or as *sub-optimal*, meaning their responses were flat or contrary to expectation. The combined statistical results of the optimal individuals were compared to that of the sub-optimal individuals, with the prediction that the optimal group would show larger presentiment effects.

RESULTS

Participants

From April 1996 through March 1997, data was obtained from 50 volunteers, all of whom used the same hardware, software, and target stimulus set. Of the 50 participants, 38 were run through the experiment by a research assistant⁸ and 12 by the author. Pilot data from four people were collected in April and May 1996, and the bulk of the remaining data collected between October 1996 and March 1997. A total of 51 people participated in the study, but hardware irregularities occurred during data collection with three people, forcing one dataset to be discarded and resulting in only 19 out of 80 valid trials in the remaining two datasets. Thus, rather than collecting the planned 2,000 trials (40 trials \times 50 people), a total of 1,939 trials were collected in which 1,283 targets were calm and 656 targets were emotional.

Distribution of targets

It was especially important to ensure that the sequence of calm and emotional targets was unpredictable, because any consciously or unconsciously detected regularities could influence participants' expectations, which would drive physiological anticipation. To study whether the obtained sequence was unpredictable, we took the obtained sequence of 1,939 calm and emotional targets generated during the experiment across all participants, and coded them as 0 for calm and 1 for emotional. This formed a binary array with 1,939 elements. Figure 4 graphs the lengths between successive 1's in this array along with the expected sequence lengths based on $p(1) = 1/3$ and $p(0) = 2/3$. This shows that the PRNG used to select the targets was adequately random, and that the sequence of actually used targets provided no statistical hints about the identity of the upcoming pictures.

However, this distribution also shows that the likelihood of obtaining progressively longer sequence lengths rapidly decreases. This may lead to the gambler's fallacy, in which a participant might feel that since the last three or four targets were say, calm, then surely the *next* target would be emotional. This might result in progressively increasing anticipation about each successive target until an emotional picture appeared, whereupon the "statistical counter" in the person's mind would reset to zero. Such an anticipatory strategy can be simulated to see if it would produce results that mimic the postulated presentiment effect. In the computer simulation, instead of increasing anticipation linearly according to the sequence length, we increased it geometrically to reflect the expectation that one's nervous anticipation might increase more as the sequence length increased.

⁸ I thank Jannine Rebman for her assistance.

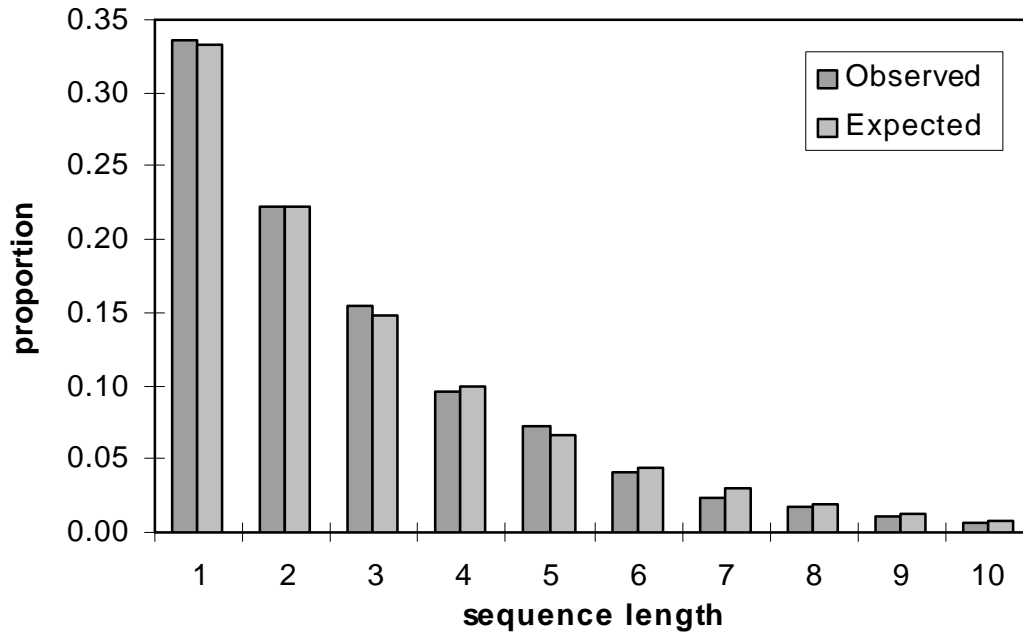


Figure 4. Sequential distribution of emotional targets.

Mean calm and emotional activation levels resulting from the simulation are graphed in Figure 5. Despite the fact that calm and emotional targets occurred with different frequencies, the resulting means are indistinguishable. This indicates that even a consistently applied counting strategy would not generate a significant difference in physiological arousal. Of course, on reflection this should be expected because otherwise the gambler’s fallacy would not be a fallacy. That is, if conscious or unconscious counting strategies really could be used to out-guess a random sequence, then one could eventually win at roulette simply by taking its past history into account.

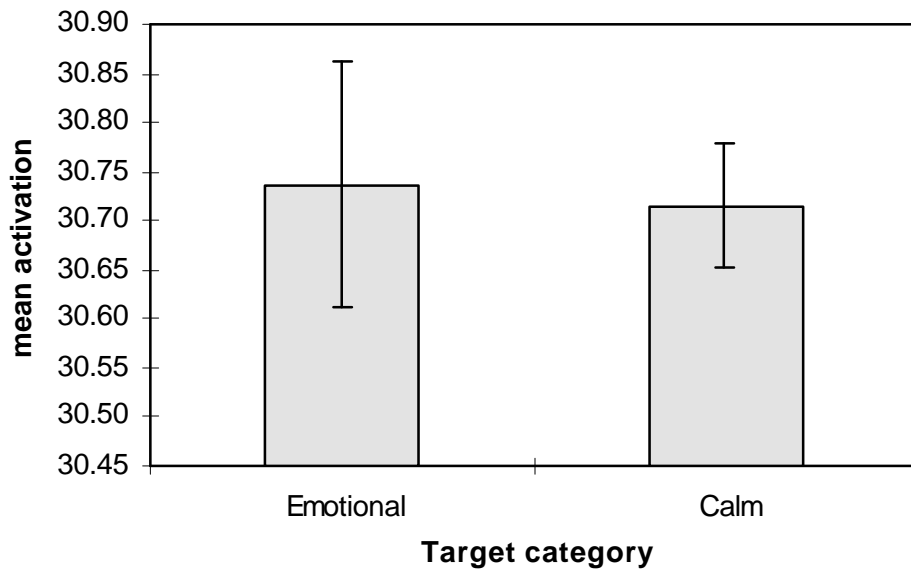


Figure 5. Results of anticipation simulation as mean activation for emotional and calm targets, and one standard error bars. The smaller standard error for the calm mean reflects the fact that a larger number of trials contributed to that mean estimate (two-thirds of the trials). Another sequence that can be examined is the subjectively rated *valences* among successive

targets.⁹ If this sequence were structured in any non-random way, it might provide hints about the upcoming targets. To examine this sequence, autocorrelations up to lag 15 were determined. These are shown in Table 1, along with their associated probabilities. None of the autocorrelations are significant, indicating no unusual sequential structure among successive target valences.

<i>Lag</i>	<i>Autocorrelation</i>	<i>p</i>
1	-0.010	0.648
2	-0.014	0.747
3	-0.012	0.838
4	-0.001	0.932
5	-0.003	0.972
6	-0.015	0.970
7	0.047	0.580
8	0.002	0.684
9	-0.058	0.201
10	0.029	0.182
11	0.007	0.239
12	-0.021	0.253
13	-0.037	0.177
14	0.014	0.213
15	0.008	0.263

Table 1. Autocorrelations of subjectively rated target valences in the observed sequence.

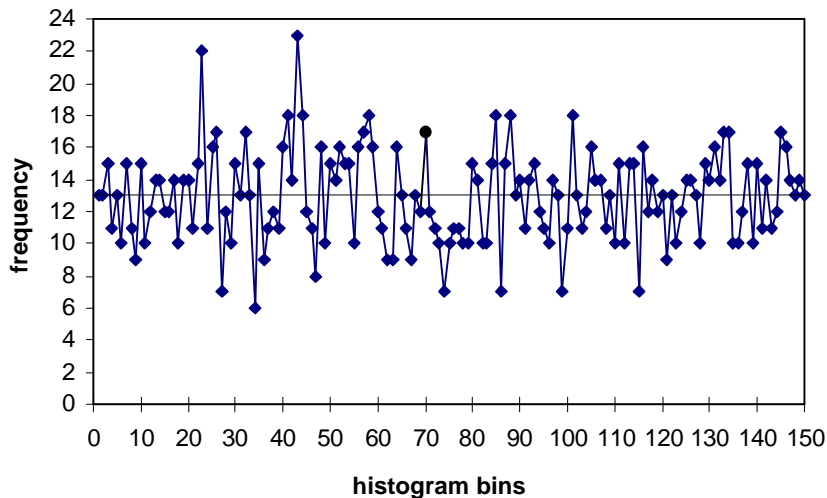


Figure 6. Distribution of targets used in the experiment.

We can further examine whether the 150 target pictures were selected uniformly at random. Figure 6 plots these frequencies. A chi-squared test of equal distribution results in $\chi^2 = 97.5$, 150 df, $p = 0.99$, indicating that the targets were uniformly distributed by the PRNG close to chance expectation. Having established that both the target sequences and their overall distributions were adequately random, and that no ordinary guessing strategies could outwit the sequence, we are

⁹ Recall that subjective ratings for target emotionality and valence were closely correlated in our target pool.

now prepared to examine the physiological results.

Hypothesis 1

H1 states that autonomic measures will show a greater degree of arousal before exposure to emotional targets than before calm targets. As a first step towards evaluating this hypothesis, Figure 7 summarizes the results of the EDA (electrodermal activity), BVP (blood volume pulse) and HR (heart rate) measurements for all trials generated by all 50 participants in the form of a superposed epoch analysis. These graphs were produced in the following way: For each trial, three continuous sequences of 90 12-bit numbers (0 to 4096) were output by the physiological monitor at a rate of 5 per second, one sequence each for EDA, BVP and HR. Due to limitations of the hardware, the first two samples returned by the monitor were unstable, so they were ignored from further analysis.

To take into account physiological drifts from one recording epoch to the next, the average of samples 3 to 25 (i.e., before stimulus display) was determined, and this mean was subtracted from samples 3 to 90. Then a grand mean of the resulting mean-difference values was formed for all calm trials, and similarly for all emotional trials. Finally, the *differences* between sample 3 value and the remaining values were taken for the grand mean calm, and grand mean emotional values. This formed two curves indicating how mean physiology *changed* from the moment the participant pressed the button to begin a trial. These grand mean-difference curves are shown in Figure 7.

The ordinates in these graphs are based on the actual values returned by the physiological monitor rather than values of conductance, beats per minute, and blood volume. The integers can be translated into physiological measures with appropriate constants, but since such a translation would not alter any of the following analyses, for the sake of simplicity the original values were retained.

The following observations can now be drawn: Independent of the future target category, the mean EDA, BVP and HR responses fluctuated in ways that are consistent with the expected orienting reflex in a two-stimulus psychophysiological experiment (e.g., Bohlin & Kjellberg, 1979.) In particular, the results observed before the stimulus are consistent with the type of responses expected for what Sokolov (1963, p. 163) called a *signal stimulus*: “Those stimuli which evoke a reaction in anticipation of external agents likely to appear in the future.” For example, as reported in other psychophysiological studies, when phobic subjects are monitored in anticipation of seeing phobic or neutral stimuli, and they know what they are about to see, their heart rate significantly accelerates more before they see the phobic stimuli than before they see the normal stimuli (Bohlin & Kjellberg, 1979, 185). In the present experiment, we see a greater heart rate acceleration before emotional pictures than before calm pictures, but, of course, these participants did not know what was about to be shown.

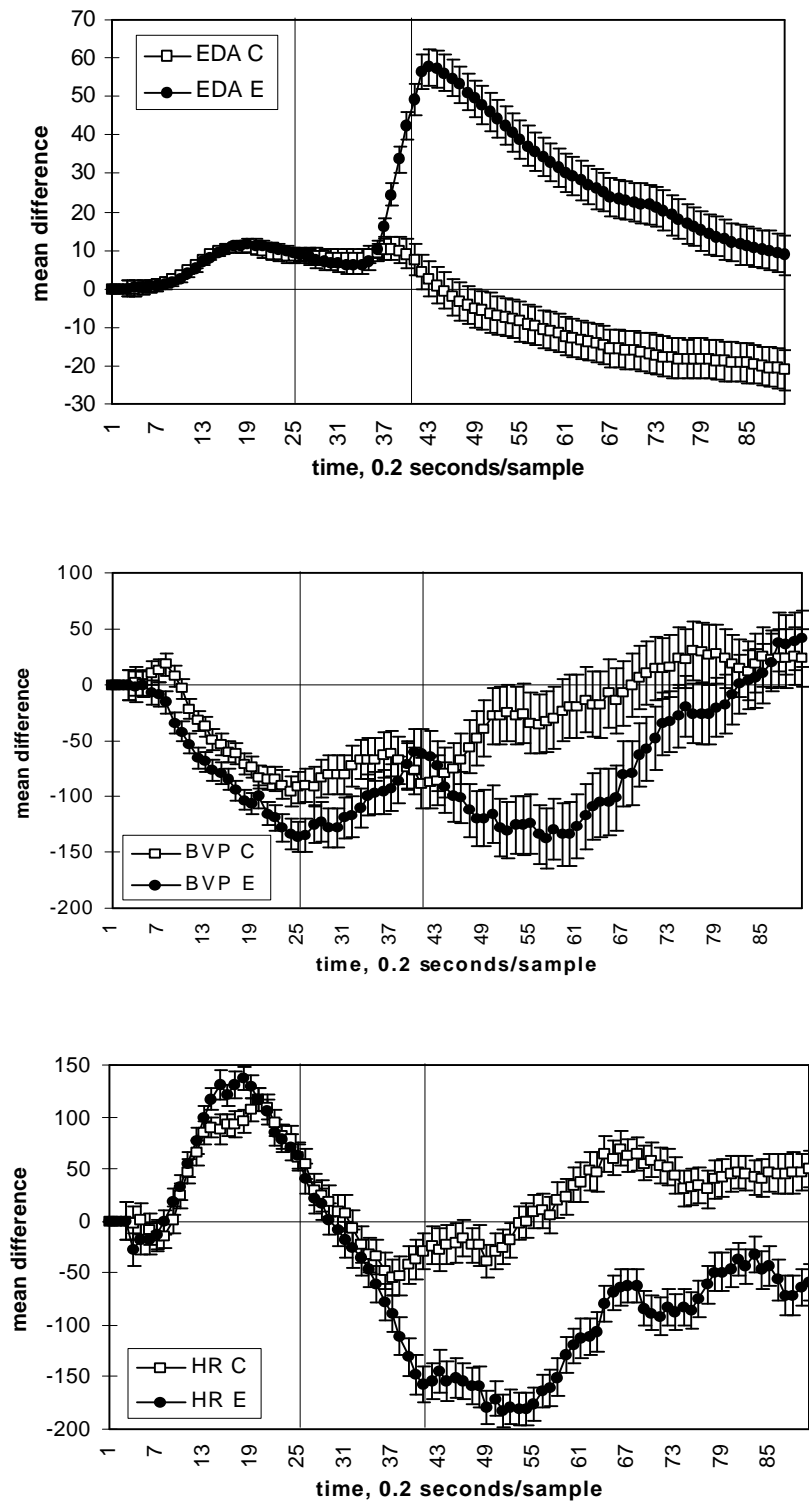


Figure 7. Superposed epoch analyses for mean EDA, BVP and HR, all trials, all 50 subjects, with one standard error bars. The three segments in these graphs represent *before* (samples 1 - 25), *during* (samples 25 - 40) and *after* (samples 41 - 90) display of the target picture.

Now, to create a distribution-free estimate of statistical significance, we used a randomized resampling method, recommended by Blair & Karniski (1993) for use in evaluating physiological difference waveforms. To do this, we took the sum of the observed physiological differences

between emotional targets and calm targets over the last 22 samples of the before-stimulus period, then compared that summed difference with the distribution of similar summed-differences formed by randomly permuting the original emotional/calm target assignments.

To illustrate this method, say that one individual’s experimental session resulted in a randomly ordered presentation of 30 calm and 10 emotional targets. For each trial, we determine the difference between sample 3 and the remaining samples. The shape of the resulting difference curve is identical to the shape based on the actual values, except that the first value is clamped to zero. Then we form a superposed epoch mean of the 30 zero-offset calm trials, and form a similar mean for the 10 zero-offset emotional trials. Each epoch mean curve now consists of 88 individual means (samples 3 to 90). We take the difference between mean calm sample 3 and mean emotional sample 3, the difference between mean calm sample 4 and mean emotional sample 4, and so on, up to sample 25, and then sum those mean-differences. The resulting *summed-mean-difference* score becomes our single measure of presentiment performance for this individual, and this particular physiological measure, in this session.

Now we scramble the original target assignments to create 30 new “pseudo-calm” and 10 new “pseudo-emotional” trials. That is, the original data stays the same, but the *assignments* of which trials were calm and which were emotional are randomized. We evaluate the data as before, finding summed-mean-difference values, and repeat the process say, a thousand times, building up a distribution of randomized resampled summed-mean-differences. We now sort this distribution from smallest to largest summed-mean-differences, and see where our original value resides in this distribution.

If our original value is greater than 950 of the samples in the randomly resampled distribution, then we know that even taking into account a non-normal underlying distribution and autoregressive dependencies among adjacent samples, that our original summed-mean-difference was greater than 95% of the empirical distribution of similarly-created scores. That is, we can assign $p = .05$ to our original summed-mean-difference score. The probability resulting from this resampling method is then transformed into a one-tailed standard normal deviate (z score).

Applying this resampling procedure to all 50 sessions resulted in 50 sets of three z-scores, one z-score for each physiological measure. We then formed a single Stouffer z score to combine the 50 z-scores for each physiological measure, and combined the three Stouffer z’s into a single, “systemic” Z score reflecting how the *differences* in EDA, BVP and HR were expected to behave during the before-display period, i.e., as $Z = (Z_{EDA} + Z_{HR} - Z_{BVP})/\sqrt{3}$. In Table 2, we see that the overall systemic z score for presentiment was $p = 0.017$. All three physiological measures went in the direction predicted by a differential anticipatory effect, but of the three, only BVP was independently significant.

Condition	BVP	EDA	HR	Combined	N	ES
z scores	-2.16	0.61	0.92	2.13	50	0.30
p (one tail)	0.015	0.272	0.179	0.017		

Table 2. Z and p values for physiology measures, and combined, for all data. ES is effect size, calculated as z/\sqrt{N} .

Hypothesis 2

H2 states that presentiment is related to the strength of the future emotion. We predicted that the effect size of the presentiment effect would be enhanced by comparing the physiological differences between the most negative and most positive targets.¹⁰ To do this, we rank-ordered the subjectively-rated valence for all 150 targets to select out the 10 most positive and 10 most negative targets. That is, we selected 20/150 or 13% of the original targets for analysis, about half calm and half emotional.¹¹ The superposed epoch analysis of this data, consisting of 278 trials, 147 calm and 133 emotional, is shown in Figure 8.

Examination of these graphs shows that the differential anticipatory effect is clearer despite a reduction in sample size. From Table 3, we see that the resampling method applied to these high contrast trials for all 50 datasets resulted in an overall (i.e., systemic) presentiment $z = 3.52$, $p = 0.0002$, and the effect size per subject increased from $e = .30$ (Table 2) with all targets to $e = .52$ for the high contrast targets. In Table 3 note that only 45 of the 50 datasets could be subjected to the resampling method because 5 participants did not observe at least one each of the top 10-rated high and low valence targets.

Measure	BVP	EDA	HR	Combined	N	ES
z	-1.77	1.48	2.85	3.52	45	0.52
p (one-tail)	0.038	0.069	0.002	0.0002		

Table 3. Results of applying the randomized resampling method to the high contrast trials from individual sessions.

Hypothesis 3

H3 states that the response after seeing a target will be related to the pre-presentation before, with focus on idiosyncratic response styles rather than on differences in target emotionality and valence. To test this prediction, we separated people into “optimal” and “sub-optimal” responders by examining how each person responded *after* seeing the calm and emotional targets. To do this, we rank-ordered each individual according to the magnitude of their systemic (combined) orienting responses for BVP, EDA, and HR, and labeled “optimal” those who produced a positive systemic response after exposure to the stimulus, and as “sub-optimal” those who produced a negative or flat response. Examination of the per-person systemic response data indicated that 32 people could be classified as optimal responders and 18 as sub-optimal. Blocking individuals’ data by this criterion, as shown in Table 4, shows that optimal responders did in fact show a larger presentiment result as compared to the sub-optimal responders, with z (difference) = 4.023, $p = 2.86 \times 10^{-5}$.

¹⁰ We equate emotional with negative, and calm with positive in this target pool because of the $r \sim 0.9$ correlation between the two subjective ratings.

¹¹ That is, we selected 20% of the emotional targets (10/50) and 10% of the calm targets (10/100) out of a pool containing twice as many calm trials as emotional trials, thus the expected number of targets in this subset of high-contrast targets would be the same.

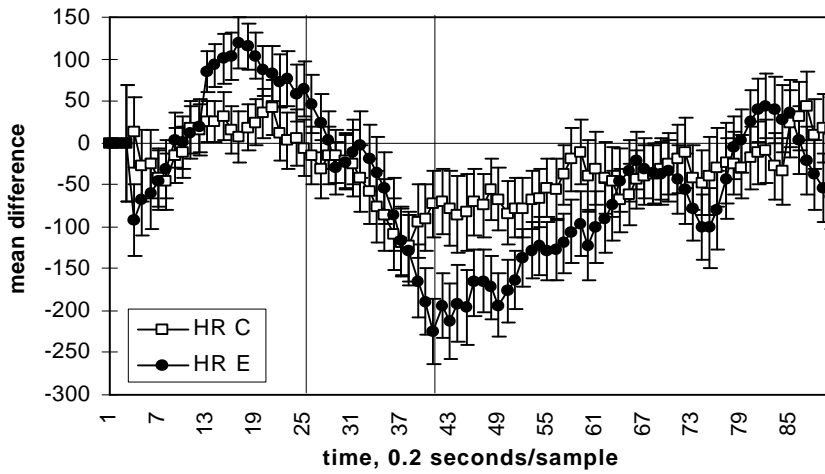
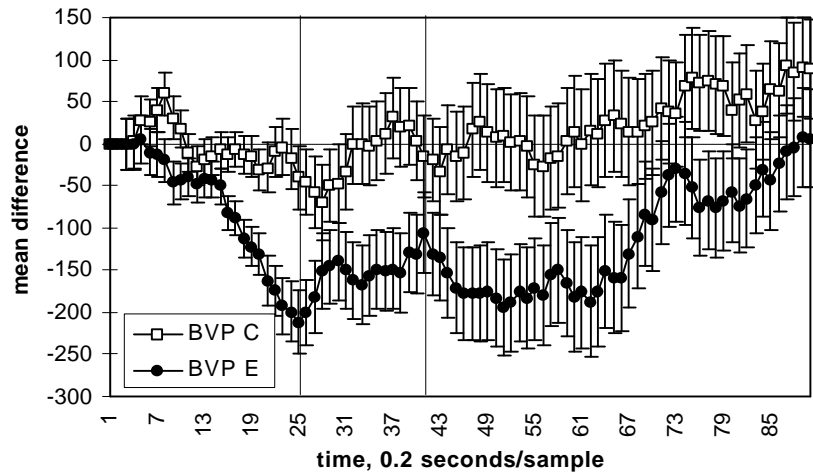
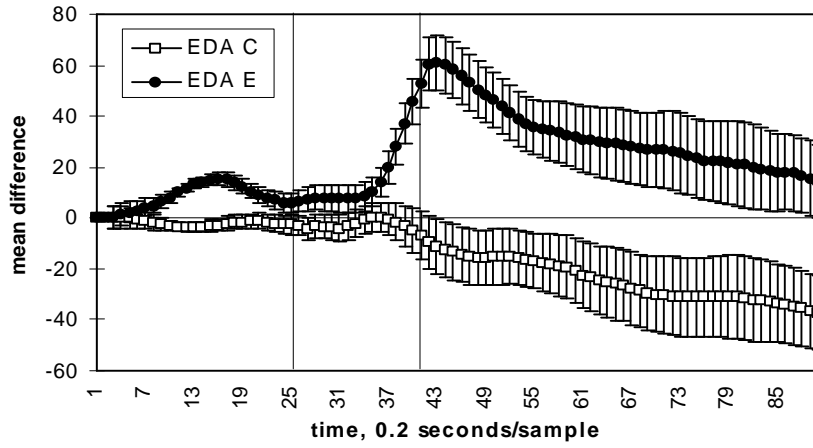


Figure 8. EDA, BVP and HR superposed epoch analyses for high contrast targets.

Condition	BVP	EDA z	HR z	Combined	Combined	pN	ES
-----------	-----	-------	------	----------	----------	----	----

Anticipatory responses to future emotions

	z			z score	(one-tail)		
Sub Optimal	0.68	-1.20	-1.11	-1.73	0.958	18	-0.41
Optimal	-3.21	1.66	1.98	3.96	0.00004	32	0.70

Table 4. Results of optimal and sub-optimal individuals.

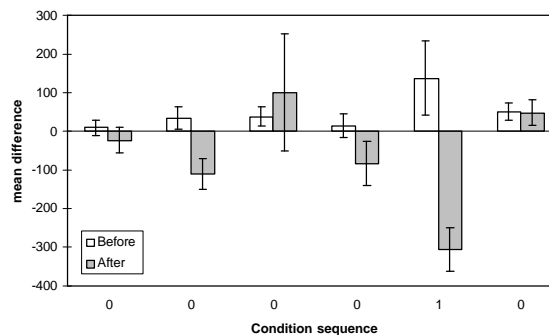
DISCUSSION

The present study, combined with the results of past replications and conceptual similar experiments, indicates that the human autonomic nervous system unconsciously responds to future emotional events. The magnitude of the presentiment effects were predicted to be associated with the strength of the future emotion, and stronger in individuals who responded in alignment with the orienting response. These predictions were confirmed.

Two potential flaws were identified that might result in spurious results. The first was inadvertent sensory cueing about the upcoming target, and the second was a statistical anticipatory effect which took advantage of the participants' progressive realization that there were more calm targets than emotional targets.

Sensory cueing was eliminated as a viable explanation because the target identity was not selected until 20 msec before it was shown. Thus, no sensory cues were possible. Most unconscious anticipatory strategies can be eliminated by referring to the gambler's fallacy, as discussed earlier, and by showing that the distribution and sequences of targets were indeed random. However, these arguments were not based on participants' actual physiological performance. Thus, to examine this possibility in more detail (post-hoc), we examined the average physiological responses in a sequence of six successive trials.

Figure 9 shows the before and after HR, BVP and EDA means for all instances (N = 74) where the 50 individuals obtained sequences of four calm (C) targets, an emotional target (E), and a calm target, i.e., a CCCCEC sequence. A progressive arousal strategy would predict that HR before the series of four calm targets would progressively increase until the emotional target was observed, peak at its highest level at the emotional target, and then drop for the remaining calm target. But the means shown in Figure 10 show no such systematic pattern. The largest HR mean occurred just before the emotional target. Similarly, BVP did not show progressively lower dips before the first four calm targets, nor did EDA show progressively larger increases. This suggests that either conscious or unconscious anticipatory schemes were *not* systematically used to try to "outguess" the pseudorandom target sequence.



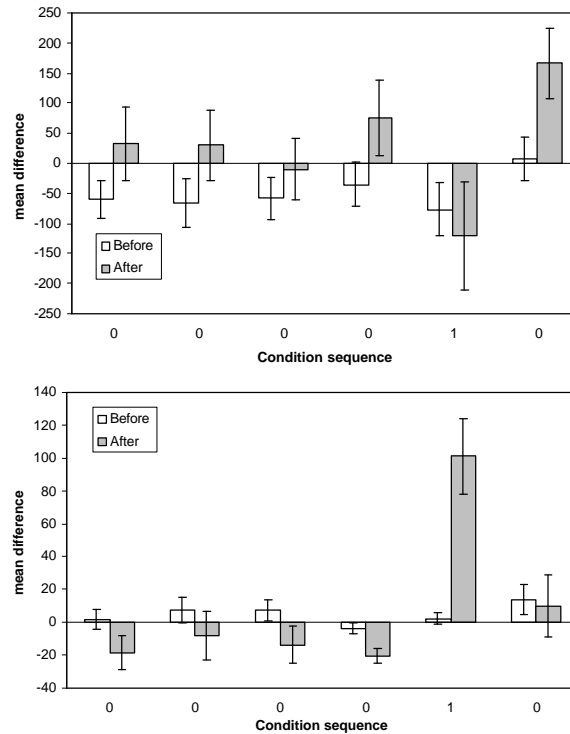


Figure 9. Mean HR, BVP, EDA (with one standard error bars) for before and after responses to calm (0) or emotional (1) targets, over all 50 individuals. These means are based on 74 instances when the target sequence 000010 occurred. “Before” is defined here as the HR measured at sample 15, three seconds into the before-display period, and “after” as the HR measured at sample 55, three seconds after the display period ended.

CONCLUSION

The present results are consistent with the earlier studies (Radin, 1997, Bierman & Radin, 1997), with conceptually similar studies (e.g., Bechara et al, 1997; Klintman, 1983, 1984), and with case studies of psi and intuitive hunches (e.g., Rhine, 1969). It is unlikely with the present design that the results are artifacts of sensory leakage given that the targets were selected in the future, and several analyses indicate that it is also unlikely to be due to inadequate target randomization or simple anticipatory strategies.

This experimental method may be unusually successful, not only in the statistical outcomes but in the ease of conducting the study, for two reasons: First, the literature suggests that emotional, meaningful targets are much better than abstract or neural targets (Moss & Gengerelli, 1968; Moss, 1969; Eisenberg & Donderi, 1979). In the present case, the use of a target set with a very high degree of contrast between calm and emotional stimuli seems to have provided a potent differential “signal” to work with. Second, rather than having to rely on conscious reporting, the use of unconscious psychophysiological measures may have provided a way to bypass a panoply of defense mechanisms that seem to block conscious reporting of psi perception.

The present design provides a promising method of exploring the nature of precognition. Of perhaps greater sociological importance for parapsychology, the technique is virtually identical to well-understood and well-replicated effects in conventional psychophysiology. The combination of commonly used, conventional methods with unexpected, psi-like outcomes may provide an easy way for psi researchers to bring their work to the attention of mainstream scientists.

REFERENCES

- Barron, F. & Mordkoff, A. M. (1968). An attempt to relate creativity to possible extrasensory empathy as measured by physiological arousal in identical twins. *JASPR*, **62**, 73-79.
- Bechara, A., Damasio, H., Tranel, D., Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, **275**, 28 February 97, 1293-1295.
- Beloff, J., Cowles, M. & Bate, D. (1970). Autonomic reactions to emotive stimuli under sensory and extrasensory conditions of presentation. *JASPR*, **64**, 313-319.
- Bierman (1997). Personal communication.
- Bierman, D. J. & Radin, D. I. (1997). Anomalous anticipatory response on randomized future conditions. *Perceptual and Motor Skills*, **84**, 689-690.
- Bierman, D. & Radin, D. I. (1998). Non conscious processes and intuition: Is there an anomalous component? *Consciousness Research Abstracts, Tuscon III: Towards a science of consciousness*. University of Arizona, Tuscon, April 1998.
- Blair, R. C. & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, **30**, 518-524.
- Bohlin, G. & Kjellberg, A. (1979). Orienting activity in two-stimulus paradigms as reflected in heart rate. In Kimmel, H. D., Van Olst, E. H. & Orlebeke, J. F. (1979). *The orienting reflex in humans*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Bradley, M. M., Greenwald, M. K. & Hamm, A. O. (1993). Affective picture processing, In N. Birbaumer & A. Ohman, (Eds.), *The Structure of Emotion*, Seattle, WA: Hogrefe & Huber Publishers.
- Braud, W. G. & Schlitz, M. (1989). A methodology for the objective study of transpersonal imagery. *Journal of Scientific Exploration*, **3**(1), 43-63.
- Braud, W. G. & Schlitz, M. J. (1991). Consciousness interactions with remote biological systems: Anomalous intentionality effects. *Subtle Energies*, **2** (1), 1-46.
- Braud, W. G. (1981). Psi performance and autonomic nervous system activity, *Journal of the American Society for Psychological Research*, **75**(1), pp. 1-35.
- Braud, W. G. (1993). On the use of living target systems in distant mental influence research. In L. Coly (Ed.), *Psi research methodology: A re-examination*. New York: Parapsychology Foundation.
- Braud, W., Shafer, D. & Andrews, C. S. (1992). Further studies of autonomic detection of remote staring: Replications, new control procedures, and personality correlates. *Proceedings of Presented Papers, Parapsychological Association 35th Annual Convention*, pp. 7-21;
- Claridge, G. S. (1967). *Personality and arousal*. New York: Pergamon Press, p. 27.
- Coles, M. G. H. (1983). Situational determinants and psychological significances of heart rate change. In Gale & Edwards, 1983.
- Dean, E. D. (1962). The plethysmograph as an indicator of ESP. *Journal of the Society for Psychological Research*, **41**, 351-353.
- Dean, E. D. (1966). Plethysmograph recordings as ESP responses, *International Journal of Neuropsychiatry*, **2**, 439-446.
- Delanoy, D. & Sha, S. (1994). Cognitive and physiological psi responses to remote positive and neutral emotional states. In D. Bierman (Ed.) *Proceedings of Presented Papers of the 37th Annual Convention of the Parapsychological Association*, 128-137.

- Delanoy, D. (1989). Characteristics of successful free-response targets: Experimental findings and observations. In L. A. Henkel and R. E. Berger (Eds.) *Research in Parapsychology 1988*, Metuchen, NJ: Scarecrow Press.
- Duane, T. D. & Behrendt, R. (1965). Extrasensory electroencephalographic induction between identical twins, *Science*, **150** (3694). pp. 367
- Eisenberg, H. & Donderi, D. C. (1979). Telepathic transfer of emotional information in humans. *Journal of Psychology*, **103**, 19-43.
- Esser, A. H., Etter, T. L. & Chamberlain, W. B. (1967). Preliminary report on physiological concomitants of "communication between isolated subjects." *International Journal of Parapsychology*, **9**, 53-56.
- Figar, S. (1959). The application of plethysmography to the objective study of so-called extrasensory perception. *Journal of the Society for Psychical Research*, **40**, 162-172.
- Grinberg-Zilberbaum, J., Delaflor, M., Arellano, M. E. S., Guevara, M. A. & Perez, M. (1992). Human communication and the electrophysiological activity of the brain, *Subtle Energies*, **3** (3), pp. 25-44.
- Hamm, A. O. & Vaitl, D. (1993). Induction of emotions via visual stimuli: Validation of an induction method of 3 response levels. *Psychologische Rundschau*, **44** (3), 143-161.
- Haraldsson, E. (1972). *Vasomotor reactions as indicators of extrasensory perception*. Doctoral dissertation, Albert-Ludwigs-Universität, Freiburg, Germany.
- Haraldsson, E. (1980). Confirmation of the percipient-order effect in a plethysmographic study of ESP. *Journal of Parapsychology*, **44** (2), 105-124.
- Hartwell, W. (1978) Contingent negative variation as an index of precognitive information. *European Journal of Parapsychology*, **2**, 83-103.
- Hettinger, J. (1952). *Telepathy and spiritualism*. London: Rider.
- Kelly, E. F. & Lenz, J. (1976). EEG changes correlated with a remote stroboscopic stimulus: A preliminary study. In *Research in Parapsychology 1975*, pp. 58-63.
- Kimmel, H. D., Van Olst, E. H. & Orlebeke, J. F. (Eds.) (1979). *The orienting reflex in humans*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Klintman, H. (1983). Is there a paranormal (precognitive) influence in certain types of perceptual sequences? Part I. *EJP*, **5**, 19-49.
- Klintman, H. (1984). Is there a paranormal (precognitive) influence in certain types of perceptual sequences? Part II. *EJP*, **5**, 125-140.
- Levin, J. & Kennedy, J. (1975). The relationship of slow cortical potentials to psi information in man. *Journal of Parapsychology*, **39**, 25-26.
- Lloyd, D. H. (1973). Objective events in the brain correlating with psychic phenomena. *New Horizons*, **1**, 69-75.
- Matrick, L. E. (1955). *A demonstration of extrasensory perception using the vasomotor response*. Unpublished student research project, University of Manitoba Medical College.
- May, E. C., Targ, R. & Puthoff, H. E. (1979). EEG correlates to remote light flashes under conditions of sensory shielding. In C. T. Tart, H. E. Puthoff & R. Targ (Eds.), *Mind at large*. New York: Praeger, pp. 127-136.
- Millar, B. (1976). An attempted validation of the "Lloyd effect." In *Research in Parapsychology 1975*, pp. 25-27.
- Moss, T. & Gengerelli, J. A. (1968). ESP effects generated by affective states. *Journal of Parapsychology*, **32**, 90-100.

Anticipatory responses to future emotions

- Moss, T. (1969). ESP effects in “artists” contrasted with “non-artists.” *Journal of Parapsychology*, **33**, 57-69.
- Nash, C. B. & Nash, C. S. (1969). Coincident vasoconstrictions in pairs of resting subjects. *Journal of the Society for Psychical Research*, **41**, 347-350.
- Peterson, D. M. (1978). Through the looking-glass: An investigation of the faculty of extra-sensory detection of being stared at. Unpublished thesis: University of Edinburgh, Scotland.
- Poortman, J. J. (1959). The feeling of being stared at. *Journal of the Society for Psychical Research*, **40**, 4-12.
- Pulley, B. (1998). Gambler learns she can't beat the house in court. *New York Times*, March 19, 1998, p. A12
- Radin, D. I. (1997). Unconscious perception of future emotions: An experiment in presentiment. *Journal of Scientific Exploration*, **11** (2), 163-180.
- Radin, D. I., Taylor, R. D. & Braud, W. (1995). Remote mental influence of human electrodermal activity: A pilot replication. *European Journal of Parapsychology*, **11**, 19-34.
- Rebman, J. M., Wezelman, R. Radin, D. I., Hapke, R. A. & Gaughan, K. (1996). Remote influence of the autonomic nervous system by focused intention. *Subtle Energies*.
- Rhine, L. E. (1969). Case study review. *Journal of Parapsychology*, **33**, 228-266.
- Rice, G. (1966). Emotional closeness, communication of affect, and ESP. *Proc. Of the PA*, **3**, 25.
- Sanjar, M. (1969). A study of coincident autonomic activity in closely related persons. *JASPR*, **63**, 88-94.
- Schlitz, M. J. & LaBerge, S. (1994). Autonomic detection of remote observation: Two conceptual replications. In D. J. Bierman (Ed.), *Proceedings of Presented Papers, Parapsychological Association 37th Annual Convention*, pp. 352-360. Fairhaven, MA: Parapsychological Association, Inc.
- Schlitz, M. J. and Wiseman, R. (1996). Experimenter effects and the remote detection of staring. In E. C. May (Ed.), *Proceedings of the 39th Annual Convention of the Parapsychological Association*, Fairhaven, MA: Parapsychological Association, Inc.
- Schouten, S. A. (1976). Autonomic psychophysiological reactions to sensory and emotive stimuli in a psi experiment. *European Journal of Parapsychology*, **1**, 72-78.
- Siddle, D. A. T, Kuiack, M. and Kroese B. S. (1983). The orienting reflex. In A. Gale & J. A. Edwards (Eds.) (1983). *Physiological correlates of human behaviour*. Vol. II: Attention and Performance. New York: Academic Press
- Sokolov, E. N. (1963). *Perception and the conditioned reflex*. New York: MacMillan .
- Targ, R. & Puthoff, H. (1974). Information transmission under conditions of sensory shielding. *Nature*, **252**, 602-607.
- Tart, C. T. (1963). Possible physiological correlates of psi cognition. *International Journal of Parapsychology*, **5**, 375-386.
- Venables, P. H. & Christie, M. J. (1973) Mechanisms, instrumentation, recording techniques, and quantification of responses. In Prokasy and Raskin (1973), p. 108.
- Warren, C. A., McDonough, B. E., & Don, N. S. (1992). Event-related brain potential changes in a psi task. *Journal of Parapsychology*, **56**, 1-30.
- Wiseman, R. & Smith, M. D. (1994). A further look at the detection of unseen gaze. In D. J. Bierman (Ed.), *Proceedings of Presented Papers, Parapsychological Association 37th Annual Convention*, Fairhaven, MA: Parapsychological Association, Inc., pp. 352-360.

Wiseman, R., Smith, M. D., Freedman, D., Wasserman, T. and Hurst, C. (1995). Examining the remote staring effect: Two further experiments. In N. L. Zingrone (Ed.), *Proceedings of Presented Papers, Parapsychological Association 38th Annual Convention*, Fairhaven, MA: Parapsychological Association, Inc., pp. 480-490.

POLTERGEIST AND NONLOCALITY: ENERGETIC ASPECTS OF RSPK

William G. Roll
*Department of Psychology,
State University of West Georgia*

Michael A. Persinger,
*Neuroscience Laboratory,
Laurentian University, Ontario*

ABSTRACT

Analyses of anecdotal accounts, on-site investigations and experimental studies suggest that RSPK (recurrent spontaneous psychokinesis or poltergeist) cannot be dismissed in terms of illusion, fraud or non-personal energetic forces. To arrive at a theory for RSPK, an analysis of 116 cases made by Roll in 1974 was compared with a survey by Huesmann and Schriever in 1989 of 54 cases investigated by the Institut of Grenzgebiete der Psychologie und Psychohygiene in Freiburg, Germany. Although the two surveys overlap (six of the cases in Roll's survey were also in the Freiburg collection) they seemed sufficiently independent for the present study. On the basis of these data, supplemented by recent research, the authors suggest that RSPK is due to electromagnetic (EM) energy in the environment modulated by EM from the agent and focused on significant objects. The process is similar to the electroacoustic effect where movement is induced in the diaphragm of a loudspeaker by an electric current. But in RSPK the EM energy moves through space-time without the benefit of electric wires presumably because it is highly focused. Psychodynamic and neuropsychological studies of RSPK agents reveal psychosocial tensions and often clinical or subclinical CPS (complex partial seizures). The authors suggest that these tensions, with or without overt CPS, result in an EM profile that can modulate and focus environmental EM on significant objects. A comparison between CPS and RSPK show suggestive similarities. E.g., both are worldwide in distribution; involve equal proportions of males and females; peak in the early teens; may be precipitated by stress; and onset is associated with increases of geomagnetic activity. The psi field theory (Roll, 1964) and Jahn and Dunne's (1989, pp. 208-219) concept of consciousness waves embrace the dual aspect of RSPK. RSPK is clearly a physical phenomenon, but a physical phenomenon with a human face.

INTRODUCTION

Instances of recurrent spontaneous psychokinesis (RSPK or poltergeist) consist of observable displays of physical energy, such as the movement of household objects, usually in proximity to one or more individuals who are unaware of their role in the disturbances. RSPK research is of three types, analyses of anecdotal accounts from people in whose vicinity the phenomena erupted, field studies by researchers, and experimental work. In 1974 one of us (Roll, 1977) made an analysis for the Physical Research Foundation (PRF) of published cases for which the evidence seemed compelling. We found 116, from 1612 to 1974, some of which had been included in two previous surveys (Cox, 1961; Zorab, 1964). In 1979 Gauld and Cornell (1979) published their seminal study of 500 cases of RSPK and haunt, of which 115 were high evidence RSPK from the same sources as ours. This was followed by Huesmann and Shriever's (1989) analysis of 54 reports from the Institut of Grenzgebiete der Psychologie und Psychohygiene in Freiburg, Germany, a major center for RSPK research. The Freiburg collection included six cases we had used; otherwise the two surveys are independent. This paper explores the characteristics of RSPK

shown by the PRF and Freiburg surveys, supplemented by more recent work.

RSPK has been attributed to three known causes, (i) observational error, (ii) non-personal mechanical forces, and (iii) deception. (i) Houran (1997) and Houran and Lange (1996) have argued that the apparent objective incidents may result from misinterpretation of environmental anomalies. This is a possibility for the minor occurrences they describe but seems unlikely for reports of loud percussive sounds and household objects that move several feet, especially when the incidents are repetitive and occur in the presence of witnesses. (ii) Lambert (1955) found RSPK and haunts to cluster near water, and suggested that changing levels jolted the house, shaking things off their shelves and causing strange sounds. Dingwall and Hall (1958) similarly supposed that the phenomena were sometimes due to earth tremors. Gauld and Cornell (1979, pp. 330-337) tested these theories by subjecting a house (scheduled for demolition) to mechanical vibrations. The objects only slid short distances, none performing the lengthy movements often seen in RSPK. The authors argued that the association of RSPK with water was an artifact due to the concentration of people in such areas. Even without their house-shaking experiment, non-personal physical forces could hardly be the explanation of RSPK since this usually depends on the presence of particular persons. (iii) The fraud theory is more plausible and in fact receives some support from the reports themselves; 26% of the Freiburg cases and 14% of the PRF's included one or more fraud incidents. These, however, seemed to be examples of "imitative" rather than "total" fraud, a distinction drawn by Cox (1961). In the former, the person copies genuine effects, while cases of total fraud are entirely deceptive. It is a mistake to dismiss a case because some of the incidents are fabricated. As Bender (1974) argued, simulated incidents may throw light on the psychodynamics of the real phenomena.

EXPERIMENTAL STUDIES

In 41% of the PRF reports there were tests of various levels of sophistication, usually challenges to the supposed entity (made in the presence of the RSPK agent) to produce movements of objects or rapping sounds, 40% apparently succeeding. Similarly 57% of the Freiburg cases included "expected observations" ("erwartender Beobachtung"), these being rewarded in 44%. In recent times, tests have been introduced to avoid observational errors. In the Miami case (Roll & Pratt, 1971) objects that had been frequently affected, such as glasses and ashtrays (the focusing effect), were placed in the most active areas. Nothing happened when the targets were filmed by news reporters or watched by the investigators but several moved when Julio, the 17-year-old agent, was under observation. There was a further advance in methodology in the Resch case where the incidents persisted in a laboratory. Several targets, such as tools and small pieces of equipment, moved when Tina, the 14-year-old agent, was under observation in other parts of the building (Stewart, Roll & Baumann, 1987). As in Miami, the objects did not move when the investigators were watching the targets, and video recording also impeded the effect (a concealed camera was installed in the lab but circumstances prevented its use).

The only definitive instrumental recording of RSPK was achieved by the Freiburg team during their investigation of Annemarie, a 19-year-old employee at a legal firm (Bender, 1968; Karger & Zicha, 1968). The phenomena included repeated deflections on the mains voltage monitor, a focusing effect that enabled the researchers to capture the phenomena with their own equipment.

If a lawful process underlies RSPK, this may be revealed by examining the features of the phenomena.

CHARACTERISTICS OF RSPK

Agent: In 79% of the PRF cases, the occurrences depended on the presence of a particular individual, and sometimes on two. (In the latter case, the person most closely involved was used to determine the characteristics of the case.) One agent was identified in 63% of the Freiburg cases, two agents in 17%, while in 6% the family or group as a whole was “agent.” As notable as the presence of human agents is the absence of animals (Owen, 1978), and also of very young children.

Gender: The gender distribution in the PRF survey changed over time. Prior to 1900, there were 79% female agents and 21% male, but for 1900-1974 the sexes were evenly distributed, with 51% female and 49% male. This is comparable to a distribution of 44% female and 56% male agents in the Freiburg collection.

Age: The age of RSPK agents in our sample ranged from 8 to the late 70s, with a median of 14 for both sexes. The medians for Freiburg, 14 for males and 12-and-a-half for females, were close to ours.

When age was related to domicile, 41% of the PRF agents who were below 19 lived away from home when the disturbances began (30% were less than 14). Similarly, 38% of the Freiburg agents below 19 did not live with both parents.

Agent Awake: Owen (1978) noted that, “...no poltergeistery occurs when the poltergeist individual is in normal sleep.” This was true for all the PRF cases. However, 9% of the Freiburg reports included occurrences when the agent was probably asleep.

Agent Absent: In a few instances, object-movements were reported at the site when the agent was away, sometimes by several miles (Roll, 1970; Roll, 1993; Roll and Pratt, 1971).

Occurrences: Most of the cases, 90.5% for the PRF and 87% for Freiburg, involved movement of household contents. In addition, 52% of the PRF cases and 57% of Freiburg's included percussive sounds, such as knocks, bangs, raps, and explosions. In 8% of our cases, floating or flashing lights were mentioned, while 26% of the Freiburg reports included “optische ercheinungen” (visual manifestations). This larger figure undoubtedly was due to the inclusion of apparitions, which we listed separately.

Focusing: RSPK often singles out certain objects or areas. Repeated disturbances of the same or similar objects has been termed “object focusing” and repeated disturbances in the same area, “area focusing” (Roll, 1968). Object focussing was evident in 85% of the PRF reports, and area focusing in 26%, either in conjunction with object focusing (10%), or by itself (16%). Focusing was independent of the agent's proximity to the objects. The Freiburg study does not mention focusing, but 41% of the cases were concentrated in certain areas.

Geomagnetism: In the 30 PRF cases where the date of onset was known, this was compared with geomagnetic values (Roll and Gearhart, 1974). In 60% the onset was on days above the yearly average and 40% on days below, a noninsignificant difference. The correlation improved when changes in activity were used: in 73% there was an increase in geomagnetic disturbance between the days preceding and following the onset of RSPK. When Gearhart and Persinger (1986) increased the sample size (to about 50), the correlation was significant.

Weight (Mass) of Object: We did not distinguish objects by weight except for the attenuation analyses (see below), which were confined to items weighing a pound or less. However, pieces of furniture nearly always moved shorter distances than smaller objects. (More light than heavy objects were affected in the Freiburg cases, perhaps because there are more light objects in most dwellings.)

Attenuation: In ten cases, where the distances were measured between the agent's location and the starting points of object-movements, frequency declined with distance (Artley and Roll, 1968; Bender, 1968; Eisler, 1974; Palmer, 1974; Roll, 1968, 1970; Roll, Burdick & Joines, 1973, 1974; Solfvin, Harary & Batey, 1976). Experimental macro-PK may also show attenuation (e.g. Brookes-Smith and Hunt, 1970).

Exponential Decay: For a more exact analysis of attenuation the data of five PRF cases, which we had personally investigated, were plotted against two common energetic functions, the inverse square function and the exponential decay function (Roll, 1968, 1970; Roll, Burdick & Joines, 1973, 1974; Roll & Stump, 1969). The latter gave a better fit in four cases, the former in one. The data in this case were probably skewed because of double agency (Roll, 1970). The inverse square function describes the dispersal of energy such as light from a point source in empty space; the exponential decay function describes the dispersal when a significant amount of the energy is absorbed by another medium (e.g., light moving through water is converted to heat and thereby attenuated more quickly).

Rotating Beam: Although there were fewer occurrences remote from the agent, the remote objects moved further, which seemed to suggest that the energy was stronger at the distant locations. To explore this apparent anomaly, the vectors of object-movements were analyzed in the two cases with enough incidents (Roll, Burdick & Joines, 1973; 1974). In both, objects close to the agent tended to move in short, clockwise and outward patterns, while distant objects moved farther, were counter-clockwise and pointed inward. This pattern could result from a rotating beam of energy centered on the agent and produced by two sources that were slightly out-of-phase.

Uneven Trajectories: In 41% of the PRF reports, the movements were described as floating, zigzag or fluttering. Changing speeds were mentioned as well as objects coming around corners or rotating in the same spot. Similarly Freiburg had 45% with “*ungewöhnlichen Flugbahnen*” (unusual trajectories).

Teleportation (Apport): In 17% of the PRF cases, some objects appeared in or disappeared from closed rooms or other confined space. Freiburg had a larger number, 37% with “*Penetrationen*,” probably because objects were included that appeared in open areas.

Apparitions and Voices: In 23% of the PRF cases, apparitions of human figures, animals, demons, hands, fingers or amorphous shapes were reported. In 11% there were intelligible voices and in 7% sounds of wind, whistling, laughter, groans, or screams. In 24% the experiences were confined to the agent, but in 51.5% they were had by other members of the household (either by themselves or in addition to the agent), and in 24%, surprisingly, by outside witnesses. The Freiburg cases showed 17% with voices but only 12% with apparitions. This figure would undoubtedly increase if their “*optische Ercheinungen*” of figures were included.

Intelligible Information: Aside from the visions and voices, in 40.5% of the PRF cases intelligible information came through writings on walls, knocks spelling out messages, or mediumistic-type communications by the agent. In about half, incorporeal entities were identified, mostly as demons and creatures of fantasy but also as living and, in 11 cases, as deceased people. Some of the information about the deceased was correct and some false or unverifiable.

Onset, Duration and Termination: In 41% of our data, onset followed moves, illness, or other events that may be upsetting to people, while 20% of the Freiburg cases began when the agent expressed “rage, disappointment or great frustration.” Both surveys showed November with more RSPK onsets than other months. In Germany, Huesmann and Schriever note, November is dark, snow-less, and lacks holidays and is therefore more depressing than other months. The same is true for most of the PRF sites.

The PRF survey showed the duration of RSPK to range from one day to six years, with an average of 5.1 months and a median of two. The median reported for Freiburg was rather larger, 5 months.

Exorcism: Religious intervention was tried in 26% of the PRF and 41% of the Freiburg cases, with a successful outcome of 13% and 10%. These figures, modest though they be, are questionable since the RSPK might have run its course when the rites were used. (In some instances, these were followed by a temporary cessation of activity, perhaps indicating an observer effect). On the other hand, changes in the composition of the group may have shortened duration. In 37% of the PRF cases where the agent or someone else moved away, the median dropped from two to 1.13 months, a suggestive but nonsignificant difference.

Inverse Relation Between RSPK and Medical or Mental Symptoms: Some cases showed an inverse relation between RSPK and medical or mental conditions. In the Bell Witch case (Ingram, 1894), the knocking sounds ceased when 12-year-old Betsy Bell had her “spells”; in the Wem case (Podmore, 1896-97) a young servant girl, who had been healthy before the RSPK, “...afterwards fell into ill-health and became subject to fits”; for Edwin March (Thacker, 1910) the incidents commenced after the boy's epileptic seizures had abated; the knocking sounds around a young woman, Karin, (Bjerre, 1947) stopped during convulsions; the two grand-mal seizures of a 21-year-old male (Solfvin and Roll, 1976) were followed by the only two periods without RSPK; and for a woman of about 40 (Roll and Tringale, 1983) migraine attacks alternated regularly with RSPK. Owen (1978) observed that “In the few poltergeist cases where anything resembling a trance is reported, the phenomena...do not seem to happen during the trance but at other times.”

Observer Effects: Type I: The poltergeist has a reputation of disappearing in the presence of witnesses. This was true only for 14% of the PRF reports. Type II: In 78% the observers either did not inhibit the occurrences or seemed to enhance them (this includes studies where the visitors seemed to suppress the activity at first but then enhanced it, e.g., Roll and Pratt, 1971; Roll, 1993). This value, however, may be inflated since the survey was limited to cases with witnessed incidents. Observer effects are not mentioned in the Freiburg study. Type III: Although the phenomena rarely decreased when the observers were watching the *agent*, in 45% of the PRF cases visual observation of the *objects* seemed to impede movement, but had no effect in 41%. When attempts were made to film active objects, this too suppressed activity (e.g. Stewart, Roll & Baumann, 1987; Roll & Pratt, 1971). Filming of suspended lamps and pictures already in motion has been reported (Bender, 1974; Rosenberg, 1974).

Voluntary RSPK: After the initial RSPK outburst, several agents produced macro-PK more of less at will (Bjerre, 1947; Fodor, 1958; Grant, 1965; Ingram, 1894; Owen, 1978; Price, 1933, pp. 228-273; M. Roll, 1969; Thurston, 1954, pp. 32-36); some making a career of exhibiting PK, including D.D. Home (Crookes, 1889-90), Matthew Manning (Owen, 1978) and Eusapia Palladino (Feilding, 1963). Attempts to elicit voluntary macro-PK in subjects with no history of PK have also provoked RSPK. In Batcheldor's (1966, 1983) "sitter group" tests, there were spontaneous movements of furniture and rapping sounds in addition to experimental effects. Brookes-Smith (1973) and Brookes-Smith and Hunt (1970) also evoked macro-PK, including violent RSPK incidents. In the "Philip" experiments by Owen and Sparrow (1976), an imaginary ghost manifested as RSPK-type thumps and raps, flickering of lights and movements of a table. (In Batcheldor's experiments a PK agent was identified, in the others the group as a whole seemed to be "agent").

Crude or Smart Poltergeists: A contrast has been made between RSPK that shows little evidence of intelligence, such as explosive sounds and random object-movements, "crude RSPK," and incidents that reflect intelligence, "smart RSPK" (Montagno and Roll, 1983).

ENERGETIC ASPECTS

Four forms of energy are known to science, the electromagnetic, the gravitational, the strong atomic force and the weak atomic force. Results in tests for micro-PK, where the targets are generated by random firings of electrons, may suggest an interaction between human operators and the weak atomic force. Deviations from chance expectancy in such tests are referred to as micro-PK. RSPK is a form of macro-PK, that is, it consists of movements of household objects, percussive sounds and other occurrences that (for all practical purposes) do not happen "by chance" but require an obvious source of energy.

Results in micro-PK suggest the non-local or non-continuous character of quantum mechanics (e.g., Jahn and Dunne, 1987, pp. 91-115) since the effect does not seem to diminish with distance. In the world of large objects, however, "...classical continuum behavior obtains" (p. 263). RSPK may be governed by the laws of classical physics because large objects are affected.

RSPK is usually a display of kinetic energy, but this can hardly be the source of the phenomena because kinetic energy is transmitted by tangible contact (discounting quantum mechanics). It seems more likely that RSPK is due to the electromagnetic force (EM) since EM is intangible, is produced by the body, and may be transformed to kinetic energy, including sound and motion.

The patterns of object-movements seem consistent with this view. If EM is converted to kinetic energy, the attenuation should follow the exponential rather than inverse function because EM energy would be lost in the transformation process. If there are two rather than one EM source, their interaction could result in a rotating beam that changes from short and broad to long and narrow. Since the short beam would pick up more objects and carry them shorter distances while the long beam would affect fewer objects but carry them farther, this could result in the observed patterns. The erratic trajectories and uneven speeds could be explained by the same process; as the beam sweeps back and forth, objects impacted only once might show straight trajectories and even speeds, while objects that were dropped and caught by another sweep might show complex trajectories and changing speeds.

Area focusing could be expected if the beam were reflected back from a concave barrier to a focal

point. As Joines (1975) has noted, the energy would "...produce a funnel-shaped...whirlpool which stays in one position and slowly dies out as time passes." He detected 146 Mhz radiation, lasting about a minute, where an object had just landed. The site was two feet in diameter, consistent with wavelength peaks at the 146 Mhz frequency. In other instances, the perturbation might remain longer and affect objects even when the agent was absent, as is sometimes reported. Like "a tornado, which generates an electromagnetic wave due to the swirling electrical charges," (p. 149) Joines's whirlpool might move about more or less independently of the agent, provoking RSPK along the way.

Relevant observations have been made in tests of macro-PK. Hasted has (1977) postulated a "surface of action" from the agent to the pieces of metal he used as targets. Other data suggested that "part or all of the surface of action is capable of rotation at several revolutions per second about an axis in its own plane," an idea reminiscent of Joines' "whirlpool."

Taylor (1975) thought that metal-bending and movement of light objects, as in the macro-PK tests with Kulagina and Vinogradova, could be due to simple EM. This does not seem unlikely; the magnetic component of the dipoles associated with the QRS component of the electro-cardiogram generate a rotating field that has been detected 50 cm from the body (the short range of Kulagina's PK and her increased heart rate seem consistent with Taylor's thesis). But Taylor did not see "...how enough energy can be radiated from the human brain to cause...an object to be hurled across the room" (pp. 151-152). The role of the agent, however, may not be to generate the energy but rather to modulate environmental EM. In the same way as each human face reflects sunlight in a distinctive way, a characteristic EM field is emitted by everyone and can be detected a few millimeters from the scalp with standard equipment. For this field to be reflected in the environment as in RSPK, an external source of energy would have to be present. The increases of geomagnetic energy often associated with the onset of RSPK may be one such source, local EM fields another (Nicols & Roll, 1998). The effect may be cumulative rather than sudden; like the proverbial butterfly whose wing-beat is the seed of a hurricane, a weak EM effect could grow into a powerful and semi-independent force assuming sufficient input from the environment.

In addition to object-movements and percussive sounds, RSPK may include simple EM effects, such as interference with electric power and flashing lights. In a case explored by one of us (Roll, 1972, Ch. 6) bursts of white light near a 19-year-old girl looked like electronic flashes. In an important study by Taff and Gaynor (1976) the main effects were light displays in the presence of a woman in her 30ies. The investigators saw a moving ball of blue light, a pulsing white light, and greenish net-like displays. Some of the lights moved, changed in intensity, expanded or contracted in response to requests or the emotional state of the woman. None of the visible lights turned up on photographs perhaps because they were in the ultraviolet band of the spectrum, which would have been attenuated by the camera lenses. But when the cameras were pointed to areas where a psychic said she sensed something, shapes and archs of light appeared on the photos (the investigators saw nothing). The lights were often attended by transient cold spots and unpleasant odors.

A notable study by Bayless (1980), reveals how EM may be converted to kinetic energy. The phone and paging systems in a factory went dead or were drowned by high-frequency sounds when a young female worker was present. Her electric typewriter and calculator, and Bayless' tape recorder were also affected. The year before another worker had experienced similar problems. The plant equipment was the source of strong radio waves and Bayless proposed that the two women "...modulated radio waves into a form that could be picked up physically on

electronic equipment, such as telephones and speakers.” In other words, the women had become electroacoustic sources and the phone and paging receivers electroacoustic transducers. (An electroacoustic source transforms electrical energy into kinetic energy which is heard as sound.) It seems possible that a similar mechanism is responsible for object-movements; in other words, RSPK may be an electroacoustic or “electrokinetic” phenomenon.

It may seem implausible that the relatively weak voltages presumably involved in RSPK could cause loud bangs and extensive object-movements. Electroacoustic effects, however, are sometimes out of proportion to the applied voltage. In the same way as a loudspeaker at maximum output may show nonlinear behavior, that is, produce large diaphragm movements not directly proportional to the applied current, the force exhibited in RSPK may be out of proportion to the EM energy from the source; loud sounds and extended object-movements could result from a relatively weak charge.

The voltage produced by some individuals may not be negligible. In a study by Green et al. (1991) surges of up to 221 volts were recorded from the bodies of therapeutic touch healers (median 8.3 volts). The subjects were seated on a platform supported by glass blocks to prevent EM from extraneous sources. The surges were brief, with a median duration of 3.6 seconds. The effect was not seen with regular subjects and only with six of the healers. The experimental room was an adaptation of a Tibetan meditation chamber and featured a bar magnet over the subject's chair and copper panels in front and back and on the ceiling and floor; the front panel also forming a mirror. The panels showed similar voltage increases, but they and the magnet probably did not contribute to the surges. The authors suggest that the healers had “a different 'energy structure,' or a different 'energy-handling capability'” than others. Assuming the process is real, it may be applied unconsciously by RSPK agents. If this is the case, it should be possible to detect an EM source in these individuals.

NEUROPSYCHOLOGICAL ASPECTS

Fifty-three percent of the agents in our sample had medical or mental problems. They included Bernard, a 13-year-old studied by Bender (1969) in 1952, the first time psychodiagnostic tests were used for RSPK. The boy showed “high inner tensions in connection with puberty, frustration and aggressive behavior with a tendency towards an explosive vent.” Similar observations, though not restricted to puberty, have been made of other RSPK agents (Owen, 1964; Palmer, 1974; Rogo, 1986; Roll, 1968, 1970, 1972, Ch. 13; Zorab, 1964). The most thorough psychodiagnostic studies of RSPK agents involved a 15-year-old boy and a woman of 19, agents of two highly evidential cases from Freiburg (Bender, 1968, 1969, 1974; Karger & Zicha, 1968; Mischo, 1968; Mischo, Timm & Vilhjalmsson, 1968). Mischo (1968) found ego weakness and labile personality structures, irritability, low tolerance of frustration and transient aggressive impulses, directed not at their objects (other people) but at substitute objects (objects belonging to these persons). The agents showed extreme tension due to inordinate needs for acceptance contrasted with feelings of inferiority and fears of being abandoned, a tension that was relieved by RSPK and aggressive behavior. (As suggested by the two surveys, illness, time of year and, for children, living away from home or parents, may be additional stressors.) But “For each poltergeist personality there must be thousands with similar problems who lack this capacity.” (Roll, 1972, p. 175). A case that centered on a young man with grand mal epilepsy (Solfvin & Roll, 1975), narrowed the field to individuals displaying sudden bursts of EM in their central nervous system (CNS). Twenty-four percent of the PRF agents with miscellaneous problems (53% of the total) had symptoms suggestive of epilepsy, such as seizures, convulsions and fainting

spells, or had been diagnosed with epilepsy (4%). A comparison between complex partial seizures (CPS; formerly referred to as psychomotor epilepsy) and RSPK show likenesses that may suggest similar processes (Roll and Montagno, 1983). RSPK and CPS both represent displays of physical energy, are worldwide in distribution, involve equal proportions of males and females, peak in the early teens, and occur spontaneously or in response to arousal. In both, onset is sometimes associated with illness and prolonged tension, stressful events may precipitate recurrence, and the personalities of the subjects are often alike. Elevations in geomagnetic activity increase the risk of CPS as well as RSPK. Some forms of epilepsy are limited to the patient's waking hours as is RSPK. CPS can sometimes be brought on voluntarily; the same is true for RSPK. The repetitive involvement of specific objects in RSPK are suggestive of the convulsive body movements and forced thinking in CPS; and visual and auditory hallucinations are found in both. For some CPS patients there is an inverse relation between seizures and anomalies such as schizophrenia; and for some RSPK agents there is an inverse relation between RSPK and other indicators. Epilepsy may result in a wide spectrum of symptoms, from simple tonic-clonic movements to elaborate hallucinations; similarly RSPK can be crude or smart.

The CPS model predicts that crude RSPK is associated with discharges in the motor cortices, and smart RSPK with discharges in the temporal lobes or the gray matter of the upper brainstem (Montagno and Roll, 1983). Since directions from higher brain centers to the body go through the brainstem, the CPS theory also predicts that brainstem anomalies divert the EM traffic in RSPK agents.

McHarg (1977) found the onset of RSPK to coincide with the early stage of bronchial carcinoma in the elderly agent, and cessation to coincide with his death nine months later. Since the bronchi are served by the vagus nerve, which emerges in the same area of the brainstem, this may have been affected by the disease as well.

Some of the predictions were supported by our study of Tina. Her psychological profile resembled the agents studied by Mischo (1968), including feelings of inferiority and fear of being deserted. When the child was ten months old, she contracted pneumonia and her mother brought her to a hospital and then abandoned her. Her adoptive family could not give Tina all the attention she needed, so she acted up and was punished, thereby simultaneously being attended to and rejected. Neuropsychological tests showed two major anomalies, one in the brainstem (Baumann, 1995) as predicted, and another in the right frontotemporal lobe (Persinger & Roll, 1993), which seemed consistent with this case of smart RSPK. A head injury when Tina was 12 may have lowered seizure threshold, and geomagnetic bursts and psychosocial tension may have provoked them. She had a dramatic confrontation with her father the night before the onset of RSPK and there was a geomagnetic storm at the time.

Tina's RSPK was regularly attended by headaches and fluttering feelings in the stomach, the former possibly related to the frontal anomaly and the latter to the vagus nerve. The vagus, which is close to the anomalous area in Tina's brainstem has fibres to the abdomen and to the thorax, and may have been damaged by her childhood pneumonia.

It may be significant that "auto-noetic memory," memories as experiences of *the self*, is mediated by the right frontal lobe (Wheeler, Stuss & Tulving, 1997). Auto-noetic consciousness is a unique attribute of the developed human brain and could explain why RSPK is not seen in animals or very young children.

Tina's apparent need for attention may throw light on two observer effects. The later RSPK occurrences mostly took place when she was in the presence of investigators (Type II) and when these were observing her. From being destructive (of her parents' belongings) the RSPK had become a positive response to the attention of the visitors. But when their attention instead was directed to potential RSPK targets, the activity ceased (Type III). The same process may have been at work for Julio (Roll and Pratt, 1971) and Roger (Roll, 1972, Ch. 11), the agents in the Miami and Olive Hill cases; and it may underlie voluntary RSPK as well.

Our theory requires a correlation between specific discharges in the central nervous system, which represent the agent's mentation, and RSPK incidents that reflect this same mentation. Tests by one of us (Ruttan, Persinger & Koren, 1990) have shown that applied EM pulses of certain amplitudes and frequencies may result in certain mental conditions. In RSPK, the process would be reverse, from brain to environment. While experimental evidence for this process is lacking, there is some evidence of EM effects in the environment that coincide with a person's intention to affect objects there. We mentioned Joines' (1975) recording of 146 Mhz at the site of an object-movement. A similar observation was made during an out-of-body study where Keith Harary, a member of the PRF research group, attempted to affect the behavior of his kitten 1/4 mile away (Morris et al., 1978). Mhz frequency was measured on six occasions at this site. One time the onset of 145 Mhz activity was observed, which lasted almost exactly the duration of the OBE. Four other times there was an increment in an already present activity during experimental OBE periods (but not during controls), twice at 160-165 Mhz and twice at 190 Mhz. The sixth time there was no activity during either period. The similarity of these values to Joines' observation is suggestive. Sometimes outside witnesses have apparitional experiences at RSPK sites. These could be due to suggestion or they might result from the same EM energies that bring on the RSPK effects. Experimental work (e.g. Ruttan, Persinger & Koren, 1990) and investigations of haunts (e.g., Persinger, 1984; Roll, Maher, and Brown, 1992; Roll, Shehan, Persinger, and Glass, 1996) suggest that hallucinations can be evoked by external EM fields.

DISCUSSION

Poltergeist phenomena have been reported for centuries, often under conditions when perceptual error, fraud or familiar energetic processes could be ruled out. This does not mean that RSPK is a paranormal phenomenon in the sense of falling outside the framework of established science. On the contrary, it appears that the energy underlying RSPK is both physical and psychological. In this respect, RSPK does not differ from familiar forms of behavior.

The force displayed in RSPK is also not extreme, rarely amounting to more than the energy required to lift the several pounds of one's leg. RSPK represents kinetic energy. This is usually transmitted through mechanical contact, as when the hand brings a coffee cup to the lips. In RSPK, on the other hand, things are flung across rooms without tangible contact. Although RSPK is a kinetic phenomenon, it must be derived from an intangible force. We suggest that RSPK and perhaps other types of macro-PK may be electroacoustic occurrences. The electromagnetic force that may underlie RSPK, we suggest, is primarily environmental. The agent's role may be to modulate and focus rather than generate EM energy. Neuropsychological studies point to the central nervous system as the origin of the EM discharges that may be responsible for this regulatory activity outside the body.

An early version of the CNS theory was criticized by Martinez-Taboas (1984) as too broad to be falsified. Whatever can be said in favor of this argument, the current theory is eminently testable.

Alvarado (1984) has objected to the theory that RSPK is associated with aggression in the agent because the projective tests that were often used (such as the Rorschach) are of questionable reliability and because the psychologists who interpreted the material may have been biased by the prevailing theory. Both points are well taken. At the time when most of this work was conducted, RSPK researchers were more concerned with the search for clues about RSPK than testing hypotheses.

The EM theory for RSPK must meet at least four requirements. There must be a mechanism that converts bursts of EM to the kinetic energy of object-movements and sounds; EM energy of sufficient amplitude to generate RSPK must be recorded at the site; it must be shown why only some individuals and not others are associated with RSPK; and there must be a mechanism whereby specific mental contents, e.g. anger at mother, may result in a specific related incident, e.g. the movement of mother's coffee cup, while other objects are left undisturbed. The material reviewed in this paper suggests that these goals are within reach.

One feature of RSPK cannot be accommodated by conventional concepts, the reports that objects appear in or vanish from enclosed space. Bender (1969) advanced the idea of four-dimensional space in this context: "Clearly, no room is closed if an object can take a trajectory in higher space. It will, in addition, appear or disappear instantaneously." (p. 101). The thought that the world of macro objects has four dimensions is not new (Hawking, 1988). But the possibility that the human mind and brain may interact with four-dimensional space-time has not been systematically explored as far as we know. It is an exciting prospect.

In his PA presidential address, "New Developments in Poltergeist Research," Bender (1969) found "...psyche and matter...so inseparably entangled that the discrimination between an external (physical) and an internal (psychological) aspect might...be an inadequate way to grasp what essentially happens." The psi field theory (Roll, 1964) advanced some years ago and Jahn and Dunne's (1989, pp. 208-219) more recent concept of consciousness waves embrace the dual aspect of RSPK. RSPK is clearly a physical phenomenon, but a physical phenomenon with a human face.

REFERENCES

- Artley, J.L. & Roll, W.G. (1968). Mathematical models and the attenuation effect in two RSPK (poltergeist) cases. *Proceedings of the Parapsychological Association*, 5, 29-31.
- Alvarado, C.S. (1984). Psychological evaluations of poltergeist agents: A critique. In R.A. White and R.S. Broughton (Eds.), *Research in parapsychology 1982*. Metuchen, N.J.: Scarecrow. pp. 117-118.
- Batcheldor, K.J. (1966). Report on a case of table levitation and associated phenomena. *Journal of the Society for Psychological Research*, 43, 339-356.
- Batcheldor, K.J. (1983). Contributions to the theory of PK induction from sitter-group work. In W.G. Roll, J. Beloff & R.A. White (Eds.), *Research in parapsychology 1982* (pp. 45-52). Metuchen, N.J.: Scarecrow.
- Baumann, S.B. (1995). An overview with examples including a case study of an RSPK subject. *The Parapsychological Association 38th Annual Convention*, August 5-8, 1995, 11-19.
- Bayless, R. (1980). En electronic "poltergeist"? *Theta*, 8, 9-11.
- Bender, H. (1968). An investigation of "poltergeist" occurrences. *Proceedings of the Parapsychological Association*, 5, 31-33.

- Bender, H. (1969). New developments in poltergeist research: Presidential Address. *Proceedings of the Parapsychological Association*, 6, 81-102.
- Bender, H. (1974). Modern poltergeist research: A plea for an unprejudiced approach. In J. Beloff (Ed.), *New directions in parapsychology* (pp. 122-143). London: Elek.
- Bjerre, P. (1947). *Spokerier*. Stockholm: Centrum.
- Brookes-Smith, C. (1973) Data-tape recorded experimental PK phenomena. *Journal of the Society for Psychical Research*, 47, 69-89.
- Brookes-Smith, C. & Hunt, D.W. (1970). Some experiments in psychokinesis. *Journal of the Society for Psychical Research*, 45, 265-281.
- Cox, W.E. (1961). Introductory comparative analysis of some poltergeist cases. *Journal of the American Society for Psychical Research*, 55, 47-57.
- Crookes, W. (1989-90). Notes of seances with D.D. Home. *Proceedings of the Society for Psychical Research*, 6, 98-127.
- Dingwall, E.J., & Hall, T.H. (1958). *Four modern ghosts*. London: Duckworth.
- Eisler, W. (1974a). The Bronz poltergeist. In W.G. Roll, R.L. Morris & J. Morris (Eds.), *Research in parapsychology 1973* (pp. 139-143). Metuchen, N.J.: Scarecrow.
- Feilding, E. (1965). *Sittings with Eusapia Palladino and other studies*. New Hyde Park, N.Y.: University Books.
- Fodor, N. (1958). *On the trail of the poltergeist*. New York: Citadel.
- Gauld, A. & Cornell, A.D. (1979). *Poltergeists*. London: Routledge.
- Gearhart, L., & Persinger, M.A. (1986). Geophysical variables and behavior: XXXIII. Onsets of historical and contemporary poltergeist episodes occurred with sudden increases in geomagnetic activity. *Perceptual and Motor Skills*, 62, 463-466.
- Grant, D. (1965). *The Cock Lane ghost*. New York: St. Martin.
- Green, E., Parks, P.A., Guyer, P.M., Fahriaon, S.L. & Coyne, L. (1991). Anomalous electrostatic phenomena in exceptional subjects. *Subtle Energies*, 3, 69-94.
- Hasted, J.B. (1977). Physical aspects of paranormal metal bending. *Journal of the Society for Psychical Research*, 49, 583-607.
- Hawking, S. (1988). *A brief history of time*. New York: Bantam.
- Houran, J. (1997). Ambiguous origins and indications of "poltergeists." *Perceptual and Motor Skills*, 84, 339-344.
- Houran, J. & Lange, R. (1996). Hauntings and poltergeist-like episodes as a confluence of conventional phenomena. *Perceptual and Motor Skills*, 83, 1307-1316.
- Huesmann, M. & Schriever F. (1989). Steckbrief des Spuks. Darstellung and Diskussion einer Sammlung von 54 RSPK-Berichten des Freiburger Instituts fur Grenzgebiete der Psychologie und Psychohygiene aus den Jahren 1947-1986. *Zeitschrift fur Parapsychologie und Grenzgebiete der Psychologie*, 31, 52-107.
- Ingram, M.V. (1894). *An authenticated history of the famous Bell Witch*. Nashville, Tenn.: Setliff.
- Jahn, R.G. & Dunne, B.J. (1987). *Margins of reality: The role of consciousness in the physical world*. New York: Harcourt Brace.
- Joines, W.T. (1975). A wave theory of psi energy. In J.D. Morris, W.G. Roll & R.L. Morris (Eds.), *Research in parapsychology 1974*

- (pp. 147-149). Metuchen, N.J.: Scarecrow.
- Karger, F. & Zicha, G. (1968). Physical investigation of psychokinetic phenomena in Rosenheim, Germany, 1967. *Proceedings of the Parapsychological Association*, 5, 33-35.
- Lambert, G. W. (1955). Poltergeists: A physical theory. *Journal of the Society for Psychical Research*, 38, 49-71.
- McHarg, J.F. (1977). A poltergeist case from Glasgow. In J.D. Morris, W.G. Roll & R.L. Morris (Eds.), J.D. Morris, W.G. Roll & R.L. Morris (Eds.), *Research in parapsychology 1976* (pp. 13-15). Metuchen, N.J.: Scarecrow.
- Martinez-Taboas, A. (1984). An appraisal of the role of aggression and the central nervous system in RSPK agents. *Journal of the American Society for Psychical Research*, 78, 56-69.
- Mischo J. (1968). Personality structure of psychokintic mediums. *Proceedings of the Parapsychological Association*, 5, 35-37.
- Montagno, E. de A. & Roll, W.G. (1983). A neurobiological model for psychokinesis. In W.G. Roll, J. Beloff & R.A. White (Eds.), *Research in parapsychology 1982* (pp. 272-273). Metuchen, N.J.: Scarecrow.
- Morris, R.L., Harary, J.B., Janis, J., Hartwell, J., & Roll, W.G. (1978). Studies of communication during out-of-body experiences. *Journal of the American Society for Psychical Research*, 72, 1-21.
- Nicols, A. & Roll, W.G. (1998). The Jacksonville water poltergeist: Electromagnetic and neuropsychological aspects. Unpublished paper.
- Owen, A.R.G. (1964). *Can we explain the poltergeist?* New York: Garrett.
- Owen, A.R.G. (1978). Poltergeist phenomena and psychokinesis. In M. Ebon (Ed.), *The Signet handbook of parapsychology* (pp. 365-378). New York: New American Library.
- Owen, I.M. & Sparrow, M. (1976). *Conjuring up Philip: An adventure in psychokinesis*. New York: Harper & Row.
- Palmer, J. (1974). A case of RSPK involving a ten-year-old boy: The Powhatan Poltergeist. *Journal of the American Society for Psychical Research*, 68, 1-33.
- Persinger, M.A. (1985). Geophysical variables and behavior; XXII. The tectonic strain continuum of unusual events. *Perceptual and Motor Skills*, 60, 59-60.
- Persinger, M.A. and Roll, W.G. (1993). Potential neurofunctional correlates of the Tina Resch 1984 poltergeist episode. In *Proceedings of Presented Papers, The Parapsychological Association 36th Annual Convention*, August 15-19, 1993, 483-491.
- Pratt, J.G. & Roll, W.G. (1958). The Seaford disturbances. *Journal of Parapsychology*, 22, 79-124.
- Podmore, F. (1896-97). Poltergeists. *Proceedings of the Society for Psychical Research*, 12, 45-115.
- Pratt, J.G. & Palmer, J. (1976). *An investigation of an unpublized family poltergeist*. In J.D. Morris, W.G. Roll & R.L. Morris (Eds.), *Research in parapsychology 1975* (pp. 109-115). Metuchen, N.J.: Scarecrow.
- Price, H. (1933). *Leaves from a psychist's case book*. London: Gollanz.
- Rogo, D.S. (1986). *On the track of the poltergeist*. Englewood Cliffs, N.J.: Prentice-Hall.
- Roll, W.G. (1964). The psi field. *Proceedings of the Parapsychological Association*, 1, 32-64.
- Roll, W.G. (1968). Some physical and psychological aspects of a series of poltergeist phenomena. *Journal of the American Society for Psychical Research*, 62, 263-308.

- Roll, W.G. (1970). Poltergeist phenomena and interpersonal relations. *Journal of the American Society for Psychological Research*, 64, 66-99.
- Roll, W.G. (1972). *The poltergeist*. New York: Nelson Doubleday. Roll, W.G. (1977). *Poltergeists*. In B.B. Wolman (Ed.) *Handbook of parapsychology* (pp. 382-413). New York: Van Nostrand Reinhold.
- Roll, W.G. (1993). The question of RSPK vs. fraud in the case of Tina Resch. In *Proceedings of Presented Papers, The Parapsychological Association 36th Annual Convention*, 456-482.
- Roll, W.G., Burdick & Joines, W.T. (1973). Radial and tangential forces in the Miami poltergeist. *Journal of the American Society for Psychological Research*, 67, 267-281.
- Roll, W.G., Burdick & Joines, W.T. (1974). The rotating beam theory and the Olive Hill poltergeist. In W.G. Roll, R.L. Morris & J. Morris (Eds.), *Research in parapsychology 1973* (pp. 64-67). Metuchen, N.J.: Scarecrow.
- Roll, W.G. & Gearhart, L. (1974). Geomagnetic perturbations and RSPK. In W.G. Roll, R.L. Morris & J. Morris (Eds.), *Research in parapsychology 1973* (pp. 44-46). Metuchen, N.J.: Scarecrow.
- Roll, W.G., Maher, M., & Brown, B. (1992). An investigation of reported haunting occurrences in a Japanese restaurant in Georgia. In *Proceedings of Presented Papers, The Parapsychological Association 35th Annual Convention*, 151-168.
- Roll, W.G. & E. de A. Montagno (1983). Similarities between RSPK and psychomotor epilepsy. In W.G. Roll, J. Beloff & R.A. White (Eds.), *Research in parapsychology 1982* (pp. 270-271). Metuchen, N.J.: Scarecrow.
- Roll, W.G. & Pratt, J.G. (1971). The Miami disturbances. *Journal of the American Society for Psychological Research*, 65, 409-454.
- Roll, W.G., Sheehan, L.C., Persinger, M.A., and Glass, A.Y. (1996). The haunting of White Ranch. *The Parapsychological Association 39th Annual Convention*, 294.
- Roll, W.G. & Stump, J. (1969). The Olive Hill poltergeist. *Proceedings of the Parapsychological Association*, 6, 57-58.
- Roll, W.G. & Tringale, S. (1983). A haunting-type RSPK case in New England. In W.G. Roll, J. Beloff & R.A. White (Eds.), *Research in parapsychology 1982* (pp. 132-136). Metuchen, N.J.: Scarecrow.
- Rosenberg, R. (1974). A haunting in New Jersey. *Theta*, 39-40, 17-18.
- Ruttan, L.A., Persinger, M.A. & Koren, S.A. (1990). Enhancement of temporal lobe-related experiences during brief exposures to milliGauss intensity extremely low frequency magnetic fields. *Journal of Bioelectricity*, 9, 33-54.
- Solfvin, G., Harary B. & Batey, B. (1976). A highly publicized case of RSPK. *Journal of Parapsychology*, 40, 48-49.
- Solfvin, G. & Roll, W.G. (1976). A case of RSPK with an epileptic agent. In J.D. Morris, W.G. Roll & R.L. Morris & (Eds.), *Research in parapsychology 1975* (pp. 115-120). Metuchen, N.J.: Scarecrow.
- Stewart, J.L., Roll, W.G., & Baumann, S. (1987). Hypnotic suggestion and RSPK. In D.H. Weiner & R.D. Nelson (Eds.), *Research in parapsychology 1986* (pp. 30-35). Metuchen, N.J.: Scarecrow.
- Taff, B.E. & Gaynor, K. (1976). A new poltergeist effect. *Theta*, 4, 1-7.
- Taylor, J. (1975). *Superminds*. New York: Viking.
- Thacher, G.A. (1910). The March poltergeist case. *Journal of the American Society for Psychological Research*, 4, 561-636.
- Thurston, H. (1954). *Ghosts and poltergeists*. Chicago: Regner.

Wheeler, M.A., Stuss, D.T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autoegetic consciousness.

Psychological Bulletin, 121, 331-354.

Zorab, G. (1964). A further comparative analysis of some poltergeist phenomena: Cases from Continental Europe. *Journal of the American Society for Psychical Research*, 58, 105-127.

IS ESP A FORM OF PERCEPTION? CONTRIBUTIONS FROM A STUDY OF SEAN HARRIBANCE¹

William G. Roll
*Department of Psychology,
State University of West Georgia*

Michael A. Persinger,
*Neuroscience Laboratory,
Laurentian University, Ontario*

ABSTRACT

ESP shows several similarities to classical forms of perception, such as vision. Both are responses to objects in space-time, both incorporate implicit memory, and both focus on objects important to survival and well being. Like subliminal perception, ESP is usually associational rather than representational and multi-modal rather than modality specific (e.g., a visual target may be represented by auditory impressions). Structures in the temporal lobes, particularly the hippocampus and amygdala, may be involved in psi reception. The hippocampus, which accesses and processes long-term memory (the body's history in space-time), exhibits the associational and multi-modal features of ESP; and the amygdala displays its emotional characteristics. Research also suggests that ESP is primarily a function of the right brain hemisphere. A neuropsychological study of the psychic, Sean Harribance, who has shown significant psychic abilities in controlled tests, may provide further insights into the ESP process. Sean had difficulty acquiring language as a child, a deficit for which the development of ESP may have compensated. In any case, he advanced from dunce to the head of his class. The same year, when Sean was 12, he had his first conscious psychic impression. School problems resurfaced when he had to learn French and Latin, languages for which he had no memory, and he quit school after the sixth grade.

In the attempt to determine the nature of the ESP stimulus, in particular if it is composed of electromagnetic (EM) fields (to which the hippocampus and amygdala are particularly sensitive), EM pulses were directed to Sean's right hemisphere. This evoked experiences similar to those he reports during ESP and may suggest that the ESP stimulus to which Sean responds is composed of EM fields in space-time. The tests further indicate that the occipitoparietal interface in Sean's right hemisphere, which normally is concerned with the perception of visual gestalts, has assumed the function of psi detector for Sean. A structural change in this part of the cortex, probably initiated in childhood to compensate for cognitive deficits, may have expanded Sean's perceptual reach.

INTRODUCTION

From a neuropsychological perspective, changes in brain processes originate from an input or afferent function, or from an output or efferent function. Modifications in the input characteristics of a stimulus occur as the information moves from the sensors into the integration areas of the brain that are associated with the various forms of human behavior (Edelman, 1989). Changes in

¹Thanks for support from the Parapsychology Foundation, Inc., New York; the Sean Harribance Institute for Parapsychology, Trinidad; and the Sean Harribance Institute for Parapsychology Research, Inc., Texas.

output characteristics occur as information moves from a central locus towards different endpoints that may involve movement of an arm, in speech, in changes of heart rate or in emission of heat from the body. A similar distinction has been made between receptive psi or ESP and expressive psi or PK. This paper is about receptive psi.

Perceptual objects are located in (past) space-time. The concept of space-time as psychological reality is less foreign than it once was. Most people accept as a fact that the night sky displays the distant past of the galaxy, not the present (in relation to the observer's time frame). Because of the time it takes light to travel, the present cannot be seen, only the past, sensory processing adding further delays. Even at earthly distances what is seen is not, strictly speaking, what is there. The classical senses only show the near past, sufficient to navigate in the near environment. To perceive objects deeper in space-time and what these objects portend for the future, the organism may rely on receptive psi.

Some psi studies suggest causal reversal. For instance in the tests by Schmidt and Stapp (1993) the subjects evidently affected sequences of random numbers that had already been recorded. This effect may be limited to the quantum-mechanical level. According to Jahn and Dunne (1987), "In the physical domain, quantum effects become evident when the dimensions of the object of interest and of its relevant environment are of the same or smaller scale than the de Broglie wavelength of the wave functions that represent it; when the dimensions of the object and its environment are much larger than its associated wavelengths, classical continuum behavior obtains." (p. 263) In other words, ESP and PK that involve human-scale objects may follow classical rules.

The human world has three dimensions of space and one dimension of time. In this world, the arrow of time would point in one direction only, from past to future. As Hawking (1988) states, "This is the direction of time in which the universe is expanding rather than contracting." (p. 145). He also notes that the second law of thermodynamics points from past to future because disorder, entropy, increases with time. Finally, the psychological arrow of time points to the future: "This is the direction in which we feel time passes, the direction in which we remember the past but not the future." (p. 145). Hawking mentions string theory with its ten or 26 dimensions, but says that such a space would be curved up into a size too small to be experienced: "...we see only one time and three space dimensions, in which space-time is fairly flat...It seems clear then that life, as we know it, can exist only in regions of space-time in which one time and three space dimensions are not curled up small." (pp. 163-65) In these regions of space-time, ESP may not differ significantly from classical sense perception since both would involve the apprehension of human-scale objects.

In this paper we take a closer look at ESP with the help of neuropsychological studies of the psychic, Lalsing (Sean) Harribance.

COGNITIVE ASPECTS OF SEAN HARRIBANCE'S LIFE

Lalsing Harribance is of East Indian descent and was born in the village of Masabood Junction, Trinidad, the West Indies, November 11, 1939. He converted to Christianity from Hinduism when he was a young man, and took the name Sean, which is how he likes to be addressed. His father owned a shop in the village and ran other small businesses. Sean was the fourth child of seven.

Sean had difficulties with the acquisition of language. His parents delayed sending him to the school in their village till he was seven, but even so he did not do well. After the second grade, his father sent him to another school hoping for an improvement, but there was little change. "The

teacher punished me and the children ridiculed me because I was so dull...My only memories from this time of my life involve punishment and feelings of loneliness, fear and inadequacy. I was pushed on with my class only because I was too big to be held back" (Harribance & Neff, 1976, pp. 22-23). Sean had similar problems at home. Although his grandfather was a leader in the Hindu temple the family attended, "The Hindu god had no attraction for me and I had great difficulty understanding the Indian language which was part of the religious ritual. I could not sing the songs and I did not understand the words used in Hindu worship" (p. 23). Then a change occurred, Sean learnt about Christianity in the new school. He liked the stories about Christ and vividly remembers the first picture he was shown of Christ. When the teacher said that Christ would help if one prayed to him, Sean made up a special prayer that he become a better student. The repetition of this prayer took on a mantra-like quality as the boy repeated it walking to school. It was a long walk, six miles each way, so there was much time for praying. In addition, Sean sometimes would go off into wooded areas to pray. "As I lay in bed at night I would talk with Christ, too." (p. 23) Sean was nine when he started praying and soon thereafter he began to improve at his studies. "A year later...I had become an average student. The year after that when I was eleven I was in the upper 20 percent of my class. And one year later I came in first in my class" (pp. 23-24).

Sean attributes his success at school to knowing the answers by ESP. Perhaps the same happened as with Shields' (1976) learning disabled children who excelled in telepathy. Shields suggests that the children may have responded telepathically to the mind of their teacher, they "...borrow' knowledge symbiotically in lieu of having knowledge themselves." (In Shields' study, and for Sean as well, the "telepathic sender" was in the same room as the child, so the success may have been due to sensory cues--a possibility that does not substantially affect Shields' suggestion.)

The same year that Sean's ESP apparently brought him to the head of his class, he had his first conscious psi experience. He was walking home from school with a friend when he had an impression of the boy's mother wearing a red dress with green stripes and cooking a special meal. The woman did not own such a dress but it turned out that his uncle had come on an unexpected visit and had brought the dress and the food. Other impressions about his schoolmates and a teacher followed. "Now everything was going my way; I was absolutely brilliant in my school work; I stood first in my class. Answers in classroom discussions and on written quizzes came to me automatically, so I could work very fast." (op. cit., p. 26).

During the same period another world opened up for the boy. When Sean was praying alone outside, "...I had the very strange and wonderful experience of feeling in tune with the wind and the flowers, the vegetation and the birds. I felt that life was pulsing in all things, and I could feel a mystical bond linking that life to mine...I continue to feel this mystical unity of life when I am alone and in prayer" (p. 24).

Sean had distanced himself from the Hindu religion, but his practice of saying a special prayer and picturing the figure of Christ was similar to Hindu mantra meditation and visualizing a chosen deity. Sean had two attributes of the Hindu guru, psychic awareness and spiritual experiences. But he had left Hinduism for a religion where both are often regarded not as signs of spiritual attainment but as the work of the devil, a problem he has often had to deal with.

Sean's academic progress did not last long. When French and Latin became part of the curriculum, "...I could not master them. My psychic ability provided no help for me." (p. 26). At the same time his father needed help running his store in the village, so Sean quit school after the sixth

grade. Instead of the long walk to school, he played cricket and soccer, favorite sports in Trinidad, excelling at both. He also played cards and became known as a card-shark. When he came to the U.S., Sean was still a card-shark but in the rather different setting of the ESP laboratory.

Sean says that his psi experiences occur as very quick images. He often hears a voice in the right ear and experiences pictures in the upper left visual field. The scenes come in rapid sequences like on a TV screen. He frequently senses a presence behind him. During a lecture or reading he feels as if a higher energy or entity is telling him what to say. While his clients often comment on Sean's ability to tell them about their homes and surroundings, he says he becomes lost or disoriented a few hundred meters from his own home.

Sean came to the U.S. in 1969 so that his abilities might be scientifically verified. Extensive testing was done at the Psychical Research Foundation and the Institute for Parapsychology in Durham, N.C. Following this period of research, Sean settled near Houston, Texas, as a professional psychic. By this time he was married, having met his wife Chris at the PRF where she was a volunteer. They have two children. Although Sean has a demanding practice, his desire for scientific understanding has persisted over the years and he has succeeded in setting up two foundations, in Trinidad and in Texas, to support this work.

Most of the neuropsychological studies of Sean were done at the Neuroscience Laboratory, Laurentian University, May 27-29, 1996, and June 2-6, 1997. During the latter period, a SPECT examination (single-photon emission computerized tomography) was performed by Dr. David Webster at Sudbury General Hospital. In addition, topographical EEG mapping was done by Cheryl Alexander, and ESP studies by Dr. John Palmer, both associated with the Institute of Parapsychology.

BRAIN FUNCTIONS SERVING RECEPTIVE PSI

Memory and the Hippocampus: Sean's inability to integrate French and Latin words into long term memory apparently had the effect of also leaving them beyond the range of his ESP. Like classical sense perception, ESP appears to be a synthesis of external stimuli and implicit memories. If the memory aspect is lacking, there can be no perception. Sean's telepathy of the minds of his French and Latin teachers drew a blank because his memory had no French and Latin words.

The role of memory in ESP has been dealt with before (e.g., Irwin, 1979; Edge et al. 1986, Ch. 8; Roll, 1966). It is perhaps most obvious in forced choice tests where the range of targets is committed to memory beforehand. If the subject cannot recall the targets during the test, there can be no evidence for ESP. There can also be no evidence if the subject only repeats the order of targets as first presented; forgetting is as important as remembering for any type of perception, receptive psi included. To succeed at psi, memories must at the same time be accessible and changeable. Experiments where the response consists of autonomic reactions without awareness (e.g., Braud, Shafer & Andrews, 1993, Dean, 1966; Radin, 1996) depend equally on memory. Only memory can enable the brain to classify stimuli as important or not and to respond only to the important ones. This unconscious process is suggestive of Sean's forced choice tests where his 10 responses took 9-10 seconds (Roll and Klein, 1972), reminiscent of the automatic manner he said answers came to him at school. The memory aspect of ESP is usually also evident when this is associated with a phenomenological representation, such as a visual image. Like other

psychics, Sean's readings (Roll et al., 1973) are restricted to his own mental images and vocabulary. Regardless of how learned his client, the images and words that come to Sean's mind are strictly his own.

We now know that the ESP response in the free response setting rarely corresponds to the target in the one-to-one fashion of familiar forms of perception. In dream and ganzfeld studies, the subject's mentation hardly ever provides an isomorphic or "photographic" representation of the target (e.g. Targ and Kutra, 1998). What is usually experienced are mental images that represent the meaning of the target to the subject rather than a copy. (The associational feature of ESP is absent from forced-choice tests because the responses are restricted to a set range.) In spontaneous ESP, the response may not even be in the perceptual mode of the target. A scene that would usually be represented in visual terms, such as an automobile accident involving a relative, may take the form of a call for help from the person that never occurred (e.g. Roll, 1966). Receptive psi is multimodal and associational, rather than mode-specific and representational.

The observation that receptive psi is composed of long-term memories and is often multimodal has led to the temporal lobes, the biological substrate for memory, as the likely receptor (Roll, 1966). In particular, ESP shows the earmarks of two structures in the temporal lobes, the hippocampus and the amygdala. The hippocampus processes memories, the amygdala meaning.

The output of the hippocampus is long-term memories and it is multimodal. The hippocampus has been compared to a memory index system because of its intrinsic structure as well as the manner in which it connects to the cortical manifold. Whereas specific items of information are associated with cortical space, abbreviated codes for access to these larger items are maintained within the hippocampal formation.

The entorhinal gyrus (parahippocampal cortex), is the primary input into the hippocampus and to the amygdala. These structures are considered multimodal because they are affected by all known forms of sensory transduction. For information from different sensory systems to be integrated into a common fabric of experience, the modality-specific information is translated into a common code. At least one manifestation of this common code consists of complex frequencies and patterns of frequencies of weak electromagnetic (EM) discharges from brain neurons (John, 1990).

The brain space accessed by the entorhinal-hippocampal system is regarded as four-dimensional. This is because long-term memories reflect the history of the organism in the three dimensions of space and the one dimension of time. The hippocampal map is therefore uniquely suited to the four-dimensional space where ESP seems to operate.

Activation of the entorhinal-hippocampal system depends on specific patterns of EM signals whose space-time configuration may be described as "burst firing." Burst firing opens the system to recruitment of neuronal aggregates within the cerebral cortices that represent previously consolidated information, in other words, conscious memories.

Tests of autonomic ESP suggest that the normal human brain unconsciously, and perhaps continuously, scans space-time for objects that may have a significant impact on the organism. Since the recognition of such objects would be accomplished by implicit memories, the process would involve the hippocampus (and the amygdala). For the normal individual in ordinary states

of consciousness, the process is unlikely to be reflected in awareness but may still affect behavior so that the object is avoided or approached. In altered states of awareness, such as REM dreams, and for psychics in the waking state, burst firing may open the cerebral cortices with the result that impressions associated with the target achieve phenomenological representation as mental images.

Meaning and the Amygdala: Memory is usually about emotionally meaningful objects. Meaning is what makes things memorable. Meaningless objects may impinge on the senses but are rarely remembered. Objects that advance a person's purposes receive positive meanings and those that oppose them negative meanings. In force-choice ESP tests, targets with positive meanings are often associated with positive scores, with deviations above chance expectancy, while targets that carry negative meanings are often associated with negative scores, with deviations below chance.

Spontaneous ESP is usually about important occurrences. Most often it is about events associated with anxiety and apprehension, such as sudden accident and death, befalling significant others (e.g. Sannwald, 1963; Schouten, 1979, 1981, 1982; Persinger, 1974, I, p. 72ff). Ullman (1977/1986) has found psychic dreams to mostly occur when "significant relations are threatened, impaired, or destroyed." Van de Castle (1994) reports that "Spontaneous psi dreams generally deal with highly charged topics and characters to whom the dreamer has important emotional ties" (p. 409). More than 50 percent of the dreams were about death, followed by non-fatal accident and injury. Seemingly trivial events, however, are not infrequent and have been noted in 21% of Persinger's cases (1974, I, p. 72) and also in 21% of Sannwald's (1963). Ehrenwald (1977/1986) refers to ESP about meaningless events as "flaw determined," (p. 722) the result of minor defects in the subject's perceptual defenses.

The centrality of emotional meaning for psi is evident also in experimental work. For instance, Dean (1966), in a test of autonomic ESP, used personal names as targets, half of which were of persons known to the subject and half of unknown individuals. The results were restricted to the known names. Similarly, the subjects of Radin (1996) only responded to targets showing erotic or violent themes, ignoring emotionally neutral themes.

While the hippocampus provides the cognitive map, emotional meaning is determined by the amygdala. This pair of almond-shaped structures is the body's principal organ for ascribing meaning to objects and to alert the body to the presence of important objects in the environment. (The ventral striatum or extended amygdala is involved in this process.)

Feelings of fear, anxiety and apprehension are signatures of the amygdala and they are common in receptive psi. Feelings related to aggression, such as depression, and feelings related sexual expression, such as affection, are also associated with the amygdala; they too are seen in receptive psi.

Psi researchers, not surprisingly, are usually more interested in fostering affection than antipathy in their subjects, more interested in positive than negative scores. Feelings of fondness, affection, and love are regarded as conducive to positive scores. Jahn (1996) regards "*a resonance*, or bond, or sharing of identity" (p. 35) between percipient and agent a main ingredient for results. When there is no agent, the successful percipient feels bonded with the physical target. Ehrenwald (1977/1986) concludes that "...in the absence of emotionally-charged attitudes and motivations, e.g., in the experimenter-subject, or agent-percipient relationship, psi phenomena are unlikely to

make their appearance." (726) This was true for our tests with Sean. Among the experimental team at the PRF his results were largely confined to tests where Judy Klein (Roll & Klein, 1972) or John Stump (Stump, Roll & Roll, 1970) were co-experimenters.

RESULTS OF NEUROPSYCHOLOGICAL STUDIES OF SEAN

We turn now to the question which structures in Sean's cortex may be engaged during ESP. Since Sean seemed to possess superior psychic abilities in comparison to the general population and lesser language skills, we expected to find evidence that the left hemisphere of his brain would be less developed or active than his right since the left brain processes language while the right may be involved in ESP (e.g. Broughton, 1976, 1977, 1984; Ehrenwald, 1987). We hoped to repeat the finding of an association between ESP and alpha EEG that we saw at the PRF (Morris et al., 1972). We expected that Sean's right brain would be more engaged during ESP tasks than his left and that this would show up in increased blood flow to the right hemisphere, hence our decision to use the SPECT procedure. In particular, we expected to find evidence of the engagement of the right temporal and parietal lobes since these structures enable the brain to perceive objects in space-time.

Finally, we anticipated evidence of complex partial epileptic (CPE) seizure activity because this has been observed in individuals who may possess above average ESP (Persinger and Makarec, 1993). The tests did not support the CPE hypothesis for Sean.

Our examination showed several other features that may be related to Sean's experiences and psychic performance. (1) There was evidence for mild hypofunction in the left prefrontal region in the form of deficits for right finger and foot agility; for right finger and toe graphaesthesia (the ability to discern symbols drawn with a stylus on the tip of the finger and toe by the examiner); for visuomotor coordination by the right hand (assessed by the speed with which the subject moves a stylus through a maze); and for the capacity to rapidly alternate between conceptual tasks. (2) There was evidence for hypofunction or an unusual structural feature in the left superior temporal gyrus in the form of deficits for association of phonemes with their printed equivalents and for the detection of words by the right ear (during a dichotic listening task). These deficits had undoubtedly contributed to Sean's difficulties in school; but the fact that he now reads and understands English at grade equivalents that exceed his formal education suggests that other cortical structures have taken over the critical language functions. (3) There was evidence for hypofunction or an unusual structural feature in Sean's right occipitoparietal region in the form of deficits in a visual orientation task and in copying a complex visuospatial pattern. These deficits are consistent with Sean's experience of being lost a short distance from home. On the other hand, Sean showed above average finger gnosis for a man his age (the ability to discern which finger is tactually stimulated). Since the cortical representation of finger gnosis is usually within the occipitoparietal interface, and since early damage to the cerebrum may result in a diminution of the function of the damaged area but enhanced capacity in areas adjacent to the damage, a structural change in this region of Sean's brain may have taken place. The unusual structures indicated in Sean's brain may account for the phenomenological patterns of his psychic experiences, assuming these coincide with periods of hypermetabolism. The probable altered structure (on the synaptic level) within the right parietal region would likely lead Sean to perceive normal sensory information differently than the non-psychic. During periods when these structures are activated, "flashes of images" are to be expected within the left visual field. The images would be experienced as ego-alien, intrusive or "psychic" because they would involve the co-activation of the left superior temporal gyrus. Stated alternatively, Sean would experience such images apart

from their left hemispheric equivalents and also apart from the sense of self that is associated with the left hemisphere. Whereas the average person would perceive a spatial pattern or object and then the "word image" associated with this pattern or object, psychics such as Sean may not make this association.

A structural change within the left superior temporal gyrus would be consistent with Sean's experience of hearing a voice in his right ear, with a sense of presence on Sean's left side or behind, and with the experience of images within the left visual field. The strong activation of this region may also evoke the sensation that an external entity is telling Sean what to say during his lectures. The unusual structural features of Sean's brain may be due to different arrangements of fibers within the splenium of the corpus callosum (which connects the two brain hemispheres) or within the adjacent dorsal hippocampal commissures.

The flashes of images seen by Sean indicate burst firing. Sean's EEG was consistent with this. While interhemispheric comparisons between the left and right temporofrontal regions were coherent (dominated by fast beta activity), the occipitotemporal comparisons were anomalous. Sean's left hemisphere displayed more frequent episodes of slow alpha rhythms than the right. Furthermore, the patterns within the right hemisphere were less symmetrical; a higher frequency was found to be superimposed upon the fundamental alpha. The discrepancy was conspicuous and suggested unusual input from the hippocampal/parahippocampal region. The unusual brain activity was consistent with Sean's experience of flashes of images in the upper left visual field.

When Sean was exposed to weak, complex EM fields that simulated the firing patterns of neurons, he sensed the same "presence" that he associates with ESP. This effect suggests that his brain is highly sensitive to complex weak (less than one microTesla) magnetic fields. The results raise the possibility that the ESP stimulus to which Sean responds may be composed of configurations of EM fields in space-time.

ALPHA PERCENTAGE DURING READINGS OF PHOTOGRAPHS

The alpha wave, which indicates reduced engagement of the classical senses and increased attention to internal states, may be conducive to receptive psi. In Morris' (1977/1986) overview of the evidence, alpha abundance was associated with high ESP scores, "especially for subjects preselected for expertise at the production of one or both." One of the subjects was Sean (Morris, et al., 1972). We hoped for a repetition of this finding in the present study.

During both visits to Laurentian University, ESP-EEG sessions were conducted. We decided to use a free response procedure with photos as targets since Sean has recently shown evidence for ESP with this procedure in informal tests. During each visit, Sean read ten people by touching their photographs while his EEG was measured over the occipital region (bipolar). The pictures were hand shuffled by an assistant and placed face down in open envelopes. During the session, the experimenter would hand Sean the envelopes, one at a time. Sean would insert his hand in the envelope, still in the presence of the experimenter, and give his impressions while touching the photo. Following this, all responses would be ranked by ten individuals who knew the target persons well (e.g., a sister or brother) according to how well the responses matched the targets. The accuracy of each comment for each target person was determined by a rank of 0 for unknown, 1 for a miss, 2 for mild similarity, 3 for moderate similarity, and 4 for impressions very specific to the individual. In the 1996 session, when the average accuracy for each of the ten photographs was correlated with the average alpha time, $r=0.53$ with an insignificant $p < .10$. The 1997 results were better. The data were analyzed according to 2-second intervals for a total of

1,159 intervals (38.6 minutes). The values were $F(4,1154)=6.61$, $p<.0001$. As can be seen from the bar graph (Figure 1) impressions judged as more accurate were associated with more alpha rhythms over the occipital lobe.

Sean's apparent ability to discern photos by touch is reminiscent of his above average results in the finger gnosis tests. It is interesting that the cortical representation of finger gnosis is within the occipitoparietal interface, the area of the cortex that may be central for Sean's ESP.

DISCUSSION

Receptive psi shows several similarities to classical sense perception. Both are responses to objects in space-time; both are cognitive processes that rely on implicit memory; and in both the hippocampus and the amygdala may be engaged. ESP may differ from familiar forms of perception mainly insofar as it is a response to objects obscured or more distant in space-time. The work with Sean fills in the picture. The tests showed that weak pulses of external EM evoked experiences similar to those he reported during ESP. This may suggest that the ESP stimulus is environmental EM. The tests also indicate that the occipitoparietal interface may have assumed the function of a psi detector for Sean. The apparent structural change in this part of the cortex, probably initiated in childhood to compensate for cognitive deficits, may have expanded Sean's perceptual horizon to his benefit and the benefit of his clients.

The research needs to be extended to other proven psychics. Perhaps ways may then be found to increase the awareness of non-psychics of space-time. The importance of ESP, however, is not so much what it tells of the past as past, but what it tells of the past as holding the key to the future. Receptive psi may be as significant to the survival and well-being of the individual and group as the other forms of perception.

REFERENCES

- Braud, W.G., Shafer, D., & Andrews, S. (1993). Further studies of autonomic detection of remote staring. New control procedure and personality correlates. *Journal of Parapsychology*, 57, 391-409.
- Broughton, R.S. (1976). Brain hemisphere specialization and its possible effects on ESP performance. In J.D. Morris, W.G. Roll, & R.L. Morris (Eds.), *Research in parapsychology 1975*. Metuchen, N.J.: Scarecrow Press.
- Broughton, R.S. (1976). Possible brain laterality effects on ESP performance. *Journal of the Society for Psychical Research*, 48, 384-399.
- Broughton, R.S. (1977). Brain hemispheric differences in psi-influenced reacted time. In J.D. Morris, W.G. Roll, & R.L. Morris (Eds.), *Research in parapsychology 1976*. Metuchen, N.J.: Scarecrow Press.
- Dean, E.D. (1966). Plethysmograph recordings as ESP responses. *International Journal of Neuropsychiatry*, 2, 439-446.
- Dixon, N. (1971). *Subliminal perception: The Nature of a controversy*. London: McGraw-Hill.
- Dixon, N.F. (1979). Subliminal perception and parapsychology: Points of contact. In *brain/mind and parapsychology*, B. Shapin and L. Coly (eds.). New York: Parapsychology Foundation.
- Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness*. New York: Basic Books.
- Edge, H. L., Morris, R.L., Palmer, J., and Keil, J. H. (1986). *Foundations of parapsychology: Exploring the boundaries of human capability*. Boston: Routledge & Kegan Paul.

- Ehrenwald, J. (1977/1986). Psi phenomena and brain research. In B.B. Wolman (Ed.), *Handbook of parapsychology*. Jefferson, N.C.: McFarland. (pp.716-729)
- Ehrenwald, J. (1984). Right-vs. left-hemispheric approach in psychical research. *Journal of the American Society for Psychical Research*, 78, 29-40.
- Harribance, S. & Neff H.R. (1976). *This man knows you*. San Antonio, Texas: The Naylor Company.
- Hawking, S. (1988) *A brief history of time*. New York: Bantam Books.
- Irwin, H. J. (1979). *Psi and the mind: An information processing approach*. Metuchen, NJ: Scarecrow Press.
- Jahn, R.G. (1996). Information, consciousness and health. *Alternative Therapies*, 2, 32-38.
- Jahn, R.G., & Dunne, B. (1987) *Margins of reality: The role of consciousness in the physical world*. New York, N.Y.: Harcourt Brace.
- John, E. R. (1990). Representation of information in the brain. In E. R. John (ed), *Machinery of the mind*. Boston: Birkhauser, pp. 27-56.
- Morris, R. L., Roll, W.G., Klein, J. & Wheeler, G. (1972). EEG patterns and ESP results in forced choice experiments with Lalsingh Harribance. *Journal of the American Society for Psychical Research*, 66, 253-268.
- Persinger, M. A. (1974). *The Paranormal: Part I; Patterns. Part II. Mechanisms and models*. M.S.S. Information Corporation: New York, N.Y.
- Persinger, M. A. & Makarec, K. (1993). Complex partial epileptic-like signs as a continuum from normals to epileptics: Normative data and clinical populations. *Journal of Clinical Psychology*, 49, 33-45.
- Radin, D.I. (1996). Unconscious perception of future emotions: An experiment in presentiment. In *Proceedings of Presented Papers, The Parapsychological 39th Annual Convention*, pp. 171-185.
- Roll, W.G. (1966). ESP and memory. *International Journal of Neuropsychiatry*, 2, 505-521.
- Roll, W.G., & Klein, J. (1972). Further forced-choice experiments with Lalsingh Harribance. *Journal of the American Society for Psychical Research*, 66, 103-112.
- Roll, W.G., Morris, R.L., Damgaard, J.A., Klein, J., & Roll, M. (1973). Free verbal response experiments with Lalsingh Harribance. *Journal of the American Society for Psychical Research*, 67, 197-207.
- Roll, W.G. & Persinger, M.A. (1998). Poltergeist and nonlocality: Energetic aspects of RSPK. Unpublished manuscript.
- Sannwald, G. (1963). On the psychology of spontaneous paranormal phenomena. *International Journal of Parapsychology*, 5, 274-292.
- Schmidt, H., and Stapp, H. (1993). PK with prerecorded random events and the effects of preobservation. *Journal of Parapsychology*, 57, 331-339.
- Schouten, S. A. (1979). Analysis of spontaneous cases as reported in Phantasms of the Living. *European Journal of Parapsychology*, 2, 408-455.
- Schouten, S. A. (1981). Analyzing spontaneous cases: A replication based on the Sannwald collection. *European Journal of Parapsychology*, 4, 9-48.
- Schouten, S.A. (1982). Analyzing spontaneous cases: A replication based on the Rhine collection. *European Journal of Parapsychology*, 4, 113-158.
- Sheilds, E. (1976). Severely mentally retarded children's psi ability. In W. G. Roll (Ed.), *Research in parapsychology 1975*. Metuchen, NJ: Scarecrow Press.

- Silberman, K. E. & Weingarten, H. (1986). Hemispheric lateralization of functions relating to emotions. *Brain and Cognition*, 5, 322-353.
- Stump, J.P., Roll, W.G., and Roll, M. (1970). Some exploratory forced-choice ESP experiments with Lalsingh Harribance. *Journal of the American Society for Psychical Research*, 64, 421-431.
- Targ, R. and Katra, J. (1998). *Miracles of mind*. Novato, CA: New World Library.
- Ullman, M. (1977/1986). Psychopathology and psi phenomena. In B.B. Wolman (Ed.), *Handbook of parapsychology*. Jefferson, NC: McFarland.
- Ullman, M. (1986). Vigilance theory and psi. Part II: Physiological, psychological, and parapsychological aspects. *Journal of the American Society for Psychical Research*, 375-392.
- Van de Castle, R. L. (1977/1986). Sleep and dreams. In B. B. Wolman (Ed.), *Handbook of parapsychology*. Jefferson, NC: McFarland.

RELATIONSHIP BETWEEN THE HYPNAGOGIC/HYPNOPOMPIC STATES AND REPORTS OF ANOMALOUS EXPERIENCES

Simon J. Sherwood¹

*Department of Psychology
The University of Edinburgh*

ABSTRACT

A whole range of anomalous experiences have been reported during the borderline hypnagogic or hypnopompic states which surround periods of sleep (e.g., Gurney, Myers, & Podmore, 1886; Rose, Hogan, & Blackmore, 1997). The question is whether these states are conducive to anomalous processes or agencies, whether normal features are being misinterpreted, or both. This paper outlines the main physiological and psychological features of these hypnagogic/hypnopompic states and considers some of the evidence to address this question.

Although the hypnagogic (and probably the hypnopompic) state has unique behavioural, electrophysiological and subjective characteristics, it is also highly variable and there are large individual differences (Rechtschaffen, 1994). During these borderline states, people can experience brief and vivid imagery in one or more sensory modalities, and also temporary paralysis. People may find these kinds of experiences rather puzzling and may be keen to find an explanation for them, particularly if they have not come across them before. Evidence for the possible conduciveness of these states to anomalous experiences is discussed with reference to experimental studies (e.g., ganzfeld), spontaneous cases and surveys. Evidence for the possible misinterpretation of hypnagogic/hypnopompic experiences is discussed with reference to experiences with anomalous interpretations found in different cultures (e.g., Old Hag attacks) which have similar phenomenology to sleep paralysis plus accompanying imagery. A number of features of hypnagogic/hypnopompic experiences are very similar to features of reported anomalous experiences, such as ESP, apparitions, and OBEs (Mavromatis, 1983, 1987).

It is concluded that hypnagogic/hypnopompic features may be both conducive to anomalous experiences and misinterpreted as involving anomalous processes or agencies (e.g., deceased persons). Either way, the experiences may be interpreted correctly or incorrectly. The interpretation may depend on the specific hypnagogic/hypnopompic features experienced, on individual knowledge and beliefs, and on the context in which the phenomena occur. Further research which addresses the decision-making processes involved in interpreting these kinds of experiences would also be useful.

What is being proposed then is that, although hypnagogic/hypnopompic imagery and sleep paralysis are relatively normal experiences, occasionally they may be influenced by anomalous processes (e.g., ESP) or may facilitate anomalous experiences. More attention to the stages, features, contents, and physiology of the hypnagogic/hypnopompic states may enable us to identify, perhaps with a greater degree of accuracy, if and when anomalous processes are operating.

¹

The author would like to acknowledge the financial assistance provided by the Economic and Social Research Council (ESRC) and the helpful comments made by referees and by colleagues at the Koestler Parapsychology Unit on an earlier version of this paper. Thanks also to Professor Robert Morris and Dr. Caroline Watt for general support.

INTRODUCTION

A whole range of anomalous experiences have been reported during the hypnagogic or hypnopompic states which surround periods of sleep. It is not uncommon for people to experience brief, vivid and often strange imagery or to find themselves temporarily unable to move or speak during these periods between wakefulness and sleep. These brief sensations and the temporary paralysis are known as 'hypnagogic' or 'hypnopompic' imagery and 'sleep paralysis',² respectively (American Sleep Disorders Association (ASDA), 1990).

A recent survey found that people who report more childhood experiences of hypnagogic/hypnopompic imagery or sleep paralysis also report a greater number of anomalous experiences during childhood or adulthood (Sherwood, 1997). More specifically, hypnagogic/hypnopompic imagery has been associated with reports of extrasensory perception (ESP), apparitions and communication with the dead, out-of-the-body experiences (OBEs), visions of past lives and experiences involving extraterrestrials (e.g., Glicksohn, 1989; Gurney, Myers, & Podmore, 1886; Leaning, 1925; McCreery, 1993; McKellar, 1957; Mavromatis, 1983, 1987; Spanos, Cross, Dickson, & DuBreuil, 1993). In addition to the above anomalous experiences, sleep paralysis has also been associated with reports of psychokinesis (PK), and near-death experiences (NDEs) (Baker, 1992; Green & McCreery, 1994; Rose & Blackmore, 1996; Rose, Hogan, & Blackmore, 1997; Spanos et al., 1993; Spanos, McNulty, DuBreuil, Pires, & Burgess, 1995).

The question is: are the hypnagogic/hypnopompic states conducive to anomalous processes and events or are normal hypnagogic/hypnopompic features being misinterpreted? Perhaps both statements are true? The aim of this paper is to consider evidence for each of these possibilities. Firstly, it is necessary to outline the characteristics of the hypnagogic/hypnopompic states and the features of some of the experiences which can occur within them.

Characteristics of hypnagogic/hypnopompic states

Most research, both experimental and survey-based, seems to have focused on the hypnagogic state (the period between wakefulness and sleep, i.e., just as you are falling asleep). Comparatively little research has been carried out on the hypnopompic state (the period between sleep and wakefulness, i.e., just as you are waking from sleep). Thus, this paper will focus mainly on the hypnagogic state. The hypnagogic state, like the sleep state, is fairly complex and contains a number of steps and stages (Mavromatis, 1983; Rechtschaffen, 1994). Hori, Hayashi, and Morikawa (1994) concluded that the sleep onset period is unique and cannot be accurately categorised as either waking or sleeping. It is very difficult to determine the precise point of falling asleep, except by using arbitrary criteria, because the transition is gradual, because the changes are not always synchronised and because there are large individual differences in when the changes occur (Lavie, 1996; Rechtschaffen, 1994).

During alert wakefulness, eye movements are fairly rapid and the normal EEG trace consists of irregular waves of high frequency (Bray, Cragg, Macknight, Mills, & Taylor, 1992). As a person

²

Hypnagogic/hypnopompic imagery and sleep paralysis can be possible symptoms of narcolepsy, a sleep disorder which is characterised by excessive sleepiness and is typically associated with cataplexy. However, it is important to note that such imagery and sleep paralysis may also occur in an isolated form independently of narcolepsy (ASDA, 1990).

relaxes or becomes drowsy, there is an increase in alpha activity (8-12 Hz) and eye movements become slower and less frequent (Parker, 1975; Rechtschaffen, 1994). In fact, the presence of slow eye movements (SEMs) is considered to be an extremely accurate indicator of hypnagogic mentation (Schacter, 1976). Stickgold and Hobson (1994) found that as the period of eyelid movement quiescence lengthens, i.e., as the eyelids move less, mentation becomes more dreamlike. This contrasts with the positive association between eyelid movements and dreamlike mentation during REM sleep. As a person passes through the hypnagogic period into the early stages of NREM sleep, there is a decline in alpha activity and a concomitant increase in slower theta activity (4-7 Hz) (Baddia, Wright, Jr., & Wauquier, 1994; Bray et al., 1992; Rechtschaffen, 1994). A person is typically considered to be asleep once they reach stage 2 sleep which is characterised by theta activity and the appearance of sleep spindles (Lavie, 1996; Rechtschaffen, 1994). During the transition from wakefulness to sleep there is also a decrease in muscle tone, a slowing of the heart and respiration rates, a reduction in blood pressure, and an increase in skin temperature (Mavromatis, 1983; Mavromatis & Richardson, 1984; Schacter, 1976). Upon awakening, these changes go in the opposite direction (Mavromatis, 1983).

During the hypnagogic/hypnopompic states, people can experience brief and vivid imagery or sensations in one or more different sensory modalities (e.g., Foulkes & Vogel, 1965; Hori et al., 1994; Mavromatis, 1987) or temporary paralysis (ASDA, 1990). Recall of hypnagogic imagery has been found to peak around the middle of standard Stage 1 sleep when the EEG mainly consists of theta activity (Hori et al., 1994).

Laboratory studies have also found that hypnagogic imagery and sleep paralysis can occur during sleep-onset REM periods (SOREMPs) and that isolated sleep paralysis is characterised by abundant alpha activity (Takeuchi, Miyasita, Sasaki, Inugami, & Fukuda, 1992; Takeuchi, Miyasita, Inugami, Sasaki, & Fukuda, 1994). SOREMPs have been associated with altered sleep schedules (Fukuda, 1994) which may predispose towards sleep paralysis (ASDA, 1990). Some consider REM sleep intrusions to be a necessary but not sufficient requirement for sleep paralysis (Takeuchi et al., 1992), though there is evidence to suggest that it is not inevitably associated with SOREMPs (see Ness, 1978). Although these SOREMP hallucinations are similar to other hypnagogic imagery, they seem to be more emotional and there is a greater awareness of the surroundings (Takeuchi et al., 1992, 1994).

Other features of the sleep onset period include a decreasing awareness of observing the contents of one's own mind, increased absorption, a loss of volitional control over mentation, inaccurate time perception, a reduction of awareness of the environment, and a reduction in reality-testing (Foulkes & Vogel, 1965; Mavromatis & Richardson, 1984; Rechtschaffen, 1994). Further features may also include the hypnagogic/hypnopompic speech phenomenon and sleep starts. The hypnagogic/hypnopompic phenomenon occurs when a person hears him/herself uttering words, which can be nonsensical or irrelevant, just as they are falling asleep or waking from sleep (McKellar, 1989; Mavromatis, 1987). Sleep starts are sudden brief muscle contractions in one or more parts of the body which occur at sleep onset (ASDA, 1990). Sleep starts are sometimes associated with hypnagogic imagery, such as illusory sensations of movement (Nielsen, 1992; Oswald, 1959).

Studies have shown that as one moves through the sleep-onset period, the amount of visual hypnagogic imagery tends to increase (Hori et al., 1994), it becomes more dream-like (Foulkes & Vogel, 1965; Stickgold & Hobson, 1994) and the image quality, vividness, luminosity and intensity of colour also increase (Mavromatis, 1987; Nielsen, 1992). McKellar (1989) suggests

that the form of hypnagogic imagery also changes from sequences of objects, faces or landscapes to more complex episodes or mini-plays.

Although the hypnagogic state (and probably the hypnopompic) has unique behavioural, electrophysiological and subjective characteristics (Hori et al., 1994), it is also highly variable and there are large individual differences (Rechtschaffen, 1994; Tart, 1969, p. 73). Before going on to consider possible relationships with anomalous experiences, it is necessary to consider the experiential features of hypnagogic/hypnopompic imagery and sleep paralysis.

Hypnagogic/hypnopompic imagery

The term 'hypnagogic imagery' was provided by Maury (1848) (cited by Mavromatis, 1987) and was defined in Warren's Dictionary of Psychology (1934) as:

"Imagery of any sense modality, frequently of almost hallucinatory character, which is experienced in the drowsy state preceding deep sleep."

The term 'hypnopompic imagery' was introduced by F. W. H. Myers (1904) who defined it as:

"pictures consisting generally in the persistence of some dream-image into the first moments of waking." (p.125).

Some writers distinguish between imagery which occurs in the hypnagogic and hypnopompic states (e.g., Glicksohn, 1989; McKellar, 1989) but others do not (e.g., Mavromatis, 1987; Mavromatis & Richardson, 1984). It is fair to say that both types are similar and so many features will apply to both. However, certain features or experiences seem to be more common in the hypnagogic than in the hypnopompic state, and vice versa. For this reason, I think it is useful to maintain the distinction.

It has also been speculated that we may enter the hypnagogic state and experience hypnagogic phenomena at times other than just prior to nocturnal sleep (Mavromatis, 1983, 1987; Tart, 1969, p.74). This has been supported by reports from a number of participants (McKellar & Simpson, 1954). Although hypnopompic imagery was originally defined as a persistence of dream imagery into wakefulness (F. W. H. Myers, 1904), it seems that they are not always continuations of dreams since they can begin after the sleeper has awoken (e.g. Leaning, 1925; Mavromatis, 1987). Thus, hypnopompic images may not necessarily be the result of REM sleep continuation.

It is not easy to distinguish hypnagogic and hypnopompic imagery from dream imagery. Any qualitative distinctions made will depend on the defining characteristics of hypnagogic/hypnopompic imagery and dreams (Mavromatis & Richardson, 1984), both of which would benefit from stricter definitions. However, there is some evidence that hypnagogic (and also hypnopompic) imagery tend to be more vivid and realistic, shorter, more passive, have less emotion, and also tend to be more disorganised and irrelevant (Foulkes & Vogel, 1965; McKellar, 1989; McKellar & Simpson, 1954; Mavromatis & Richardson, 1984). McKellar (1989) describes how:

"To use an analogy, dreaming resembles a lecture illustrated by slides which form part of it; hypnagogic imagery is more like a display of slides meant to illustrate some other lecture. Moreover, the slides have been mixed up, and follow one another in random." (p.103).

With hypnagogic experiences, there also seems to be a greater awareness of the true situation and more reality testing compared with dream experiences (Mavromatis & Richardson, 1984). Some people also claim to be able to generate and/or control their hypnagogic/ hypnopompic imagery to some extent (Mavromatis, 1987, p. 71). Necessary requirements seem to include a receptive attitude and passive volition.

Timing

It seems that episodes of hypnagogic/hypnopompic imagery often occur fairly sporadically although concentrated series of episodes can also occur (Mavromatis, 1987). Hypnagogic and hypnopompic images are typically very brief and dynamic and may last no more than a second or two (Nielsen, 1992).

Frequency

Generally speaking, hypnagogic imagery seems to be more common than hypnopompic imagery. Early surveys (e.g. Galton, 1883; Müller, 1848 cited by Mavromatis, 1987) estimated that about 2% of adults had experienced hypnagogic imagery of some form. More recent surveys have estimated that c. one-third (Leaning, 1925), 61-63% (McKellar & Simpson, 1954; McKellar, 1957) or c.75% (Glicksohn, 1989; Richardson, Mavromatis, Mindel & Owens, 1981; Sherwood, 1997) of people have experienced it on at least one occasion compared with 21.4% (McKellar, 1957) to 67.6% for hypnopompic imagery (Richardson et al., 1981; Sherwood, 1997). One recent survey estimated that 37% (12.5%) of the UK population had experienced some form of hypnagogic (hypnopompic) imagery at least twice a week during the preceding year (Ohayon, Priest, Caulet & Guilleminault, 1996).

Sensory modalities

Although visual and auditory seem to be the most common forms of both hypnagogic and hypnopompic imagery (Foulkes & Vogel, 1965; Hori et al., 1994; McKellar & Simpson, 1954), olfactory (smell), gustatory (taste), tactile, thermal, bodily, movement, and synesthetic sensations (where imagery in one modality triggers modality in a different modality) may also occur (e.g., Leaning, 1925; Mavromatis, 1987; Schacter, 1976).

Visual imagery typically occurs with the eyes closed though it can occur with eyes open (Gurney et al., 1886; Leaning, 1925; McKellar, 1972, 1989; McKellar & Simpson, 1954). It often begins with reports of clouds or mists of bright colours or a circle of light. Images may quickly change from one to another and may develop into progressively more complex images (Gurney et al., 1886; Leaning, 1925; Mavromatis, 1987). Occasionally the images may be in black and white rather than in colour (McKellar, 1957; McKellar & Simpson, 1954). The images may sometimes be very small (micropsias) or gigantic (megalopsias), though changes in size and shape are possible (Leaning, 1925; McKellar, 1957; McKellar & Simpson, 1954; Mavromatis, 1987). A series of continuous repetitions (polyopsias) of the same image may also be experienced. Sometimes the images appear to be strangely illuminated or may be seen from a peculiar angle (Leaning, 1925; Mavromatis, 1987; McKellar, 1957).

Visual hypnagogic/hypnopompic imagery is often pleasant, even humorous, but it can also be terrifying (McKellar & Simpson, 1954; Mavromatis, 1987). Although hypnagogic and hypnopompic imagery are characterised by variety, Mavromatis (1987) has modified Leaning's (1925) classification scheme and identified six recurrent themes: (1) Formless, e.g. waves, clouds

of colour, (2) Designs, e.g. geometric and symmetrical patterns and shapes, (3) Faces, figures, animals, objects, (4) Nature scenes, e.g. landscapes, seascapes, gardens, (5) Scenes with people, (6) Print and writing, e.g. in real or imaginary languages. Visual hypnagogic/hypnopompic imagery has often been referred to as ‘the faces in the dark phenomenon’ because the seeing of faces is so common (McKellar, 1957).

Mavromatis (1987) provided a summary of auditory hypnagogic/hypnopompic phenomena:

“Auditory hypnagogic [and also hypnopompic] phenomena include the hearing of crashing noises, one’s name being called, a doorbell ringing, neologisms [new words or expressions], irrelevant sentences containing unrecognizable names, pompous nonsense, quotations, references to spoken conversations, remarks directed to oneself, meaningful responses to one’s thought of the moment.” (p.81).

Other reported imagery includes music, bangs and explosions. Sometimes auditory hypnopompic imagery can take the form of a warning of impending danger or an important event; other times it may just be a feeling of foreboding (Mavromatis, 1987).

Sensations of smell (e.g. cigars, roses), taste (e.g. sweet), sensations of actively touching or being passively touched by someone or something, and hot or cold sensations, sometimes moving along the body, have also been reported during the hypnagogic and hypnopompic states (Mavromatis, 1987)..

A sensation of falling seems to be the most commonly reported sensation of movement. The experience is relatively common and is often associated with a bodily jerk and visual imagery, such as falling off a cliff (Oswald, 1959). Other sensations of movement may include floating, swinging, rocking, spinning, and being in or on a moving vehicle. Bodily sensations may include, for example, a feeling of energy flowing through the body, weightlessness, heaviness, tingling, numbness, shaking/vibrating, and elongation of the body. A sense of presence has also been reported (e.g. Ohayon, Priest, Caulet, & Guilleminault, 1996) and may coincide with both imagery and sleep paralysis (Conesa, 1995; Hufford, 1982; Rose & Blackmore, 1996; Spanos et al., 1993, 1995). Feelings of foreboding or being under threat, or more general feelings of confusion and disorientation, have also been reported.

Mavromatis (1987, p. 28) also points out that “very often hypnagogic [and hypnopompic] images are symbolic or metaphoric, and not infrequently autosymbolic, and therefore not always meaningless.” The experient may sometimes become aware of the significance of the imagery (which may be known only by that person) during the experience or just afterwards. Such awareness is a characteristic of the ‘autosymbolic phenomenon’, described by Silberer (1965), which is an experience in which one’s thoughts or feelings at a given moment are translated into a symbolic form of imagery. However, Mavromatis (1987) adds that “awareness of the significance of the symbolism is not always present, and in the majority of cases imagery remains a puzzle until one begins to pay attention to it and enters into a form of ‘conversation’ with it.” (p. 59).

It is clear that hypnagogic/hypnopompic imagery may be extremely vivid, is characterised by variety, may evoke both positive and negative emotions, may occur in more than one sensory modality, sometimes simultaneously, and may have some significance for the experient.

Sleep paralysis

Definition

According to ASDA (1990):

“Sleep paralysis consists of an inability to perform voluntary movements either at sleep onset (hypnagogic or pre-dormital form) or upon awakening either during the night or in the morning (hypnopompic or post-dormital form).” (p.166).

Timing

The sleep paralysis episode usually lasts from a few seconds to a few minutes although a duration as long as 70 minutes has been reported (Goode, 1962; Schneck, 1960; Spanos et al., 1995). The experience is sometimes preceded by and/or accompanied by visual imagery which can be terrifying (Blackmore, 1996; Conesa, 1995; Goode, 1962; Penn et al., 1981; Spanos et al., 1995; Takeuchi et al., 1992). The experience may end due to efforts to overcome it by the experient, may be terminated by someone else either by touch or verbally, the experience may move into a dream (possibly lucid) or the episode may simply terminate spontaneously (Firestone, 1985; Goode, 1962; Schneck, 1960; Snyder, 1983).

Frequency

The hypnagogic form seems to be more frequent than the hypnopompic form (Conesa, 1995; Goode, 1962; Spanos et al., 1995) although the opposite has also been reported (Penn, Kripke & Scharff, 1981; Sherwood, 1997). The International Classification of Sleep Disorders estimates that isolated sleep paralysis (i.e. that which occurs independently of narcolepsy) occurs at least once in a lifetime in 40-50% of normal people. Surveys have found that between 4.7% to 49% of people have reported sleep paralysis (of one or other or both forms) though the most frequent estimates range between 30-45% (Goode, 1962; Everett, 1962; Penn et al., 1981; Spanos et al., 1995; Blackmore, 1996; Rose & Blackmore, 1996; Sherwood, 1997).

Features

Accompanying features of sleep paralysis may include: acute anxiety or terror, awareness of the surroundings, hypnagogic/hypnopompic imagery, a sense of presence, difficulty breathing or a sense of suffocation, pressure on the chest, a tendency to mentally or physically struggle to overcome it, a feeling of time distortion, sexual arousal (Blackmore, 1996; Conesa, 1995, 1997; Goode, 1962; Hufford, 1982; Liddon, 1967; Penn et al., 1981; Rose et al., 1997; Schneck, 1960, 1977; Snyder, 1983; Spanos et al., 1995; Takeuchi et al., 1992):

Interpretation of hypnagogic/hypnopompic experiences

Personal beliefs and expectations, knowledge of normal sleep-related experiences, mental set and the setting in which hypnagogic/hypnopompic experiences take place are all very important factors which can influence how these experiences are interpreted (e.g., Leaning, 1925; McKellar & Simpson, 1954; Mavromatis, 1987). People may find hypnagogic/ hypnopompic experiences rather puzzling and may be keen to find an explanation for them, particularly if they have not come across such experiences before, as this letter illustrates:

" I'm writing about a recurring experience of mine in the hope that a reader might be able to offer an explanation. It happens when I'm asleep or half asleep - although it's so real at the time that I'm convinced I'm awake. My whole body buzzes or tingles, like a bad case of 'pins and

needles'. I feel really scared, trapped and unable to move or speak. Sometimes I think I see or hear something. With concentration of strength I can escape from this state, although I always think I'm losing control. This only happens when I fall asleep on my back - which I now try to avoid! No-one I've spoken to has had the same experience. Any ideas?"

(Doubleday, 1996, p.53).

Mavromatis (1987) illustrates how, in some cases, 'visual psi experiences are practically indistinguishable from those occurring in hypnagogia both in their content and in their nature. Also, the mental state of the subject appears to be the same.' (p. 138). If psi does exist, then it seems that hypnagogic/hypnopompic experiences may be a vehicle for it but clearly there is room for misinterpretation. On a given occasion, a hypnagogic/hypnopompic experience may reflect genuine anomalous processes but it may or may not be interpreted as such. On another occasion, a hypnagogic/hypnopompic experience may not reflect genuine anomalous processes but it may be interpreted as if it does. Thus, false negatives and false positives are possible. Mavromatis (1987) does not really give any indication of how psi and non-psi hypnagogic/hypnopompic experiences might be differentiated. If reliable differences could be identified then this might help the experient make better judgements. This might be particularly useful in a ganzfeld context. Further investigation of this question is warranted.

Evidence for conduciveness of hypnagogic/hypnopompic states to anomalous processes

So what evidence is there to suggest that hypnagogic/hypnopompic states are conducive to anomalous processes or agencies? The hypnagogic state is considered to be 'unusually receptive' (Schacter, 1976, p. 468) and shares features of the psi-conducive state such as physical relaxation, reduction in sensory distraction and increased internal attention (Braud & Braud, 1975; Honorton, 1977; Mavromatis, 1987). According to Mavromatis (1983), support for a relationship between psi and hypnagogia (his generic term for hypnagogic and hypnopompic imagery) comes from the practices and literature on occultism and spiritualism, the literature on controlled psi experiments, and spontaneous cases of psi during hypnagogic practices.

Some experimental studies have found that hypnagogic/hypnopompic imagery seems to be conducive to telepathy (Gertz, 1983; Schacter & Kelly, 1975), perhaps more so than dreaming (Braud, 1977; White, Krippner, Ullman & Honorton, 1971). The ganzfeld technique, which has provided some of the best evidence for ESP, is believed to induce a hypnagogic-like state (Bertini, Lewis, & Witkin, 1969). However, although the ganzfeld procedure is often assumed to induce a hypnagogic state, the extent to which it does resemble the naturally-occurring hypnagogic state is not clear (Braud, Wood, & Braud, 1975; Schacter, 1976). If the experimentally-induced state is not radically different to the naturally-occurring state at sleep onset, then this suggests that hypnagogic experiences can occur at different times of the day and in different settings. Palmer, Bogart, Jones & Tart (1977) report a ganzfeld study (and a previous unpublished study, by Palmer) which found significant correlations between ESP z performance (psi-hitting or psi-missing) and scoring on altered state of consciousness/hypnagogic imagery scales.

Bem and Honorton (1994) reviewed ganzfeld studies carried out up until the Hyman (1985) and Honorton (1985) meta-analyses, and also a group of subsequent studies carried out by Honorton which were designed to meet the more stringent standards specified by Hyman and Honorton (1986). Bem and Honorton (1994) concluded that Honorton's studies had met the required standards and had produced results consistent with the previous ganzfeld studies, though they also acknowledged the need for further replication by a broader range of investigators. Milton and

Wiseman (1997) conducted a meta-analysis of other ganzfeld studies reported since the joint Hyman-Horton guidelines were published and concluded that these studies did not offer any evidence for ESP, though they suggested that certain psi-conducive factors may not have been exploited as effectively in these studies. Although previous ganzfeld research has provided some of the best evidence for ESP, Milton & Wiseman (1997) concluded that "what is clear is that use of the ganzfeld alone is no guarantee of a successful study." (p. 277).

In terms of spontaneous cases, there are a number of well-documented cases of ESP and crisis apparitions which have occurred during the hypnagogic/hypnopompic states (Gurney et al., 1886, Chapter IX, "Borderland" Cases, pp. 251-285). For example, a father reported a hypnopompic image involving his son:

"I was suddenly awake by hearing his voice, as I fancied, very near me. I saw a bright, opaque, white mass before my eyes, and in the centre of this light I saw the face of my little darling, his eyes bright, his mouth smiling. The apparition, accompanied by the sound of his voice, was too short and too sudden to be called a dream : it was too clear, too decided, to be called an effect of imagination. So distinctly did I hear his voice that I looked around the room to see whether he was actually there." (p.277).

The father received a letter the following day informing him that his son was ill but later found out that he had died at the time of the apparition. Some writers believe that hypnagogic visions might be an early form of ESP (Leaning, 1925; Mavromatis, 1983). In support of this, developing psychics often experience an increase in hypnagogic phenomena (Mavromatis, 1987). Gifted subjects also use hypnagogic imagery (White, 1964); for example, well-known psychics, such as Garrett and Northage, have described personal examples of telepathy and clairvoyance during the hypnagogic and hypnopompic states (Mavromatis, 1987).

Moody with Perry (1993) also described a number of cases of visionary encounters with departed loved ones inside a psychomanteum chamber which share characteristics of hypnagogic imagery. OBEs also tend to occur spontaneously during the hypnagogic/ hypnopompic states (Mavromatis, 1983); McCreery (1993) found a positive relationship between number of hypnagogic imagery episodes and number of OBEs. Mavromatis (1983, 1987) also lists a number of hypnagogic phenomena which have been reported during OBEs: e.g., sensations of floating, sinking, drifting, seeing lights, images, landscapes, hearing noises, music, name being called, sensation of being touched.

Rose et al. (1997) found that sleep paralysis has been associated with reports of a number of anomalous experiences, such as ESP, PK, OBEs, NDES, apparitions, past life experiences, and extraterrestrials. It is possible that anomalous experiences and sleep paralysis episodes might be related to each other because they might both be affected by a third variable such as the earth's geomagnetic field. A number of studies have found that subjective (e.g., Persinger, 1985) and experimental GESP experiences (Berger & Persinger, 1991) and sleep paralysis (Conesa, 1995, 1997) tend to occur when geomagnetic activity is relatively low.

If the hypnagogic state is conducive to anomalous processes perhaps this could be due to the initial increase in alpha or the later increase in theta activity (Healy, 1986) which occurs during this period (e.g., Davis, Davis, Loomis, Harvey, & Hobart, 1938; Baddia et al., 1994). Experienced meditators have also been found to show the alpha-theta progression which characterises the transition through the hypnagogic state towards sleep (Mavromatis, 1987).

Meditation has been associated with reports of a variety of anomalous experiences (Eysenck & Sargent, 1993).

There is evidence that alpha activity might be conducive to anomalous processes and experiences. Tart (1968) found that Miss Z's OBEs tended to occur during stage 1 sleep which was dominated by alphoid activity. Laboratory episodes of sleep paralysis have also been found to be characterised by abundant alpha activity (Takeuchi et al., 1992). In terms of theta, Stanford and Stevenson (1972), cited by Healy (1986), found some evidence to suggest that lower EEG activity, such as theta, might facilitate telepathy performance. Unusual theta activity has also been found in individuals who report mediumship ability or OBEs (Nelson, 1970; Tart, 1967, 1968; Palmer, 1979; all cited by Healy, 1986).

It has been suggested that it is not so much altered states per se which are psi-conducive but the degree or the rapidity of the transition from one state to another (e.g., Murphy, 1966; Honorton, 1973; Honorton, Davidson, & Bindler, 1971; all cited by Parker, 1975). Physiological monitoring of participants in the hypnagogic/hypnopompic states might be useful in that it could potentially identify the precise point, or at least the optimal physiological conditions, at which psi processes might operate. However, it is recognised that this might be difficult to investigate experimentally given that, for many people, hypnagogic/hypnopompic imagery and sleep paralysis may occur only sporadically and tend to be rather involuntary when they do occur. But, if the ganzfeld conditions are indeed very similar to the naturally-occurring hypnagogic state then this might be a suitable compromise.

In summary, evidence for the conduciveness of the hypnagogic/hypnopompic states to anomalous processes comes from the fact that these states have physiological and psychological features believed to be psi-conducive in other contexts, from experimental studies using both naturally-occurring and induced states, from spontaneous case reports of a variety of different phenomena, and from biographical accounts of gifted subjects and psychics.

Evidence for misinterpretation of hypnagogic/hypnopompic experiences

Hypnagogic/hypnopompic experiences may also have been misinterpreted as ESP, apparitions, visions of previous lives or other worlds, alien abductions, witchcraft or attacks by evil spirits or demons etc. (Baker, 1992; Blackmore, 1996; Dahlitz & Parkes, 1993; Hufford, 1982; Leaning, 1925; Liddon, 1967; McKellar, 1957, 1989; Spanos et al., 1993, 1995; Wilson & Barber, 1983; Wing et al., 1994; Zusne & Jones, 1989). Such experiences may initiate or sustain beliefs in the paranormal and the supernatural and may have contributed to mythology and folk-lore (Fukuda et al., 1987; Liddon, 1967; McKellar & Simpson, 1954; Mavromatis, 1983; Ness, 1978). Blackmore and Rose (1996) found that many people were scared by sleep paralysis and some were worried that they were going mad or being visited by supernatural entities. It is also possible that knowledge and beliefs may influence the content of hypnagogic/hypnopompic experiences (Hufford, 1982; Spanos et al., 1993).

Assuming that the hypnagogic/hypnopompic experiences do not reflect anomalous processes, are there any general characteristics of the hypnagogic/hypnopompic states which might facilitate misinterpretations, regardless of individual knowledge, beliefs and expectations? There may be reduced sensory input from the environment and some ambiguity of external stimuli, especially if the person is in bed and it is dark. This may interfere with accurate reality-testing (Spanos et al., 1993) which also tends to reduce during the sleep-onset period (Foulkes & Vogel, 1965). It is

sometimes difficult, subjectively, to distinguish wakefulness from sleep (Rechtschaffen, 1994); false awakenings and lucid dreams are a good illustration of this (Green & McCreery, 1994). One of the features of hypnagogic/ hypnopompic imagery which may lead people into believing in their reality and veracity may be that they feel awake throughout. In their survey of hypnagogic experiences, McKellar & Simpson (1954) found that

'Among the reasons given for believing oneself to be awake were: being able to have ordinary perception at the same time (the commonest reason given); being able to have separate thoughts; being able to engage in conversation; being able to open eyes, close them, and continue with the image, etc.' (p. 270).

Awareness of the surroundings may be reduced during the hypnagogic/hypnopompic states to some extent (Foulkes & Vogel, 1965; Rechtschaffen, 1994) but it is an important feature, particularly during sleep paralysis episodes (e.g., Conesa, 1995; Goode, 1962; Liddon, 1967; Schneck, 1960). Mavromatis (1987) points out that 'The 'sense of reality, of life-likeness' pointed out by many subjects in reference to their hypnagogic imagery often expands into 'feelings of *heightened reality*'.' (p. 30). Visual imagery can sometimes contain more detail than one might observe in more usual circumstances (Leaning, 1925). Hypnagogic/ hypnopompic imagery and sleep paralysis are also spontaneous vivid, realistic, intense and often frightening (Conesa, 1995; Schacter, 1976; Zusne & Jones, 1989). The unfamiliarity and involuntary nature of the imagery might facilitate external attributions.

Evidence for possible misinterpretation of normal hypnagogic/hypnopompic experiences is perhaps best illustrated by example. The Old Hag attack, well-known in the Canadian province of Newfoundland, is believed, by some people, to be caused by a supernatural creature, by a human in spirit form (e.g. a witch) or a combination of the two (Firestone, 1985; Hufford, 1982; Ness, 1978). The main features of an Old Hag attack are an impression of wakefulness and an accurate perception of the real environment, paralysis and fear; secondary features, which may be experienced with eyes open or closed, include a sense of presence, imagined sounds, visual images of a human (e.g. an old woman) or non-human attacker, a sense of motion, pressure (e.g. on the chest), difficulty breathing, odours and other bodily sensations (Hufford, 1982). Experiences with similar phenomenology have also been reported in Japan and China. These experiences are known as 'kanashibari' and 'ghost oppression attacks', respectively, and are believed by some to be caused by evil spirits or possession by a ghost (Fukuda et al., 1987; Wing et al., 1994). However, there is evidence to suggest that such beliefs may be more common among people who have not had the experiences themselves (Wing et al., 1994).

It is also possible that hypnagogic/hypnopompic imagery and sleep paralysis may account for some intense UFO reports and abductions (e.g., Baker, 1992; Spanos et al., 1993). Abductions are often reported around the time of sleep and may feature paralysis, awareness of surroundings, a sense of presence, bright lights and figures in the room, humming and buzzing sounds, and sensations of floating (e.g., Baker, 1992; Spanos et al., 1993).

Visual hypnagogic/hypnopompic imagery might also facilitate interpretations in terms of ghosts or apparitions. Faces may be experienced which range from the beautiful and the pleasant to the hideous and the terrifying (Leaning, 1925; McKellar, 1957; Mavromatis, 1987). These faces are often characterised as being extremely life-like and often seem to be looking at the observer (Gurney et al., 1886; Leaning, 1925). These faces can also develop into figures which may move towards the observer. Such faces/figures can be singular or in groups, of known or unknown,

living or dead persons and may sometimes seem to represent particular moods and emotions (Leaning, 1925; Mavromatis, 1987). The experient may also hear their name being called which might be interpreted as attempts at communication by deceased persons.

Features which are similar to OBE/NDE accounts include feelings of floating or weightlessness, sensations of movement, changes in body image, awareness of the surroundings, and experiencing visual images such as land/seascapes, faces/figures (perhaps from an unusual angle).

There are also hypnagogic/hypnopompic features which might facilitate ESP interpretations. Hypnopompic imagery, in particular, tends to anticipate forthcoming daily events, and in connection with actual later events it may be considered to be precognitive (Zusne & Jones, 1989). Hypnopompic imagery may also appear to be warning of imminent or future danger. Visual imagery involving complex scenes characterised by movement and life may also be experienced (Leaning, 1925; Mavromatis, 1987).

In summary, it seems possible that features of the hypnagogic/hypnopompic states can facilitate possible confusions between reality and imagination in some instances. There are also a number of specific features which may facilitate anomalous interpretations. This may be more likely if a person has little knowledge of normal hypnagogic/hypnopompic features and/or if a person is within a group or culture which has certain explanations for particular phenomena, or if the person already believes in anomalous phenomena.

In conclusion, it is possible that the hypnagogic/hypnopompic states may be both conducive to, and also misinterpreted as involving, anomalous processes and agencies. In the absence of more objective information, the decision as to which interpretation is taken may depend on the individual and the context in which the experiences take place. Further research which addresses the decision-making processes involved in interpreting these kinds of experiences would be useful. Ideally, more naturalistic, experimental testing of participants who regularly experience hypnagogic/hypnopompic experiences is required. More attention to the hypnopompic state and the extent to which experimentally-induced states are physiologically and psychologically equivalent to the naturally-occurring states would also be beneficial. More attention to the physiology, features and content of the hypnagogic/ hypnopompic states may enable us to identify, with a greater degree of accuracy, if and when anomalous processes are operating.

REFERENCES

- American Sleep Disorders Association. (1990). International classification of sleep disorders: Diagnostic and coding manual (ICSD). Rochester, MN: Author.
- Baddia, P., Wright, Jr., K. P., & Wauquier, A. (1994). Fluctuations in single-Hertz EEG activity during the transition to sleep. In R. D. Ogilvie, & J. R. Harsh (Eds.), Sleep onset: Normal and abnormal processes (pp. 201-218). Washington, DC: American Psychological Association.
- Baker, R. A. (1992). Alien abductions or human productions? Some not so unusual personal experiences. Unpublished manuscript.
- Bem, D. J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. Psychological Bulletin, 115, 4-18.

- Berger, R. E., & Persinger, M. A. (1991). Geophysical variables and behaviour: LXVII. Quieter annual geomagnetic activity and larger effect size for experimental psi (ESP) studies over six decades. Perceptual and Motor Skills, 73, 1219-1223.
- Bertini, M., Lewis, H. B., & Witkin, H. A. (1969). Some preliminary observations with an experimental procedure for the study of hypnagogic and related phenomena. In C. T. Tart (Ed.), Altered states of consciousness (pp. 93-111). New York: John Wiley & Sons.
- Blackmore, S. (1996, September). On the edge of reality. Paper presented at the British Association Annual Festival of Science, Birmingham, England.
- Braud, W. (1977). Long-distance dream and presleep telepathy. In J. D. Morris, W.G. Roll, & R.L. Morris (Eds.), Research in Parapsychology 1976 (pp. 154-155). Metuchen, NJ: Scarecrow Press.
- Braud, W. G., & Braud, L. W. (1975). The psi-conductive syndrome: Free response GESP performance following evocation of 'left-hemispheric' vs. 'right-hemispheric' functioning. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), Research in Parapsychology 1974, (pp. 17-20). Metuchen, NJ: Scarecrow Press.
- Braud, W. G., Wood, R., & Braud, L. W. (1975). Free response GESP performance during an experimental hypnagogic state induced by visual and acoustic ganzfeld techniques: A replication and extension. Journal of the American Society for Psychical Research, 69, 105-114.
- Bray, J. J., Cragg, P. A., Macknight, A. D. C., Mills, R. G., & Taylor, D. W. (Eds.) (1992). Lecture notes on human physiology (2nd ed.). Oxford, England: Blackwell.
- Conesa, J. (1995). Relationship between isolated sleep paralysis and geomagnetic influences: A case study. Perceptual and Motor Skills, 80, 1263-1273.
- Conesa, J. (1997). Isolated sleep paralysis, vivid dreams and geomagnetic influences: II. Perceptual and Motor Skills, 85, 579-584.
- Dahlitz, M., & Parkes, J. D. (1993). Sleep paralysis. Lancet, 341, 406-407.
- Davis, H., Davis, P. A., Loomis, L., Harvey, N., & Hobart, G. (1938). Human brain potentials during the onset of sleep. Journal of Neurophysiology, 1, 24-38.
- Doubleday, J. (1996, August). Scares in my sleep. [Letter]. Fortean Times, 89, 53.
- Everett, H. C. (1962). Sleep paralysis in medical students. Journal of Nervous and Mental Disease, 3, 283-287.
- Eysenck, H. J., & Sargent, C. (1993). Explaining the unexplained: Mysteries of the paranormal. London: PRION.
- Firestone, M. (1985). The "Old Hag:" Sleep paralysis in Newfoundland. The Journal of Psychoanalytic Anthropology, 8, 47-66.
- Foulkes, D., & Vogel, G. (1965). Mental activity at sleep onset. Journal of Abnormal Psychology, 70, 231-246.
- Fukuda, K. (1994). Sleep paralysis and sleep-onset REM period in normal individuals. In R. D. Ogilvie, & J. R. Harsh (Eds.), Sleep onset: Normal and abnormal processes (pp. 161-181). Washington, DC: American Psychological Association.
- Fukuda, K., Miyasita, A., Inugami, M., & Ishihara, K. (1987). High prevalence of isolated sleep paralysis: Kanashibari phenomenon in Japan. Sleep, 10, 279-286.
- Gertz, J. (1983). Hypnagogic fantasy, EEG, and psi performance in a single subject. Journal of the American Society for Psychical Research, 77, 155-170.
- Glicksohn, J. (1989). The structure of subjective experience: Interdependencies along the sleep-wakefulness continuum. Journal of Mental Imagery, 13, 99-106.

- Goode, G. B. (1962). Sleep paralysis. Archives of Neurology, 6, 228-234.
- Green, C., & McCreery, C. (1994). Lucid dreaming: The paradox of consciousness during sleep. London: Routledge.
- Gurney, E., Myers, F. W. H., & Podmore, F. (1886). Phantasms of the living. London: Trübner, 2 vols.
- Healy, J. (1986). Hippocampal kindling, theta resonance, and psi. Journal of the Society for Psychical Research, 53, 352-368.
- Honorton, C. (1977). Psi and internal attention states. In B. B. Wolman (Ed.), Handbook of parapsychology (pp. 435-472). New York: Van Nostrand Reinhold.
- Honorton, C. (1985). Meta-analysis of psi ganzfeld research: A response to Hyman. Journal of Parapsychology, 49, 51-91.
- Hori, T., Hayashi, M., & Morikawa, T. (1994). Topographical EEG changes and the hypnagogic experience. In R. D. Ogilvie, & J. R. Harsh (Eds.), Sleep onset: Normal and abnormal processes (pp. 237-253). Washington, DC: American Psychological Association.
- Hufford, D. J. (1982). The terror that comes in the night: An experience-centered study of supernatural assault traditions. Philadelphia: University of Pennsylvania Press.
- Hyman, R. (1985). The ganzfeld psi experiment: A critical appraisal. Journal of Parapsychology, 49, 3-49.
- Hyman, R., & Honorton, C. (1986). A joint communiqué: The psi Ganzfeld controversy. Journal of Parapsychology, 50, 351-364.
- Lavie, P. (1996). The enchanted world of sleep. New Haven: Yale University Press.
- Leaning, F. E. (1925). An introductory study of hypnagogic phenomena. Proceedings of the Society for Psychical Research, XXXV, 287-411.
- Liddon, S. C. (1967). Sleep paralysis and hypnagogic hallucinations: Their relationship to the nightmare. Archives of General Psychiatry, 17, 88-96.
- McCreery, C. (1993). Schizotypy and out-of-the-body experiences. Unpublished doctoral thesis. University of Oxford, England.
- McKellar, P. (1957). Imagination and thinking: A psychological analysis. London: Cohen & West.
- McKellar, P. (1989). Abnormal psychology: Its experience and behaviour. London: Routledge.
- McKellar, P. & Simpson, L. (1954). Between wakefulness and sleep: Hypnagogic imagery. British Journal of Psychology, 45, 266-276.
- Mavromatis, A. (1983). Hypnagogia: The nature and function of the hypnagogic state. Unpublished doctoral thesis. Brunel University, England.
- Mavromatis, A. (1987). Hypnagogia: The unique state of consciousness between wakefulness and sleep. London: Routledge & Kegan Paul.
- Mavromatis, A., & Richardson, J. T. E. (1984). Hypnagogic imagery. International Review of Mental Imagery, 1, 159-189.
- Milton, J. & Wiseman, R. (1997). Ganzfeld at the crossroads: A meta-analysis of the new generation of studies. Proceedings of the Parapsychological Association 40th Annual Convention held in conjunction with The Society for Psychical Research, Brighton, England, 267-282.
- Moody, R. with Perry, P. (1993). Reunions: Visionary encounters with departed loved ones. London: Little, Brown and Company.

- Myers, F. W. H. (1904). Human personality and its survival of bodily death (Vol. I.) New York: Longmans, Green, & Co..
- Ness, R. C. (1978). The Old Hag phenomenon as sleep paralysis: A biocultural interpretation. Culture, Medicine and Psychiatry, 2, 15-39.
- Nielsen, T. A. (1992). A self-observational study of spontaneous hypnagogic imagery using the upright napping procedure. Imagination, Cognition and Personality, 11, 353-366.
- Ohayon, M. M., Priest, R. G., Caulet, M., & Guilleminault, C. (1996). Hypnagogic and hypnopompic hallucinations: Pathological phenomena? British Journal of Psychiatry, 169, 459-467.
- Oswald, I. (1959). Sudden bodily jerks on falling asleep. Brain, 82, 92-103.
- Palmer, J., Bogart, D. N., Jones, S. M., & Tart, C. T. (1977). Scoring patterns in an ESP Ganzfeld experiment. Journal of the American Society for Psychical Research, 71, 121- 145.
- Parker, A. (1975). States of mind: ESP and altered states of consciousness. New York: Taplinger.
- Penn, N. E., Kripke, D. F., & Scharff, J. (1981). Sleep paralysis among medical students. Journal of Psychology, 107, 247-252.
- Persinger, M. A. (1985). Geophysical variables and behavior: XXX. Intense paranormal experiences during days of quiet, global, geomagnetic activity. Perceptual and Motor Skills, 61, 320-322.
- Rechtschaffen, A. (1994). Sleep onset: Conceptual issues. In R. D. Ogilvie & J. R. Harsh (Eds.), Sleep onset: Normal and abnormal processes (pp. 3-18). Washington, DC: American Psychological Association.
- Richardson, J. T. E., Mavromatis, A., Mindel, T., & Owens, A. C. (1981). Individual differences in hypnagogic and hypnopompic imagery. Journal of Mental Imagery, 5, 91-96.
- Rose, N. & Blackmore, S. (1996, August). Two pilot surveys of unusual personal experiences. Paper presented at the 20th International Conference of the Society for Psychical Research, Cirencester, England.
- Rose, N., Hogan, J. & Blackmore, S. (1997). Experiences of sleep paralysis. Proceedings of the Parapsychological Association 40th Annual Convention held in conjunction with The Society for Psychical Research, Brighton, England, 358-369.
- Schacter, D. L. (1976). The hypnagogic state: A critical review of the literature. Psychological Bulletin, 83, 452-481.
- Schacter, D. L., & Kelly, E. F. (1975). ESP in the twilight zone. Journal of Parapsychology, 39, 27-28.
- Schneck, J. M. (1960). Sleep paralysis without narcolepsy or cataplexy. Journal of the American Medical Association, 173, 1129-1130.
- Sherwood, S. J. (1997). Relationship between childhood hypnagogic/hypnopompic experiences, childhood fantasy proneness and anomalous experiences and beliefs: An exploratory WWW survey. Proceedings of the Parapsychological Association 40th Annual Convention held in conjunction with The Society for Psychical Research, Brighton, England, 370-386.
- Snyder, S. (1983). Isolated sleep paralysis after rapid time zone change ('jet lag') syndrome. Chronobiologia, 10, 377-379.
- Spanos, N. P., Cross, P. A., Dickson, K., & DuBreuil, S. C. (1993). Close encounters: An examination of UFO experiences. Journal of Abnormal Psychology, 102, 624-632.
- Spanos, N. P., McNulty, S. A., DuBreuil, S. C., Pires, M. & Burgess, M. F. (1995). The frequency and correlates of sleep paralysis in a university sample. Journal of Research in Personality, 29, 285-305.

- Stickgold, R. & Hobson, J. A. (1994). Home monitoring of sleep onset and sleep-onset mentation using the Nightcap©. In R. D. Ogilvie & J. R. Harsh (Eds.), Sleep onset: Normal and abnormal processes (pp. 141-160). Washington, DC: American Psychological Association.
- Takeuchi, T., Miyasita, A., Sasaki, Y, Inugami, M., & Fukuda, K. (1992). Isolated sleep paralysis elicited by sleep interruption. Sleep, 15, 217-225.
- Takeuchi, T., Miyasita, A., Inugami, M., Sasaki, Y, & Fukuda, K. (1994). Laboratory-documented hallucination during sleep-onset REM period in a normal subject. Perceptual and Motor Skills, 78, 979-985.
- Tart, C. T. (Ed.) (1969). Altered states of consciousness. New York: Wiley & Sons.
- 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.
- White, R. A., Krippner, S., Ullman, M., & Honorton, C. (1971). Experimentally-induced telepathic dreams with EEG-REM monitoring: Some manifest content variables related to psi operation. In W. G. Roll, R. L. Morris, & J. D. Morris (Eds.), Proceedings of the Parapsychological Association, Number 5, 1968 (pp. 85-87). Durham, NC: Parapsychological Association.
- Wilson, S. C. & Barber, T. X. (1983). The fantasy-prone personality: Implications for understanding imagery, hypnosis, and parapsychological phenomena. In A. A. Sheikh (Ed.), Imagery: Current theory, research, and application (pp. 340-387). New York: John Wiley & Son.
- Wing, Y. K., Lee, S. T., & Chen, C. N. (1994). Sleep paralysis in Chinese: Ghost oppression phenomenon in Hong Kong. Sleep, 17, 609-613.
- Zusne, L., & Jones, W. H. (1989). Anomalistic psychology: A study of magical thinking (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

DREAM GESP STUDY II USING DYNAMIC VIDEO-CLIPS: INVESTIGATION OF CONSENSUS VOTING JUDGING PROCEDURES AND TARGET EMOTIONALITY

Sherwood¹, S. J., Dalton, K., Steinkamp, F., & Watt, C.

*Department of Psychology,
The University of Edinburgh,
Scotland, UK.*

ABSTRACT

This study was a partial replication and extension of a previous study (Dalton, Steinkamp, & Sherwood, in press). It investigated whether individual versus small group consensus judging procedures, and/or the emotionality of the dynamic target video clips, would affect the frequency of correct identification of the target in a free-response dream GESP study.

Two people located in Edinburgh, Scotland (Edinburgh participants) and a third person located in Derby, England (remote participant) acted both as experimenters and as participants. On each of the 28 trial nights, a one minute video clip, which was randomly-selected from 72 possible target clips, was shown repeatedly in a locked room in the Department of Psychology between 3.00-4.30am. Participants slept in their respective homes and recorded their dreams upon awakening. The following morning the participants viewed all four video clips from the selected target pool (i.e., 3 decoys plus the target) in a random order and then judged the correspondences between the clips and their dream mentation.

The Edinburgh participants obtained a greater number of direct hits using their group judgements than by using their own individual judgements. The discussion method was also marginally more successful than the objective consensus method. As hypothesised, using the discussion consensus judging method, the Edinburgh participants scored 12 direct hits during the 28 trials which is significantly higher than the mean chance expectation (exact binomial $p= 0.0294$, $ES(h)= 0.38$). The objective consensus judging method yielded 11 direct hits which is approaching a significant deviation from the mean chance expectation (exact binomial $p= 0.0679$, $ES(h) = 0.30$). As individuals, none of the participants obtained a significant score: the Edinburgh participants both scored 7 direct hits (exact binomial $p= 0.5721$, $ES(h)= 0.00$), the remote participant scored 9 direct hits (exact binomial $p= 0.2499$, $ES(h)= 0.16$).

Participants, both as a group and as individuals, obtained a greater proportion of direct hits when the target was emotionally negative than when it was either positive or neutral. The direct hit rate for negative emotion targets (5 direct hits during 8 trials) was significantly higher than the mean chance expectation for both group judging methods and for one of the individuals (FS) (exact binomial $p= 0.0273$, $ES(h)= 0.79$) which partially supported our hypothesis.

Post hoc analyses involving correlations between the global aa index of the Earth's geomagnetic field and target rankings were also reported.

¹

We would like to acknowledge the Economic and Social Research Council (ESRC) for their financial support of Simon Sherwood, the Society for Psychical Research and the Institut fur Grenzgebiete der Psychologie und Psychohygiene for their support of Kathy Dalton, and Fundação Bial for their support of Fiona Steinkamp. Also, Drs. Jessica Utts and John Palmer for statistical advice.

INTRODUCTION

A large proportion, estimated to be 33-65% (Van de Castle, 1977), of spontaneous cases of extrasensory perception (ESP) have been reported during dreams (e.g., Gurney, Myers, & Podmore, 1886). Such cases often feature information about close relatives or friends and often feature negative life events (Ullman, Krippner with Vaughan, 1989; Van de Castle, 1977). Thus, emotional, often negative contents are fairly common.

Some experimental research has found that dreams (e.g., Braud, 1977; Child, Kanthamani, & Sweeney, 1977; Dalton et al., in press; Ullman et al., 1989; Van de Castle, 1971) and hypnagogic/hypnopompic imagery (Braud, 1977; Gertz, 1983; White, Krippner, Ullman, & Honorton, 1971) seem to be conducive to ESP. Indeed, some studies have suggested that dreams may be more conducive to ESP than ganzfeld conditions (Kanthamani, Khilji, & Rustomji-Kerns, 1989; Kanthamani & Khilji, 1990; Kanthamani & Broughton, 1992). The Maimonides dream studies (Ullman et al., 1989), which mainly investigated dream telepathy, are undoubtedly the most well-known and most successful (Child, 1985). A recent meta-analysis of 450 Maimonides dream telepathy sessions found the overall hit rate to be 63% (mean chance expectation (MCE)= 50%) with odds against chance of 75 million to one (Radin, 1997). Attempted replications of the Maimonides dream studies have not been so successful, though the conditions surrounding these attempts may not have been particularly conducive (see Van de Castle, 1977).

Dynamic targets

The dream ESP studies, and later the ganzfeld studies, represented a move away from the early forced-choice methods towards free-response methods. Most dream GESP studies have used static target materials, such as art prints or slides, though some studies have used more dynamic target materials such as projector slides with an accompanying soundtrack, a person performing mimes, or films/videos. Other free-response GESP research has suggested that dynamic and multi-sensory targets might be more conducive to ESP than static targets (Dalton & Utts, 1995; Delanoy, 1989; Honorton, Berger, Varvoglis, Quant, Derr, Schechter, & Ferrari, 1990). The present study used dynamic video clips with accompanying soundtracks.

Emotionality of targets

Spontaneous cases of dream GESP frequently contain emotional information which is often negative. Evidence from experimental studies of GESP suggests that emotional target materials are more conducive than neutral materials (Bierman, 1995, 1997; Moss & Gengerelli, 1968 cited by Gelade & Harvie, 1975; Radin, 1997). Other writers have agreed with the potential conduciveness of emotional target material (e.g., Dalton, 1997a; Ullman et al., 1989; Van de Castle, 1977; Watt, 1989). Krippner (1975) suggested that emotional stimuli are more effective in dreams than in non-dream experiments. However, it is not clear whether positive or negative emotional targets are more psi-conducive. Some studies have found that positive materials may have a larger effect than negative materials (Dalton, 1997b;

Radin, 1997); others have found the reverse to be true (Dalton et al., in press). This study attempted to address this question by comparing the direct hit rate for positive, negative and emotionally neutral target video clips. More specifically, it was hypothesised that the group would score significantly higher than the overall chance expectation when the targets were negative as this was the finding in a similar study carried out at Edinburgh (Dalton et al., in press).

Small group participation and consensus judgement methods

Some studies have used majority-vote or pooled rating/ranking procedures in order to try to maximise ESP performance (Fiske & West, 1956, 1957; Kennedy, 1979; Ryzl, 1966; Taetzsch, 1962; Thouless, 1960 - cited by Carpenter, 1991; Braud, 1977; Kanthamani et al., 1989; Kanthamani & Khilji, 1990). Carpenter (1995) has been using group consensus judgements of the contents of regular quasi-psychotherapeutic group meetings as a method of facilitating ESP. Ullman (1989) has also been conducting exploratory research with small groups to see if their dreams are conducive to ESP. A previous dream GESP study using a consensus vote judging procedure found that this method yielded a higher number of direct hits than individual judgements (Dalton et al., in press). The present study also aimed to explore whether a group consensus based on a discussion of each individual's judgements would be more or less successful than determining the consensus purely on the objective pooling of individual judgements.

To summarise, this study was a replication and extension of our previous study (Dalton et al., in press). The aim was to investigate whether dreams (and also hypnagogic/ hypnopompic imagery) might be conducive to general extrasensory perception. Further aims were to investigate whether individual versus small group consensus judging procedures and/or the emotionality of the dynamic target video clips would influence task performance.

Hypotheses for the present study were

- H₁ The direct hit rate for the group discussion consensus judging method would be significantly higher than the mean chance expectation.
- H₂ The direct hit rate for the negative emotion targets would be significantly higher than the mean chance expectation.
- H₃ The direct hit rate for judgements attributed to hypnagogic or hypnopompic imagery would be significantly higher than the mean chance expectation.

Exploratory investigations

1. To explore whether there would be a difference between the direct hit rates for two group judging methods: discussion consensus and objective consensus.
2. To explore whether the group direct hit rates would be higher than the individual's hit rates.
3. To explore whether the direct hit rates would differ according to the emotionality of the target.
4. To explore whether a participant who had obtained a significant direct hit rate in a previous dream GESP study (Dalton et al., in press) would be able to obtain similar

results from a remote location.

METHOD

Design

This study used a clairvoyance design (i.e., no sender), although possible precognition could not be ruled out. Twenty-eight trials were prespecified as part of a repeated measures design. There was one pilot trial on the 4th September 1997. Two trials during the course of the study had to be aborted¹. The pilot and the aborted trials were not included in the analysis. The experimental trials were carried out between 14th September and 13th December 1997. Typically, there were 2-3 trials per week.

The independent variables were (1) the judging method used (discussion consensus, objective consensus, individual) and (2) the emotionality of the target clips (positive, negative, neutral). There were two different participant locations, approximately 220 miles apart: Edinburgh, Scotland (SS, FS); Derby, England (KD).

The dependent variables were (1) the accuracy classification per trial (a direct hit or a miss), (2) the dream-mentation-video-clip correspondence rating per trial (1-99), and (3) the dream-mentation-video-clip correspondence ranking (1-4) per trial.

Experimenters and participants

Three of the authors (SS, KD, FS) acted both as experimenters and as participants. The fourth author (CW) did not act as a participant or experimenter. CW collated and looked after both the remote participant's (KD) dream mentation and judgements and stored a copy of the computer printout for each trial. CW was also responsible for checking the raw data and the statistical analyses. Two of the experimenters (SS, FS) set up the system for the selection and subsequent display of the target for each trial. These experimenters set up 14 trials each on a mostly alternate basis.

All three participants typically report that they experience more than one dream per night on a regular basis and are able to recall at least some of the contents the following morning. Typically, the participants reported 2-4 different dreams per night on trial nights.

Apparatus

The free-response testing system at the University of Edinburgh is designed to be used under a variety of experimental designs. It is a computer-based system that can provide automatic data recording, highly effective shielding against sensory cues, and resistance to both participant and intentional experimenter bias. The system has been used in a number of automated ganzfeld studies which have been conducted at Edinburgh (e.g. Dalton, 1997b). For more details of the equipment, the laboratory, and security measures, see, for example,

¹ On the first occasion, the experimenter discovered that the computer program had not been able to complete its task of showing the target clip because both VCR's had been switched off. On the second occasion, an experimenter error meant that the computer program moved onto the judging sequence before the four possible video clips had been shown.

Dalton et al. (in press), Dalton (1997b) or Dalton, Morris, Delanoy, Radin, Taylor, & Wiseman (1994).

Pseudo-random number generator

Following recommendations for the reporting of randomisation methods (Milton & Wiseman, 1997), the target pool and the target clip within the selected pool for each trial were determined using a pseudo-random number generator (pseudo-RNG). The pseudo-RNG was a computer algorithm. The program used the RANDOMIZE TIMER command to generate a seed number for the random number function, RND. Initially, the computer program generated a random number to determine which of the 18 pools of clips would be used for the forthcoming trial; the program then generated a second random number to determine which of the four clips within the selected target pool would be used as the target clip. The computer program also utilised this procedure to determine the order of presentation of the video clips for the judging process.

Global tests of the randomness of the pseudo-RNG output generated by this method on the same equipment have been carried out in the past at periodic intervals and have not revealed any significant deviations from the expected distribution (Dalton, 1997b). Global tests of the random number output for selection of the target pool (1-18) and the target clip (1-4) were also carried out after the data for this study had been collected. A test program was used to generate two series of 10000 random numbers, one series between 0-18 and one series between 0-4. A chi-square goodness of fit test (carried out on Microsoft Excel for Windows 95 version 7.0 using the CHITEST function) found that neither of these series deviated from the expected distribution to a significant extent.

Target materials

This study used 18 separate target pools each containing four video clips taken from films, TV programmes, and cartoons. Thus, the target clip on any given trial was one out of a possible 72 video clips. Each video clip was one minute in duration and had an accompanying soundtrack. This set of target clips has been used in a previous sender/no sender ganzfeld study (Morris, Dalton, Delanoy, & Watt, 1995). The remote participant had a duplicate set of video clips which were viewed on her home VCR and television set.

One of the authors (KD) had previously had the 72 video clips viewed and categorised in terms of its emotionality (i.e. positive, negative, neutral) by three independent judges (see Dalton, 1997b). There were 24 positive, 26 negative, and 22 neutral video clips according to these three judges.

PROCEDURE

Selection and displaying of the target video clip

The selection of the target clip was performed by a pseudo-random number generator (RNG) which was controlled by a Microsoft Quick Basic computer program. The displaying of the target clip during the night and the displaying of the complete target pool during the judging procedure was also controlled by the computer program.

Prior to each trial the dream study computer program had to be activated by one of the

Edinburgh experimenters (SS or FS). This was usually done towards the end of the working day to avoid possible disruption to other ongoing research. Trials set up during the weekend were done at the experimenters' convenience. In order to run the dream program, the experimenter needed to enter a password (known only to SS, FS, & KD).

The computer created a file on the hard disk in which the details of the trial would be recorded. Note that this file could not be accessed by the experimenter without disrupting the experimental set-up. The experimenter was therefore blind as to which pool of target clips and which target clip had been selected for the forthcoming trial. A 3.5" floppy disk was also placed into the floppy disk drive (a copy of the trial datafile would later be transferred to it). Having checked that all relevant equipment was switched on, the experimenter initiated the dream program.

The target clip was later shown a total of 20 times between approximately 3.00-4.30am (via the computer-controlled VCR) on the television monitor and stereo headphones in the Target room.

Recording of dream mentation

During each trial night the Edinburgh participants (SS, FS) and the remote participant (KD) slept at their respective homes and wrote down the contents of any dreams or hypnagogic or hypnopompic imagery that they could recall either throughout the night or the following morning.

Judging procedure

Edinburgh participants - The following morning, the two participants (SS, FS) came into the Koestler Parapsychology Unit's (KPU) experimental suite bringing with them their hand-written dream mentation reports.

The name of each video clip in the selected target pool appeared on the computer monitor in turn and each clip was then displayed on the TV monitor.² Once the names and the order of the four video clips were known, the experimenter telephoned the remote participant and communicated this information to her.

The Edinburgh participants then watched each of the video clips in the computer-designated order whilst reviewing their dream mentation. Participants were allowed to view any or all of the clips again. Each individual then gave a dream-mentation-video-clip correspondence rating from 1-99 to each clip ('99' indicating a perfect correspondence). The clips were then placed in rank order with a rank of '1' being given to the clip with the highest correspondence to the dream mentation, '2' to the next highest correspondence and so on. Thus, the lower

²

From trial 3 inclusive onwards, the television monitor remained switched off until after the order of the video clips in the target pool had been communicated to the remote participant. In trials 1 and 2, the TV monitor was switched on at the beginning which meant that the participants viewed each of the four clips before the order of clips had been communicated to the remote participant. The experimenters recognised the possibility that they might inadvertently communicate cues about the correspondence between clip contents and their own dream mentation during the telephone conversation with the remote participant and thus changed the procedure accordingly.

the rank assigned to the clip, the higher the correspondence with the dream mentation and the higher the expectation that it was the actual target. The participants then swapped their dream reports with each other (without having discussed or having mentioned their own dream mentation) and performed a similar rating and ranking procedure. Note that participants' judgements of their own mentation were recorded on separate pieces of paper so that they were blind as to the other participant's judgements at this stage. Participants were not permitted to go back and alter their own judgements.

Having read and judged the correspondence between the dream mentation and the four video clips in the target pool, the Edinburgh participants discussed their own and each other's mentation until they had reached a consensus regarding which clip they considered to have been the target. This was known as the 'discussion consensus' judgement method. The individual's 'own' and 'other' correspondence ratings and rankings were then collated and entered onto a record sheet (by SS) and checked by FS. The sum of the ranks given to each individual clip (based on both participants' 'own' and 'other' judgements) was calculated. The lowest sum of ranks was given a rank of '1', the next lowest a rank of '2' and so on. This was known as the 'objective consensus' judgement method. If there were any ties to the sums of ranks, the sum of the ratings was calculated for each tie and the highest sum of ratings was allocated the lower rank. If the sums of ratings tied, the range of ratings was calculated and the smaller range was given the lower rank. The objective consensus rankings were then entered into the computer. The individual ratings and rankings, the discussion consensus and the objective consensus rankings were recorded by hand on the record sheet.

Remote participant - In the morning, the remote participant (KD) entered her dream mentation into an e-mail message which was sent to CW via KD's connection to a public internet service provider. Note that this was carried out before KD was informed of the names of the clips for the trial. Once she had been informed of the names and the order of the video clips by the experimenter, the remote participant manually found the position of each clip on her duplicate videotape using her VCR and viewed each one on her TV set in the designated order. The remote participant was free to watch each clip more than once if she wished to do so. The remote participant then followed the same individual rating and ranking procedure as the Edinburgh participants. Once the judging procedure had been completed, the remote participant entered her correspondence ratings and rankings into an e-mail message which was then sent to CW. Thus, the remote participant kept a hand-written and a computer hard disk copy of her mentation and correspondence ratings and rankings. Once this had been completed, the remote participant awaited a telephone call from the experimenter.

Once the objective consensus rankings had been entered into the computer, the experimenter telephoned the remote participant and checked that she had completed her judging procedure and had e-mailed a copy of her mentation and judgements to CW.³ The experimenter then instructed the computer to show which of the video clips had been the target, whilst on the phone to the remote participant, so that all three participants obtained feedback regarding their performance simultaneously.

Once the trial was completed, the experimenter instructed the computer to save the trial data to a floppy disk and also to print out three hard copies of the trial datafile. One copy was

³

Although the remote participant also verbally informed the experimenter of her rankings of the four clips prior to the target being revealed, this was not actually recorded by the experimenter.

stapled to the record sheet and the hand-written reports from the Edinburgh participants and kept in a folder in FS's office in the experimental suite. The second copy was kept in SS's office; the third copy was kept in CW's office. Note that the hard and floppy disk files and the printouts contained only the objective consensus rankings. The remote participant's dream mentation and judgements were kept in the form of e-mail messages which were stored in CW's e-mail account. These e-mail messages were printed out as they came in.

Statistical analysis

A table of raw data for the individual and consensus judging methods was produced from the experimental record sheets and the printouts of the remote participant's judgements by the first author (SS). SS also noted which experimenter had set up each trial, the name of the target clip and the emotionality of the target clip. This raw data table was checked against a similar table produced by KD. The trial by trial raw data records from the remote and the Edinburgh participants, the exact binomial probabilities, effect size calculations, sum of ranks analyses and global tests of randomness for the pseudo-RNG were double-checked by CW.

Pre-planned analyses

The critical ratio (CR) test, which is usually used to calculate the probability of gaining at least a given number of direct hits (assuming the null hypothesis is true), was not used as it is not recommended if np (number of trials \times probability of a direct hit) is less than 10 (Palmer, 1986, p.148). Instead, the exact binomial probabilities for obtaining at least a given number of direct hits were calculated using the BINOMIST function of Microsoft Excel for Windows 95 version 7.0. The planned analysis of the direct hit rate for judgements attributed to hypnagogic/hypnopompic imagery (H_3) was not carried out due to insufficient data.

In terms of calculating the effect size, previous studies, including our own, have used Cohen's h (Cohen, 1977) to index the difference between two proportions (or hit rates). Calculations of Cohen's h and its 95% confidence limits were carried out by hand (by SS) using the formulae given by Rosenthal & Rosnow (1991, p.449).

Post-hoc analyses

It has been suggested that sum of ranks analyses might generally be expected to be a more sensitive measure of ESP performance than direct hits because it uses more information per trial (Milton, 1997). An inconclusive review of ganzfeld studies found that sums of ranks did outperform direct hits measures in terms of deviations from chance and effect sizes but not to a significant degree (Milton, 1997). A sum of ranks analysis was also carried out (Palmer, 1986) in this study as it may facilitate further consideration of this issue.

After the data had been collected, the authors also decided to investigate a possible relationship between dream GESP performance (in terms of rankings) and the earth's geomagnetic field (GMF) as they had done in a previous study (Dalton et al., in press). The daily average antipodal (aa) index, which is the average change in global GMF from midnight to midnight on the day in question, was selected as the GMF measure (data was obtained by KD once the study was completed). A Spearman's rho correlation coefficient was calculated for the relationship between the GMF measure and each of the objective-consensus and individual target rankings.

RESULTS

Group versus individual judging methods

Table 1 shows that using the discussion consensus judging method, the Edinburgh participants scored 12 direct hits (i.e., correct identification of the target) during the 28 trials (direct hit rate= .43 or 43%, mean chance expectation (MCE)= .25 or 25%). This is significantly higher than the mean chance expectation (exact binomial $p= 0.0294$ for ≥ 12 direct hits, $ES(h)= 0.38$, 95% C.I. (h)= 0.00–0.75). The objective consensus judging method yielded 11 direct hits during the 28 trials (direct hit rate= .39 or 39%) which is approaching a significant deviation from the mean chance expectation (exact binomial $p= 0.0679$ for ≥ 11 direct hits, $ES(h)= 0.30$, 95% C.I. (h)= -0.10–0.66). Thus, our hypothesis H_1 that the group would perform significantly better than the mean chance expectation using the discussion consensus judging method, was supported.

Table 1 also shows the two Edinburgh participants both scored 7 direct hits during the 28 trials (direct hit rate= .25 or 25%, MCE= .25 or 25%) which is clearly not higher than chance expectations (exact binomial $p= 0.5721$ for ≥ 7 direct hits, $ES(h)= 0.00$, 95% C.I. (h)= -0.44–0.34). The remote participant scored higher than the Edinburgh participants and obtained 9 direct hits during the 28 trials (direct hit rate= .32 or 32%) which is also non-significant (exact binomial $p= 0.2499$ for ≥ 9 direct hits, $ES(h)= 0.16$, 95% C.I. (h)= -0.25–0.50).

The Edinburgh participants obtained a greater number of direct hits by using their group judgements than by using their own individual judgements. The discussion consensus method also yielded a marginally greater number of direct hits than the objective consensus method.

	Correspondence rankings				Row total (sum of ranks)
	1	2	3	4	
Expected	7	7	7	7	28 (70)
Discussion	12	---	---	---	---
Objective	11	5	8	4	28 (61)
Edin. (SS)	7	13	4	4	28 (61)
Edin. (FS)	7	8	8	5	28 (67)
Remote	9	7	3	9	28 (68)

Note: a rank of 1 = a direct hit

Table 1: Number of direct hits and distribution of ranks assigned to the target clips by the Edinburgh participants (both as a group and as individuals) and by the remote participant.

Post hoc sum of ranks analyses (Palmer, 1986, p.148) gave the following results: objective consensus ($CR(z)= 1.44$, $p= 0.0749$, one-tailed); Edinburgh participant (SS) ($CR(z)= 1.44$, $p= 0.0749$, one-tailed); Edinburgh participant (FS) ($CR(z)= 0.42$, $p= 0.3372$, one-tailed); remote

participant (KD) ($CR(z) = 0.25, p = 0.4013$, one-tailed). These results show that, as with the dichotomous scoring method, the objective consensus method yielded a result which was approaching a significant deviation from the chance expectation. Using the sum of ranks analysis, the Edinburgh participants' performance appeared to be slightly better than it did using the dichotomous scoring method; however, the remote participant's performance appeared to be slightly worse using the dichotomous scoring method. All individual performances were still non-significant using the sum of ranks analysis method.

Emotionality of the target clips

Table 2 shows that the majority of the target clips in the study were emotional, either positive (57.1%) or negative (28.6%). However, both as a group and as individuals, participants obtained a greater proportion of direct hits when the target was negative than when it was either positive or neutral. The direct hit rate for negative emotion targets was significantly higher than the mean chance expectation for both group judging methods and for one of the individuals (FS) (exact binomial $p = 0.0273$ for ≥ 5 direct hits, $ES(h) = 0.79$, 95% C.I.(h) = 0.09–1.69) which partially supports H_2 .

	Emotionality of the target clips			Row total
	Positive	Negative	Neutral	
Times selected as target	16	8	4	28
Target type availability	24	26	22	72
% of category	16/24 = 66.7%	8/26 = 30.8%	4/22 = 18.2%	-----
% of total no. of trials	16/28 = 57.1%	8/28 = 28.6%	4/28 = 14.3%	100%
	No. of direct hits per emotionality type (%)			
Discussion consensus	6 (37.5%)	5 (62.5%)	1 (25.0%)	12
Objective consensus	5 (31.3%)	5 (62.5%)	1 (25.0%)	11
Edinburgh (SS)	4 (25.0%)	3 (37.5%)	0 (-----)	7
Edinburgh (FS)	2 (12.5%)	5 (62.5%)	0 (-----)	7
Remote (KD)	4 (25.0%)	4 (50.0%)	1 (25.0%)	9

Table 2: Number and proportion of direct hits by emotionality of the target clips

The direct hit-rates for the positive emotion targets were all non-significant. However, it is important to note that the different target types were not selected an equal number of times and there are differing numbers of each target type within the 18 pools of video clips and so the probability of each target type being selected was not equal.

Relationship with the Earth's geomagnetic field

The results of the post hoc analysis indicated that the correlations between the global aa index and target rankings were as follows (all two-tailed): group objective consensus $r_s = -.327$, $p = .090$; Edinburgh participant (SS) $r_s = -.163$, $p = .407$; Edinburgh participant (FS) $r_s = -.013$, $p = .947$; remote participant (KD) $r_s = .075$, $p = .706$. Thus, the Edinburgh participants, both as a group and as individuals, demonstrated a negative relationship between GMF and dream-GESP target rankings, i.e., as the Earth's geomagnetic field increased the dream-GESP performance increased. The remote participant demonstrated a small positive relationship. The relationships were all non-significant, although the relationship for the group objective-consensus judging method was approaching significance.

DISCUSSION

The results of this dream GESP study found that only the group consensus judging methods yielded a performance, in terms of number of direct hits, which was greater than chance expectations to a significant (discussion consensus, $p = 0.0294$, $ES(h) = 0.38$), or approaching significant (objective consensus, $p = 0.0679$, $ES(h) = 0.30$), extent. If ESP was operating in this study, the effect size was small; a value of $h = 0.20$ is considered to be a small effect, $h = 0.50$, a medium effect size (Rosenthal & Rosnow, 1991). Further exploration of the data revealed that, as individuals, neither the Edinburgh participants nor the remote participant scored significantly higher than chance expectations in terms of the number of direct hits. The group consensus judgements obtained a greater number of direct hits than the individual judgements which is a replication of our previous findings (Dalton et al., in press). The discussion consensus judging method also obtained a marginally greater number of direct hits than the objective consensus judging method.

Both the individual and the group judgements obtained a greater number of direct hits when the targets were emotional as opposed to neutral; more specifically, performance was best when the targets were negative. However, the emotionality results should be interpreted with some caution (see below).

A post hoc sum of ranks analysis also found that the objective consensus and individual scores did not reach significance, although the probabilities of the Edinburgh participant scores were much closer to the criterion level for significance using this technique. This provides some support for previous research which has suggested that sum of ranks analysis may be more sensitive than that based on direct hits and may yield greater deviations from chance and larger effect sizes (Milton, 1997).

Group versus individual judging methods

This study does provide some limited support for the occurrence of GESP during dreams but it seems that information received by more than one person was required in order to be able to make judgements which were more accurate than chance expectations. Our finding that consensus judging procedures might be conducive to ESP, perhaps more so than individual procedures, supports previous findings for both dream GESP (e.g., Braud, 1977; Dalton et al., in press; Kanthamani et al., 1989; Kanthamani & Khilji, 1990) and waking GESP (e.g., Fiske & West, 1956, 1957; Kennedy, 1979). There did not appear to be much advantage to be gained by discussing possible target-mentation correspondences in order to reach a consensus

as opposed to determining the consensus simply by pooling the independent individual judgements. The practicality of each of these consensus judging methods may depend on the choice of judging requirements. Whilst an objective consensus method could easily be used for either hit/miss or rank ordering scoring methods, the discussion consensus method takes much longer and is much more difficult.

So, why might group judgements be more successful than individual judgements? If ESP has a relatively weak effect then a combination of extrasensorially perceived information relating to the same target from more than one person may be required in order to boost the accuracy of target judgements to a significant level. However, looking at the success of a judging procedure which involves a single overall consensus call per trial, based on several individuals, is different from looking at the number of hits per trial based on the individual calls of several individuals per trial. In the latter case, the overall score may be partially due to a stacking effect. If more than one call is made per trial then the likelihood of a hit is increased because there is more than one chance of getting it correct. It must be recognised that our overall results may also have been partially biased by the stacking effect since we reported both individual and group performance which are not independent.

Emotionality of the target clips

Participants tended to perform more successfully with emotional as opposed to neutral targets which supports previous research (e.g., Bierman, 1995, 1997; Dalton et al., in press; Moss & Gingerelli, 1968; Radin, 1997). The fact that participants tended to perform better with negative as opposed to positive targets supports our previous dream GESP study (Dalton et al., in press) but is in contrast to other findings from ganzfeld and presentiment studies (Dalton, 1997b; Radin, 1997). Previous research has also found that spontaneous cases of dream ESP often feature negative life events (Ullman et al., 1989; Van de Castle, 1977). However, our findings regarding the emotionality of the target material should be treated with caution for these reasons: (1) the comparison is based on a small number of trials, (2) the different types of target emotionality were not expected or observed to appear an equal number of times throughout the study because they were not counterbalanced across the target pools, (3) the emotionality categorisations were determined by a small number of independent judges - it is not known whether the participants' categorisations would have been in agreement, (4) perhaps there was a participant response bias towards selecting negative clips as being the target.

So, why might sleep, and also emotional target material, be conducive to dream ESP? In evolutionary terms, some organisms are likely to be at risk whilst they are asleep. It has been suggested that the periodical appearance of REM sleep and dreams provides a vigilance mechanism, which is oriented to the present and the future, and also an anticipatory arousal mechanism which counteracts these risks (Tolaas, 1986; Ullman, 1990). The dream vigilance mechanism would come into operation during each sleep cycle and would result in awakening, if a potential threat was identified and was considered to be of sufficient importance, or a continuation of the sleep cycle and a return to non-REM sleep. There might be an ESP component within this mechanism which monitors information which is distant in terms of space and time (Ullman, 1986). It has further been suggested that, as society has developed, the dream vigilance mechanism has become focused on potential psychological rather than physical dangers and, in particular, to events which might disrupt connections with significant others (Ullman, 1986). If such a mechanism does exist, it would explain why many

spontaneous cases of dream ESP feature emotionally-close persons and negative life events. It would also explain why negative target materials may be more conducive to dream ESP in the laboratory than positive or neutral materials.

Relationship with the Earth's geomagnetic field

The negative relationship between the GMF index and the group's objective-consensus target rankings ($r_s = -.353$) was in the opposite direction to the relationship obtained in a previous Edinburgh dream-GESP study ($r_s = .224$). The finding in this study is also contrary to previous research which has found that dream ESP is more likely when the Earth's geomagnetic field is lower (e.g., Persinger & Krippner, 1989). However, recent research suggests that the relationship between GMF and free-response ESP performance may depend on the Local Sidereal Time (LST) at the time of the trials (Spottiswoode, 1997). This factor may explain the discrepancy with the previous study - further analysis would be required to address this issue.

CONCLUSION

This study was a fairly successful replication of our previous dream GESP study (Dalton et al., in press). However, the effect sizes for the small group judging were slightly lower in this study (discussion $ES(h) = .38$, objective $ES(h) = .30$) than in the previous study ($ES(h) = .46$). Effect sizes for individual judgements were also much lower in this study. However, the group direct hit rates per emotional target type were of a similar magnitude. Possible reasons for these differences, apart from the procedural differences and the use of a different target pool, may be that in this study the schedule for the judging periods was more variable and involved earlier morning meetings (sometimes around 7am) due to increased demands on the laboratory facilities. Another contributing feature may have been a lack of novelty or a change in the group dynamics. However, one of the positive aspects of this study, from the group participants' perspective at least, was that it maintained a relaxed and fairly informal atmosphere, features which have been found to be conducive in other small group dream ESP research (Ullman, 1989). It seems that home-based dream GESP research may continue to be a fruitful endeavour, particularly if it utilises emotional dynamic target material and consensus judging procedures.

REFERENCES

- Bierman, D. J. (1995). The Amsterdam ganzfeld series III & IV: Target clip emotionality, effect sizes and openness. Proceedings of the 38th Annual Convention of the Parapsychological Association, Durham, NC, 27-37.
- Bierman, D. J. (1997). Emotion and intuition I, II, III, IV & V: Unravelling variables contributing to the presentiment effect. Proceedings of the 40th Annual Convention of the Parapsychological Association in conjunction with The Society for Psychical Research, Brighton, UK, 49-62.
- Braud, W. (1977). Long-distance dream and presleep telepathy. In J. D. Morris, W.G. Roll, & R.L. Morris (Eds.), Research in Parapsychology 1976 (pp. 154-155). Metuchen, NJ: Scarecrow Press.
- Carpenter, J. C. (1991). Prediction of forced-choice ESP performance part III: Three attempts to retrieve coded information using mood reports and a repeated-guessing technique. Journal of Parapsychology, *55*, 227-280.
- Carpenter, J. C. (1995). Small group interaction as a medium for ESP response. Proceedings of the 38th Annual Convention of the Parapsychological Association, Durham, NC, 79-87.
- Child, I. L. (1985). Psychology and anomalous observations: The question of ESP in dreams. American Psychologist, *40*, 1219-1230.
- Child, I. L., Kanthamani, H., & Sweeney, V. M. (1977). A simplified experiment in dream telepathy. In J. D. Morris, W.G. Roll, & R. L. Morris (Eds.), Research in Parapsychology 1976 (pp. 91-93). Metuchen, NJ: Scarecrow Press.
- Cohen, J. (1977). Statistical power analysis for the behavioral sciences (rev. ed.). New York: Academic Press.
- Dalton, K. (1997a). Is there a formula to success in the ganzfeld? Observations on predictors of psi-ganzfeld performance. European Journal of Parapsychology, *13*, 71-82.
- Dalton, K. (1997b). The relationship between creativity and anomalous cognition in the ganzfeld. Unpublished doctoral dissertation, Edinburgh University, Edinburgh, Scotland.
- Dalton, K., Morris, R. L., Delanoy, D. L., Radin, D., Taylor, R., & Wiseman, R. (1994). Security measures in an automated ganzfeld system. Proceedings of the 37th Annual Convention of the Parapsychological Association, Amsterdam, Holland, 114-123.
- Dalton, K., Steinkamp, F., & Sherwood, S. J. (in press). A dream GESP experiment using dynamic targets and consensus vote. Journal of the American Society for Psychical Research.
- Dalton, K., & Utts, J. (1995). Sex pairings, target type, and geomagnetism in the PRL automated ganzfeld series. Proceedings of the 38th Annual Convention of the Parapsychological Association, Durham, NC, 99-112.
- Delanoy, D. L. (1989). Characteristics of successful free-response targets: Experimental findings and observations. In L. A. Henkel, & R. E. Berger (Eds.), Research in Parapsychology 1988 (pp. 92-95). Metuchen, NJ: Scarecrow Press.
- Gelade, G., & Harvie, R. (1975). Confidence ratings in an ESP task using affective stimuli. Journal of the Society for Psychical Research, *48*, 209-219.
- Gertz, J. (1983). Hypnagogic fantasy, EEG, and psi performance in a single subject. Journal of the American Society for Psychical Research, *77*, 155-170.
- Gurney, E., Myers, F. W. H., & Podmore, F. (1886). Phantasms of the living. London: Trübner, 2 vols.

- Honorton, C., Berger, R., Varvoglis, M., Quant, M., Derr, P., Schechter, 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.
- Kanthamani, H., & Broughton, R. S. (1992). An experiment in ganzfeld and dreams: A further confirmation. Proceedings of the 35th Annual Convention of the Parapsychological Association, Las Vegas, NV, 59-73.
- Kanthamani, H., & Khilji, A. (1990). An experiment in ganzfeld and dreams: A confirmatory study. Proceedings of the 33th Annual Convention of the Parapsychological Association, Chevy Chase, MD, 126-137.
- Kanthamani, H., Khilji, A., & Rustomji-Kerns, R. (1989). An experiment in ganzfeld and dreams with a clairvoyance technique. In L. A. Henkel, & R. E. Berger (Eds.), Research in Parapsychology 1988 (pp. 84-88). Metuchen, NJ: Scarecrow Press.
- Krippner, S. (1975). Dreams and other altered conscious states. Journal of Communication, *25*, 173-182.
- Milton, J. (1997). An empirical comparison of the sensitivity of direct hits and sums of ranks as outcome measures for ganzfeld studies. Proceedings of the 40th Annual Convention of the Parapsychological Association in conjunction with The Society for Psychical Research, Brighton, UK, 226-239.
- Milton, J. & Wiseman, R. (1997). Guidelines for extrasensory perception research. Hatfield, England: University of Hertfordshire Press.
- Morris, R. L., Dalton, K., Delanoy, D. L., & Watt, C. (1995). Comparison of the sender / no sender condition in the ganzfeld. Proceedings of the 38th Annual Convention of the Parapsychological Association, Durham, NC, 244-259.
- Palmer, J. (1986). Statistical methods in ESP research. In H. T. Edge, R. L. Morris, J. Palmer, & J. H. Rush, Foundations of parapsychology: Exploring the boundaries of human capacity (pp. 138-160). London: Routledge & Kegan Paul.
- Persinger, M. A., & Krippner, S. (1989). Dream ESP experiments and geomagnetic activity. Journal of the American Society for Psychical Research, *83*, 101-116.
- Radin, D. (1997). The conscious universe: The scientific truth of psychic phenomena. New York: Harper Collins.
- Rosenthal, R., & Rosnow, R. L. (1991). Essentials of behavioral research: Methods and data analysis. (2nd ed.). New York: McGraw-Hill.
- Spottiswoode, S. J. P. (1997). Geomagnetic fluctuations and free-response anomalous cognition: A new understanding. Journal of Parapsychology, *61*, 3-12.
- Tolaas, J. (1986). Vigilance theory and psi. Part I: Ethological and phylogenetic aspects. Journal of the American Society for Psychical Research, *80*, 357-373.
- Ullman, M. (1986). Vigilance theory and psi. Part II: Physiological, psychological, and parapsychological aspects. Journal of the American Society for Psychical Research, *80*, 375-391.
- Ullman, M. (1989). A group approach to the anomalous dream. In M. Ullman, S. Krippner with A. Vaughan, Dream telepathy: Experiments in nocturnal ESP (2nd ed.), (pp.217-228). Jefferson, NC: McFarland
- Ullman, M., & Krippner, S., with Vaughan, A. (1989). Dream telepathy: Experiments in nocturnal ESP (2nd ed.). Jefferson, NC: McFarland.
- Ullman, M. (1990). Dreams, species-connectedness, and the paranormal. Journal of the American Society for Psychical Research, *84*, 105-125.

- Van de Castle, R. L. (1971). The study of GESP in a group setting by means of dreams. Journal of Parapsychology, *35*, 312.
- Van de Castle, R. L. (1977). Sleep and dreams. In B. Wolman (Ed.) Handbook of Parapsychology (pp. 473-499). Jefferson, NC: McFarland.
- Watt, C. (1989). Characteristics of successful free-response targets: Theoretical considerations. In L. A. Henkel, & R. E. Berger (Eds.), Research in Parapsychology 1988 (pp. 95-99). Metuchen, NJ: Scarecrow Press.
- White, R. A., Krippner, S., Ullman, M., & Honorton, C. (1971). Experimentally-induced telepathic dreams with EEG-REM monitoring: Some manifest content variables related to psi operation. In W. G. Roll, R. L. Morris, & J. D. Morris (Eds.), Proceedings of the Parapsychological Association, Number 5, 1968 (pp. 85-87). Durham, NC: Parapsychological Association.

POSITIVE THERAPEUTIC EFFECT OF DISTANT HEALING IN AN ADVANCED AIDS POPULATION.

F. Sicher, E. Targ, D. Moore, & H. Smith

California Pacific Medical Center.

ABSTRACT

Nonlocal or distant healing (DH), defined as an act of mentation to benefit another person's physical and/or emotional well-being at a distance, has been found in some form in nearly every culture since prehistoric time. Anecdotal experience has stimulated a substantial body of research. However, DH remains controversial with insufficient formal research involving **extended** periods of DH when subjects **do not know** if they are receiving treatment and with **extended followup** to establish whether such efforts actually affect health.

In this research it was hypothesized that an intensive 10-week DH intervention from experienced DH practitioners located around the U.S. would benefit medical outcomes for a population of advanced AIDS patients in the San Francisco area.

Two separate randomized, double-blind studies were performed: a pilot study of 20 subjects stratified by number of AIDS defining illnesses and a replication study of 40 subjects **pair-matched** for age, T-cell count and number of AIDS defining illnesses. Relevant independent variables at baseline and ongoing during the intervention were also examined.

Subjects were assessed by psychometric testing and blood draw at enrollment, after the 10-week intervention and at 6 month followup, when blind medical chart review by a physician was also conducted to verify patient report of illnesses, doctor visits and hospitalizations.

The main outcome measures were mortality, change in T-cell count, number and severity of new illnesses, number of doctor visits, hospitalizations, and change in subjective well-being (Profile of Mood States - level of distress; Wahler Physical Symptom Inventory - perceived physical symptoms; Medical Outcome Survey/HIV Version - perceived functioning).

In the pilot study 4 of 10 control group subjects died vs. none in the treatment group, but this result was confounded by age differences. When the replication study began nearly all subjects were on a new standard of care involving triple drug "cocktail" therapies which had shown strong correlation with survival and T-cell count, and neither of these measures differed significantly for the treatment and control groups. However, treated subjects acquired significantly fewer new AIDS defining illnesses (0.1 vs. 0.9 per pt., $p=0.04$), and the illnesses acquired were significantly less severe (severity score 0.8 vs. 2.15 per pt., $p=0.03$). Treated subjects required significantly fewer outpatient doctor visits (9.2 vs. 13.0 per pt., $p=0.01$), fewer hospitalizations (0.15 vs. 0.6 per pt., $p=0.04$) and fewer days of hospitalization (0.5 vs. 3.4 per pt., $p=0.04$). Treated subjects also showed significantly improved mood compared to controls (change in POMS -26 vs. +14, $p=0.02$).

These results support a positive therapeutic effect of DH in this study. Other studies of DH have produced mixed results, and the unique parameters or descriptors of this study are discussed. Discussion focusses on the need to replicate this type of research in order to be sure the effect is real and how such replications must take fully into account the subtle interactions occurring between participants which can affect results.

A GUIDE TO INDEPENDENT CODING IN META-ANALYSIS¹

Fiona Steinkamp

University of Edinburgh

ABSTRACT

Meta-analysis is a relatively new, but increasingly popular, technique for combining research findings across studies. The most time-consuming aspect of undertaking a meta-analysis is the coding. Coding involves assigning a number to each variable that the meta-analyst wishes to investigate in the studies in the database. Thus, if the meta-analyst wishes to examine whether female experimenters are better with male participants, the meta-analyst would code for both experimenter (e.g., 1 = male, 2 = female) and similarly for participants for each study in the database. The reliability of the findings from a meta-analysis depends to a large degree on the accuracy of the coding. One way in which to determine whether or not the initial coding is reliable is to have a second person independently code the database. If the two sets of coding agree, there can be a degree of confidence about analyses performed on that database. If the two sets of coding do not agree at all, any analyses performed using the initial coding will be relatively meaningless.

Nevertheless, there are many issues to be decided when employing independent coding of a database and these are not generally outlined in the literature. There are many decisions that have to be made prior even to getting someone to code the database. These decisions include whether or not to spend time on a pilot coding, whether the independent coder should be blind to the results, how much of the database to have coded, who should do the coding, how the papers should be prepared for the independent coder and whether one person should do the coding or whether the studies should be shared out amongst a number of people. Each decision has advantages and disadvantages that may affect the eventual interrater reliability.

A second set of decisions concerns the choices available to the meta-analyst once the second coding set has been obtained. These decisions include not only the decision as to which statistical test to use, but also issues about whether to use the meta-analyst's or the independent coder's coding decisions for the analyses and whether to correct for human errors in the coding.

A final concern is that of increasing the validity of meta-analytic results. If all variables include an "unsure" category, the validity of the final meta-analysis will be improved because most of the false positives and negatives will be channeled into the unsure category. Thus the categorical coding judgements (i.e., those that are not "unsure") will be cleaner as a result. Another conclusion is that if individual studies are particularly prone to strong disagreements between coders, those studies should be omitted from the database. Further recommendations arising from the paper include the use of blind coding, the examination of any bias in coding particular variables and more detailed reporting of the methods used in procuring independently coded datasets. If these recommendations are followed, meta-analytic results are more likely to be valid.

¹ I would like to thank Julie Milton and Caroline Watt for their comments on an earlier version of this paper. This research would not have been possible without financial support from Fundação Bial.

INTRODUCTION

There is an ever-increasing use of meta-analysis in numerous fields. Before the meta-analysis can be conducted, all studies in the meta-analysis need to be coded. Coding involves assessing the studies according to a prespecified set of criteria – for example, was the experimenter male or female, were participants encouraged, was the number of trials prespecified, were the statistical analyses prespecified etc.? The answers to these questions are represented numerically (e.g., 1 for male experimenter, 2 for female experimenter) and the meta-analytic database consists of the codes for all the variables for all the studies. Each variable coded in each study essentially represents a judgement on the part of the coder. Obviously, the coder may make a number of mistakes whilst undertaking the coding – this can be due not only to human error, but also to textual ambiguities or to variables requiring some subjectivity of judgement. It is therefore advisable for a database to be coded more than once in order to iron out the number of coding errors.

Independent coding can serve a number of purposes. It can help the meta-analyst to assess the reliability of the initial coding by showing how many of the original coding decisions other people would also make given the same information. Interrater reliability is standardly calculated to show how much agreement there is between the two coders (i.e., the meta-analyst and the independent coder(s)). If the independent coding is undertaken with the second judge blind as to the studies' outcomes, independent coding can also reveal any bias that may have entered into the initial coder's judgements through this coder having known the results of the individual studies. The initial coder cannot code the studies blind because they need to calculate the effect size etc. for each study. Independent coding can also reveal other potential coding biases such as a particular person's tendency to believe that (for example) optional stopping is unlikely when there is insufficient information to assess this likelihood with any certainty.

Nevertheless, there is relatively little information available about how to implement and to use independent coding (but e.g., see Rosenthal 1991, Stock, Okum, Haring, Kinney & Ceurvorst 1982, Wolf 1984). This paper aims to fill this gap in the literature.

Issues that are relevant to the use of independent coding with studies prepared for blind judging fall into two main categories; (i) how to go about getting the studies independently coded; and (ii) what to do with the independent coding after it has been completed. This paper will review and address problems pertinent to both of these categories. The first section will list the decisions that have to be made prior to conducting independent coding of a meta-analysis and the issues that have to be considered in making those decisions. The second and third sections outline further decisions to be taken once the independent coding has been completed and suggest ways in which the independent coding can be used to improve the integrity of the meta-analytic database. The paper concludes with a summary of recommendations arising from the discussion.

It will be assumed throughout that there is only one other coding of the studies involved, although clearly many of the issues will be relevant whether there is double coding, triple coding or more. I will also assume throughout that the meta-analyst is the person who does the initial coding.

INDEPENDENT CODING - GETTING STARTED

Wanous, Sullivan & Malinak (1989) list the number of judgement calls - and hence subjectivity - that come into play when designing and implementing any meta-analysis. These judgements include not only the coding decisions themselves (see Glass, McGaw & Smith, 1981 for a synopsis), but also judgements about how to retrieve studies and how to group studies if there are multiple measures etc. In this paper I will show that there are yet further decisions to be made once a meta-analyst chooses to implement an independent coding of their database. These decisions will form the italicised subheadings of this section. Each decision will be followed by a number of choices that the meta-analyst has and the relevant advantages and disadvantages to each alternative.

When to employ independent coding

Before undertaking the initial coding of the database, the meta-analyst can decide either to have a partial pilot independent coding implemented before the meta-analyst proceeds to code the whole database (as suggested by, e.g., Stock et al, 1982) or to request an independent coding of the database only once their own initial coding is complete. Some advantages and disadvantages to each approach are listed below.

1. Have a subset independently coded before performing initial coding of whole database

Advantages: Some potential misunderstandings of the meta-analyst's coding criteria that the independent coders may have can be ironed out beforehand; i.e., the meta-analyst can clarify the coding descriptions before coding the rest of the database and can thus improve subsequent interrater reliability. The meta-analyst may think more carefully about possible alternative interpretations of the texts whilst conducting the rest of the initial coding.

Disadvantages: The coding criteria may change over time anyway as the meta-analyst codes more studies and refines coding descriptions accordingly. Thus the pilot coding may be redundant if the final coding descriptions differ from the originals.

2. Leave independent coding until all initial coding has been completed.

Advantages: The meta-analyst will be aware of many of the individual problematic coding decisions in the studies in the database. If meta-analysts refine the coding criteria whilst rating the studies, the final coding criteria that the independent coder will use should be fairly thoroughly worked out.

Disadvantages: Any unanticipated misunderstandings of the coding criteria will not be discovered. Thus interrater reliability will be lower than if the criteria had been tested at an earlier stage of the initial coding process.

Although it would save time, it is inadvisable for independent coding to be undertaken simultaneously with the meta-analyst's initial coding (even for the initial subset for pilot coding), because the meta-analyst will usually refine the coding descriptions as the initial coding progresses. The independent coders could supply a pilot coding of a subset *after* the initial coding of the whole database is completed. The meta-analyst could then amend coding descriptions accordingly. However, if coding descriptions are revised in the light of the pilot

coding, the meta-analyst's original coding decisions may no longer match the revised coding criteria. To overcome this problem the meta-analyst could re-check all existing coding for revised variables. This is possibly the best way of increasing the final interrater reliability (which is subsequently performed on the whole database). However, it would also be the most laborious one. Some form of pilot coding is recommended nevertheless.

How many studies and variables to code

Before preparing studies for independent coding, the meta-analyst needs to decide how many studies and variables should be coded for the final interrater reliability assessment. The following delineates some of the relevant possibilities.

1. All variables and all studies

Advantages: Coding reliability will be unbiased, because the database used will not run the risk of including only variables and/or studies that are relatively easy to code (or the opposite).

Disadvantages: Time-consuming, especially if the database is very large. Some variables may not require any judgement (e.g., publication date) and thus may artificially raise the interrater reliability.

2. A sample of variables and/or a sample of studies

Advantages: Quicker than the above.

Disadvantages: Some problematic variables and/or studies may be omitted (as noted e.g., by Orwin & Cordray, 1985). As a result the subsequent reliability assessment will be an inaccurate estimate of the database reliability. The study or variable sample may not be large enough to detect more subtle, but recurring misunderstandings of the coding criteria (thus failing to show interpretative problems when performing analyses with those variables).

Coding all variables and all studies is clearly the option with the least disadvantages. However it may be unpracticable for very large databases, especially if there are time constraints and/or funding restrictions (the more time devoted to independent coding, the higher the personnel costs).

Whether to have the studies coded blind

Another decision is whether or not the studies in the database should be coded blind (i.e., with the independent coder blind to the studies' results). Meta-analysts will necessarily know the studies' results for they will need to refer to the results' sections for variables such as prespecifying the number of trials. Some advantages and disadvantages to blind coding are listed below.

1. Have the studies coded blind.

Advantages: The independent coder will not be biased through knowing the results of the studies. The independent coder's task will not be so laborious because they will not have to calculate effect sizes etc.

Disadvantages: There is no check on effect size calculations. Preparing studies for blind coding presents other problems for reliability (see next decision below on preparing material).

2. Do not have the studies coded blind.

Advantages: The independent coder can assess effect sizes etc. This serves as a double-check on the meta-analyst's calculations.

Disadvantages: The independent coder will know the studies' results and this knowledge may affect the way in which they decide how to code any given study. The independent coder needs to spend considerably more time on the task.

3. Have the studies coded blind and have effect sizes etc. recalculated only after all blind coding has taken place.

Advantages: As for blind coding, plus the advantage of having a check on the effect size calculations.

Disadvantages: Laborious and time consuming for the independent coder. Laborious and time consuming for the meta-analyst because the meta-analyst will have to prepare papers for blind review and keep the original studies for the independent coder to perform the subsequent calculations.

If biases are unknown or unbalanced – and it is arguable that at least the extent of anyone's bias is unknown – blind coding is obviously the more reliable option. In the rest of this paper I will assume that blind coding will be chosen. The issue of employing a later recalculation of effect sizes will depend on the difficulty of the meta-analytic database. If the effect size measure is easily applied to all studies in the database and if the effect size can be calculated automatically on a database by using the relevant formula, an independent recalculation of the effect size may be unnecessary. However, Hunter & Schmidt (1990) list the dangers of assuming that computerised calculations will be correct. Moreover, there is particular need for caution if some studies require complicated ways of dividing data into separate studies (and thus this division and subsequent calculations may be prone to error). Similar caution is required if effect sizes are difficult to obtain. In such databases it is certainly advisable to have another person calculate the effect sizes.

How to prepare the material for coding

The first task before independent coding can be conducted blind is to prepare material for the independent coder. It is preferable for the meta-analyst to prepare each paper as soon as the meta-analyst has coded it. If the preparation is left until a later date, the meta-analyst may

forget where relevant pieces of information reside in the article and where revealing information may unexpectedly be. It can be surprisingly difficult to remove all information that refers to the study's outcome from the paper. For example, a concluding paragraph may mention the desirability of replications of the current findings, thus suggesting that the study had been successful. If the paper is not read thoroughly before being prepared for blind coding, such references can be overlooked. In all cases journal name and experimenter name should be omitted in case coders are biased for or against certain journals and/or people in the field. Below I outline some of the issues to consider when preparing material.

1. Photocopy the whole article and blank out all individual references to the study's outcome.

Advantages: Much of the information is left intact and in context. The independent coder has the closest copy possible to the information that the meta-analyst used.

Disadvantages: Much of the information may be irrelevant, giving the independent coder more work than necessary. If information is blanked out in context, the context itself may suggest what that information had been. For example, if a study is significant, phrases such as "the results for this condition are in contrast to the results in the [other] section" may be used rather than a more neutral reportive style such as "this condition got [blank] and that condition got [blank]". The emphasis in discussion sections can also vary according to what the results are like (for example, alternative explanations for the data are given only if the study is significant) and this emphasis may reveal the outcome of the study.

2. Photocopy the article, but omit the results section. Include discussion sections only if they contain relevant information. If discussion sections are included, blank out all references to the results. Omit any irrelevant sections.

Advantages: The independent coder has less to read and can concentrate fully on the relevant sections.

Disadvantages: By omitting the results section, independent coders may misinterpret the design of the study they are rating. For example, they will not know how the data have been split in the study and may, therefore, subtly misinterpret the design of the study they are coding. The results section often has implications for some of the variables in the meta-analysis. For example, the number of trials for each condition is sometimes a good indicator of whether or not the number of trials has been prespecified. Without this information independent coders may not be able to make an informed judgement about the possibility of optional stopping. Even if results are blanked out of discussion sections, if there are large blocks of missing text, independent coders will be able to surmise that the results were significant.

3. Photocopy the article as in 2 above, but include sections that the meta-analyst feels are irrelevant. Omit discussion sections. Supplement the prepared article with a sheet outlining

how the data have been split, the number of trials etc. and giving any information in discussion sections that could not be included in the photocopy.

Advantages: Independent coders should have all relevant information. The additional sheet may mean less work for independent coders if it contains just relevant information.

Disadvantages: The meta-analyst may be unconsciously selecting some data and omitting others for the additional sheet. Alternatively, the meta-analyst may phrase the missing information in an unambiguous way on the additional sheet, whereas the actual text may have had an ambiguity that the meta-analyst had missed. Preparing the additional sheet involves a lot of extra work for the meta-analyst.

Obviously, all alternatives above are open to the meta-analyst failing to spot references to the results in the text and leaving them in by accident. The options above are by no means exhaustive and can be mixed and matched appropriately. If meta-analysts outlined in their reports how they prepared papers for blind coding, future analyses of meta-analyses may be able to determine which of the disadvantages make a larger difference to interrater reliability. Although many of the possible pitfalls outlined here may seem relatively small, they may nevertheless have an impact on the overall accuracy of the independent coding.

Who should code the studies?

Once the papers have been prepared for blind coding, a decision arises as to who or what type of person should do the coding. Some of the options are considered below.

1. People in the field

Advantages: It will be easier for them to read the studies if they have an interest in the subject area. They may be better able to assess the quality of a study if inferences have to be made from the text in order to arrive at a decision. They may be better able to make a sensible “best guess” in cases where the information is ambiguous.

Disadvantages: They may already know the results of a number of studies in the database - thus “blind” coding may not be blind at all.

2. People unconnected to the field.

Advantages: They will be unaware of the results of any particular study.

Disadvantages: They may find it uninteresting. They may not have the capacity always to make informed judgements.

If more than one person independently codes the studies and if the field is a controversial one (such as parapsychology or homeopathy which arouse widely differing attitudes to the data), independent coders should consist of people with similarly differing attitudes.

One or many coders?

I assumed at the beginning of this paper that there would be only one extra set of coding. When trying to find people who are competent to code the studies in the meta-analysis, the question arises as to whether it is better to have one person provide the independent coding or whether it is best for a number of people to share the studies between them. Although this latter option is relatively unusual, it nevertheless remains as a possibility. Naturally, there are advantages and disadvantages to each option:

1. One other coder

Advantages: The independent coder is able to gain a feel for the whole database. Coding becomes easier the more familiar the person becomes with coding descriptions and the database as a whole. The meta-analyst receives a consistent set of coding to compare with their own.

Disadvantages: Consistent errors or consistent biases on the part of one person will be repeated throughout the whole database. It demands a lot of the independent coder's time and they may find the coding tedious (thus they may start to make more mistakes).

2. Several coders

Advantages: The coders will be fresh because they won't be laden down with too many studies. Consistent biases/errors will hold for only part of the database. Because many people are working on the coding at once, the coding will be completed sooner. A larger range of coding decisions will be represented.

Disadvantages: It is difficult to obtain a number of qualified people with the time to undertake the coding. In the initial stages the coders will be slow until they become accustomed to the coding descriptions and the database as a whole (i.e., total working hours will probably be longer).

Obviously, the studies will be shared amongst a number of people only if a relatively large number of studies and variables is being coded.

INDEPENDENT CODING – HOW TO PROCEED

Deciding on a statistical measure

Once the second coding set has been procured, there are a number of ways in which the reliability of the database can be assessed. It is not the aim of this paper to discuss statistical issues; however it is worth briefly mentioning the main alternatives here. Percentage agreement is a popular option. An acceptable percentage agreement rate is 80% (Hunt, 1997).

However, Rosenthal (1987) recommends using a correlation coefficient instead (usually

Pearson's r). If the coding criteria are ordinal, it is possible to take the mean of the judges' ratings for each variable for the final analyses. If this option is used, the effective reliability of the database (or of the relevant individual variables if only part of the database has ordinal coding) will be greater because more than one judge's ratings is used to determine the final coding decision for each variable. The resultant effective reliability (R) can be calculated from the correlation coefficient (Rosenthal, 1991, 1987).

Nevertheless, use of a correlation coefficient and the estimate of effective reliability depend on the coding criteria being ordinal. In many databases, however, the criteria may be categorical. For categorical variables Cohen's (1960, 1968) kappa is recommended. The aim of the following subsections is to outline conceptual problems that arise when the coding criteria are categorical.

Resolving differences of opinion

For the most part meta-analysts with categorical data have just used the independent coding as a way of assessing the reliability of their own coding rather than using the independent coding as a way of providing a cleaner data set. This seems like asking another person(s) to do a lot of work for not much overall gain. There are several options for resolving differences of opinion that I shall discuss below.

1. Through discussion. The meta-analyst and the independent coder discuss all instances of disagreement together and come to an agreement for each decision.

Advantages: The agreements are more objective and presumably the coding as a whole will be cleaner. Two heads are better than one and either side may have seen things in the database that the other person had clearly either missed or not considered.

Disadvantages: Discussion may not take place as if between equals. The meta-analyst may well press their opinion through more forcibly because it is their research project. The independent coder has to devote considerably more time to the project because of the additional discussion time needed (i.e., personnel costs will be higher).

2. The meta-analyst checks the database for human error on both their part and the independent coder's part. Human error on both sides is corrected by the meta-analyst. All cases in which the coder and the meta-analyst disagree but in which the meta-analyst can see a reason for the independent coder's decision are left as disagreements. The corrected database is then assessed for overall agreement/reliability.

Advantages: Human error is corrected without inconveniencing the independent coder. The database is cleaner and an assessment of its reliability is still meaningful.

Disadvantages: There is sometimes a thin line between a disagreement due to ambiguity and a disagreement that is actually due to error. The meta-analyst may simply not have noticed an ambiguity that the independent coder had seen and the meta-analyst may therefore believe that something was an error on the independent coder's part when it was not. The meta-analyst may not understand when they

have been in error (i.e., even some of the meta-analyst's coding mistakes may still be present in the database).

3. As in (2), but the meta-analyst corrects only human error on their part.

Advantages: As in 2.

Disadvantages: The independent coder's human errors are still present (as may be some of remaining meta-analyst's errors) and mistakenly conflicting with the meta-analyst's final coding.

4. Leave the database as it is.

Advantages: Assessment of percentage agreement more accurately reflects the actual database (assuming percentage agreement is calculated prior to amendments). No more time is needed on the coding.

Disadvantages: All errors remain in the database and the independent coding has not served as a way of pointing up errors in the meta-analyst's coding. As a result independent coding does not provide a cleaner data set.

Of course, there are subsets of these four basic strategies (for example, the independent coder and the meta-analyst could independently assess coding discrepancies for cases of human error only, make corrections as they feel appropriate and compare and discuss discrepancies afterwards). These subsets will have the appropriate advantages and disadvantages listed above. No one strategy is ideal, although the least optimal strategy is that of leaving the database as it is.

Deciding on data set

In most instances (2-4) the question remains of whether to conduct the meta-analysis on the meta-analyst's data set (i.e., coding decisions) or on that of the independent coder. Although independent coding is usually implicitly understood as a way of assessing the reliability of the meta-analyst's coding, it is far from clear that the meta-analyst's coding is necessarily the best data set to use. I will list and discuss the most obvious alternatives below:

1. Use the blind coding

Advantages: This data set should not be biased by knowledge of the studies' results and/or from any hopes or expectations regarding the final analyses.

Disadvantages: The independent coder may have consistently misunderstood the meta-analyst's reasons for certain coding criteria if the meta-analyst had failed to make them explicit (for example, if the coder did not know why the meta-analyst was interested in whether or not the number of trials were prespecified and the underlying interest was optional stopping). The independent coder may have had less interest in the study and thus coded more carelessly as a result.

2. Use meta-analyst's corrected coding

Advantages: The coding should be an improvement on the original, uncorrected data set. It will have the relevant advantages to the resolution of conflicting data strategy used. The data set will be coded as the meta-analyst intended it to be.

Disadvantages: The coding may retain some of the meta-analyst's biases. The disadvantages to the amendment of data strategy used will be in play.

3. Use only those coding decisions where both blind coder and meta-analyst agree for conducting the analyses on the variables in question. Retain the meta-analyst's coding for study quality/flaws if studies are to be weighted for quality.

Advantages: A clean data set - the meta-analyses undertaken with each moderator variable will have greater validity.

Disadvantages: On a small database this may not be practicable (there may not be enough studies to make an analysis worthwhile). Power of analyses is reduced. Blind coder and meta-analyst may agree and yet the decision may still be wrong.

I can see no advantage to using a single uncorrected data set. My own conclusion from the arguments above is that method 1 of correcting the database may be best. If any of the other methods are used, the meta-analyst's corrected coding appears to be the best data set or, where practicable, a database consisting of only agreements between both data sets.

Given that reliability of the data is so important - for if there is little agreement, there can be no confidence in the results from the meta-analysis - improving reliability and learning from the coding is crucial. In the second part of this paper I suggest ways to improve the validity of meta-analytic results by increasing database reliability further.

INDEPENDENT CODING – IMPROVING VALIDITY OF META-ANALYTIC RESULTS

This section of the paper will use for illustration the independent coding of a meta-analysis of forced-choice experiments examining both clairvoyance and precognition (Steinkamp & Milton, 1998). In this meta-analysis the independent coding of the whole database was shared out amongst five people after a small subset had been pilot coded by two test coders. Some of the issues raised will be applicable only to the methods and variables used for this particular meta-analysis. However, the issues should be generalizable to similarly conducted meta-analyses. There will be three subsections – one discussing the moderator variables, one considering the coding of safeguard variables across studies and one dealing with the coding of safeguard variables across variables.

Coding - Moderator variables

In this meta-analysis some judges had articles all of a similar type to code. For example, all the defence mechanism test papers were coded by one person and all papers using ESPerciser were coded by one judge. This was because some papers refer back to other papers for full details of the experimental procedure and thus such papers have to be considered in conjunction with each other. In cases such as these, then, any agreements or disagreements over variables will be exaggerated because the same decision will be repeated over a number

of papers. This problem could be overcome by having the initial paper prepared for blind review a number of times for reference back by a number of coders taking separate, but related studies. Care would nevertheless have to be taken to ensure that individual papers did not each provide separate snippets of information about the procedure; ensuring that all relevant details were given in the reference paper for all coders to use would be somewhat time-consuming. Moreover, the problems in general related to preparing papers for blind review (listed above) would be relevant here.

The variable of whether or not participants had been encouraged had relatively low agreement (61%). On referring back to the original papers, it became clear that various coders had interpreted encouragement differently. Therefore, a better definition of encouragement would be necessary for future meta-analyses wishing to examine this variable. For instance, both competitive situations and rewards could be defined as encouragement of participants, although the meta-analyst had not regarded them as such. The implicit assumption of the author had been that what was of interest was whether participants were encouraged to believe that psi was possible. However, this is itself ambiguous. Participants may believe that psychic functioning (psi) is possible, but not that they themselves could produce good results. Alternatively, they may think that psi is possible, but not in the laboratory situation.

This variable is a good example of how hard it is to (i) pin down what is being examined; (ii) to specify what the underlying purpose of the moderator variable is; (iii) to anticipate not only what procedures may come up in the studies to be coded but also to anticipate all the ways in which other people could possibly interpret those procedures; and (iv) to anticipate what information in a given study another person may find relevant to that specific variable.

Moreover, even if this variable were divided into several variables to take into account the full range of possibilities (for instance, differentiating between encouraging an individual to accept their own psi experiences, encouraging participants to believe that experiments investigating psi can be successful, encouraging individuals to think that they can be successful in the experiment and adopting incentives to try to increase success), the result may be that only a few studies present enough information to make a judgement possible (thus yielding too few studies for analysis on any of the variables). The specificity of the information required may also mean that broad information (e.g., “the atmosphere was one of encouragement”) may be lost amongst all the “unsure” categories.

However, the advantage of looking at an independent rater’s coding of the studies lies in discovering problems with existing coding and, therefore, ways in which that coding could be improved for the present or for future meta-analyses. Although this is time-consuming, it is vital to have as clean and as clear a data set as possible.

Coding - Safeguard variables.

The coding for the safeguard ratings generally fell into three categories - no, the safeguard was not present; unsure whether it was present or not; and yes, it was definitely present. It was hoped that by using a three-point scheme a more representative rating of the presence or otherwise of safeguards could be established. The unsure category would mean that there would be fewer false positives in either the yes or no category. The following analyses on the safeguard variables for the database will show ways in which databases using an unsure category can have a cleaner data set for the final analyses and in which a clearer understanding

of the implications of the independent coding can be gained. I shall examine this both across studies and across variables.

Safeguard variables - across studies: One important question is whether coding disagreements could significantly affect the overall quality rating of any given study. If a weighted Stouffer's z is calculated by weighting the results according to the number of study safeguards in place and if individual coders radically disagree about the overall quality rating of the individual studies, the Stouffer's z will be relatively meaningless. It will likewise be impossible to obtain any reliable assessment of the effect of moderator variables on effect size if the quality ratings are themselves unreliable.

By looking at the differences between the two coders' ratings for each study, some implications for the reliability of the meta-analytic conclusions can be ascertained. There are two main areas in which such an assessment can be informative.

Firstly, it can show whether biases are in play. Take the current database of 31 clairvoyance studies and 31 precognition studies. Each clairvoyance study could gain a maximum of 16 points for quality. Thus the maximum quality points that could be awarded to the entire set of clairvoyance studies was $31 \times 16 = 496$. Each precognition study could gain a maximum of 12 quality points. Thus the maximum quality points that could be awarded to the entire set of precognition studies was $12 \times 12 = 372$. By summing the number of safeguard points given overall by the independent coders and the meta-analyst, it can be seen whether there was a general difference in bias towards leniency or severity between the two sets of coding. In the current database the overall difference in point assignment was 11 (with the blind judges being less lenient) out of a possible total of 868 points. It appears from this database that there is little difference in coding bias towards overall leniency or severity between the meta-analyst and the independent coders.

A second and perhaps more important consideration is whether the total number of quality points assigned to individual studies differs. If there is a large discrepancy between the two coders' quality ratings of the same study, it is difficult to understand that study's viability. As a result, any meta-analytic conclusions drawn from variables present in that study will be open to doubt.

In the meta-analysis under consideration, an individual clairvoyance study could score at most 16 points for quality. However, the disagreement between the independent coder's and the meta-analyst's total quality rating for any individual study ranged from -6 points to +10 points. That is, one study in the database may have been rated by the meta-analyst as having a total of 12 quality points, whereas the independent coder would assign that same study only 6 quality points (i.e., -6). Alternatively, the meta-analyst may judge another study to be worth only 2 points, whereas the independent coder would rate it as worth 12 points (i.e., +10). Therefore at the extreme there was almost a 66% difference between blind coder and meta-analyst for safeguard ratings of individual studies. Moreover, a similar range of differences held for the precognition studies too. Here the range was from -8 to +7. The total possible quality points for an individual study in the precognition condition was 12. Again, the extreme of the range indicates a 66% difference between the blind coder's quality rating of a given study when compared to the meta-analyst's rating. This range is worrying.

Three of the four clairvoyance experiments with a large disagreement (a difference of more than +5 or -5 [31%]) had been coded three times. These particular experiments had been commented on spontaneously as especially difficult to code by the test coders. This suggests that large disagreements are due to the individual texts (such as inadequate descriptions of procedures used) rather than to coding biases.

It may therefore be advisable to discard experiments where there is a large disagreement. In any event it would be worthwhile to prespecify what to do with individual experiments in which there is little agreement over quality between coders.

These analyses show how examining the two sets of coding across studies can be useful in objectively determining problematic experiments in the database. If these problematic experiments can be deleted from the database, the overall reliability of the database - and hence of the meta-analysis itself - can be improved.

Safeguard ratings - across variables: Another way of analysing disagreements is to see whether any one particular quality variable is singularly difficult to code or is particularly prone to bias. Analyses of disagreements across variables, therefore, may be useful. Such analyses would show which variables should be considered with caution in the final meta-analysis and which variables can be considered with some confidence.

The clairvoyance data are given in Table 3. Each safeguard variable could have gained a total of 62 points if that safeguard had been present in every single study. It can be seen that the range of difference between the total number of points assigned to individual variables (total point difference) by the meta-analyst and the independent coder is less across variables than across studies; the range of total difference between coders across variables being from -6 to +3. That is, in some cases the meta-analyst may have awarded a given variable 30 points in total, whereas the secondary coder would have assigned it only 24 (-6). For another variable the secondary coder may have given a total of 26 points, whereas the meta-analyst assigned it only a total of 23 points (+3). Thus here the disagreement is at most just 10%. Apparently, most disagreement in the current database is due to individual problematic studies rather than due to the variables as such. This is obviously a useful and informative conclusion.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Overall
Total difference	-2	1	2	3	-6	-5	1	3	-3
Total N's with disagreement	3	5	9	8	8	11	15	10	69

N = number of experiments

Table 1: Disagreements between meta-analyst and coders across variables

However, it is important to see the number of studies in which the meta-analyst and the independent coders disagree about their coding of a particular variable (i.e., when total N's with disagreement is large). For example, a given variable may be difficult to code (there may be a lot of studies in which the two coders disagree), but if coders disagree half the time in both directions, there would not be a great total difference in points assigned. It therefore seems that these variables with many disagreements could not be used for showing whether or not they were responsible for an increase in effect size, because the coders disagreed so much about whether or not the variable was even present. This appears to be the case for Q7 above

where there are 15 instances of disagreement (out of a possible 31), but only one point total difference in safeguard points assigned for the variable.

However, if the disagreements in Q7 (and perhaps Q2 and Q3, which also have relatively many disagreements but only a little total point difference) are all between “don’t know” (1 point) and either “yes” or “no” (2 and 0 points respectively), the disagreements are not serious ones, but just ones of degree of caution (“cautious disagreements”). If the ratio between point difference and total disagreements is large, however, one would expect more “radical” (i.e., disagreements between “yes” [2 points] and “no” [0 points]) than cautious disagreements.

	Q1	Q2	Q3	Q5	Q7
No. of radical disagreements	1	0	1	4	2
No. of cautious disagreements	2	5	8	4	13

Table 2: Radical versus Cautious disagreements

Table 2 provides the relevant data for the clairvoyance condition and shows some support for this hypothesis. Q7, for example, has only 2 radical disagreements and a total of 13 cautious disagreements, whereas Q5 has 4 radical disagreements and only 4 cautious disagreements.

These results suggest several recommendations for future meta-analyses. Firstly, it points to the advantage of using a three-point scheme (or at least a scheme in which “unsure” can be used as a coding option) over a binary one (or one in which there is not an “unsure” coding option). In 3-point schemes disagreements can be ones of caution rather than ones of any serious consequence to the overall meta-analytic results from that variable. For although disagreements in a binary scheme may have a cancelling out effect (ambiguous cases equally being coded as mistakenly “yes” and mistakenly “no” by the meta-analyst or the blind coder), in the three-point scheme many of these disagreements will be channelled into the middle category thereby leaving a cleaner set of “yes” and “no” results for any analyses on those variables.

Another advantage of the three-point scheme is that it is easier to pin down which variables are the really problematic ones. Where there is no “unsure” category it is difficult to know which disagreements are due to ambiguity and which are clearly due to a radical disagreement, because all disagreements will be expressed as a simple “yes/no” opposition. In a three-point scheme radical disagreement will be expressed in the yes/no opposition and other disagreements will be between “yes/no” and “don’t know”. If there were a high proportion of radical disagreements for a particular variable, it is likely that the variable will provide less valid results. This would serve well as a warning to the meta-analyst to treat any conclusions from that variable with some caution.

Additionally, it suggests a quick and practical way of increasing reliability and of increasing the validity of the final meta-analyses. Namely, all instances of radical disagreement could be changed to “don’t know”. Not only would this improve reliability and produce a cleaner data set of “yes” and “no” for the meta-analysis; it would also prevent the problems outlined at the end of the first half of this paper in using discussions to settle disagreements. For a three-point scheme, this procedure would be the same as using the mean rating (Rosenthal, 1991), but for categorical variables, substituting all disagreements that were between two distinct

categories for “unsure” would provide a far cleaner set of data. However, this would require a relatively large database.

CONCLUSION

The aim of this paper was twofold. The first part outlined conceptual problems that arise prior to implementing an independent coding of the meta-analytic database. The second part emphasised ways in which independent coding can be used to improve the quality and reliability of the final meta-analysis. Given that coding is the largest and most important part of any meta-analysis, the inconvenience of analysing the independent coding as well is more than outweighed by the increased likelihood of procuring valid results from the meta-analysis. Moreover, the independent coder invests a lot of time and effort in undertaking the coding task and it is a waste of their effort if meta-analysts do not learn from the independent coding. If meta-analysts published more details about how the independent coding was undertaken, future analyses could show which follow-up studies based on meta-analyses fare better under which procedures of undertaking or using independent coding. One would hope that when the procedures recommended here are followed, resultant follow-up studies would have a greater chance of success.

In sum, the author recommends the following:

- When studies refer back to each other and several people are sharing the coding task, share the studies out rather than giving them all to one person.
- Prepare papers for blind coding as you go along.
- Conduct a meta-analysis on a subset first if you are planning a meta-analysis of a large database.
- Provide the blind coder with information about how the study data are divided; have the calculations double-checked.
- Include an “unsure” category for all variables
- Exclude individual studies that produce a large prespecified inter-coder disagreement (e.g., more than 30% of the maximum quality rating that a study can possibly have).
- Follow the above but for each variable.
- When reporting the meta-analysis, describe how many coders were used; how many variables were independently coded; how papers were prepared for blind coding if coding was performed blind; which database was used for the final meta-analysis and which changes (if any) were made to the database in order to improve the integrity of the database.

REFERENCES

- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**, 37-46.
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, **70**, 213-220.
- Glass, G.V., McGaw, B. & Smith, M.L. (1981). *Meta-analysis in Social Research*. London: Sage Publications.

- Hunt, M. (1997). *How Science takes Stock. The Story of Meta-Analysis*. New York: Russell Sage Foundation.
- Hunter, J.E. & Schmidt, F.L. (1990). *Methods of Meta-Analysis. Correcting Error and Bias in Research Findings*. London: Sage Publications.
- Orwin, R.G. & Cordray, D.S. (1985). Effects of deficient reporting on meta-analysis: A conceptual framework and reanalysis. *Psychological Bulletin*, **97**, 134-147.
- Rosenthal, R. (1987). *Judgment Studies. Design, Analysis and Meta-Analysis*. London: Cambridge University Press.
- Rosenthal, R. (1991). *Meta-analytic Procedures for Social Research*. Beverly Hills, CA: Sage.
- Steinkamp, F. & Milton, J. (1998). A meta-analysis of forced-choice experiments comparing clairvoyance and precognition. Paper presented at the Annual Convention of the Parapsychological Association, Halifax, Nova Scotia, August 1998.
- Stock, W.M., Okum, M., Haring, M., Miller, W., Kinney, C. & Ceurvost, R. (1982). Rigor in data synthesis: A case study of reliability in meta-analysis. *Educational Researcher*, **11**, 10-20.
- Wanous, J.P., Sullivan, S.E. & Malinak, J. (1989). The role of judgment calls in meta-analysis. *Journal of Applied Psychology*, **74**, 259-264.
- Wolf, F.M. (1984). *Meta-analysis. Quantitative Methods for Research Synthesis*. London: Sage Publications.

A META-ANALYSIS OF FORCED-CHOICE EXPERIMENTS COMPARING CLAIRVOYANCE AND PRECOGNITION¹

Fiona Steinkamp & Julie Milton

University of Edinburgh

ABSTRACT

This meta-analysis examined a database of studies published in the main parapsychology journals from 1935-1997 that compared outcomes of precognition and clairvoyance trials under relatively similar experimental conditions. Both the precognition and clairvoyance studies had a statistically significant cumulated overall effect, with the precognition database being more statistically significant (Stouffer $z = 4.15$, $p = 2 \times 10^{-5}$) than the clairvoyance database (Stouffer $z = 2.68$, $p = .004$) because it contained the larger number of trials. However, the mean effect sizes ($z/N^{1/2}$) of the two types of ESP arising from the 22 study pairs were very similar at 0.010 for precognition and 0.009 for clairvoyance. Even when the studies were weighted according to the number of trials (a larger number of trials being expected to yield a more reliable estimate of effect size), the effect sizes were still very similar, with a mean of 0.034 for precognition and 0.030 for clairvoyance. Thus there was no evidence from this meta-analysis to suggest that clairvoyance works better than precognition.

Studies in the database were coded for a number of moderator and safeguard procedure variables, with coding reliability between the first author's coding and that of blind coders being on average 78%, an acceptable level. There were no statistically significant correlations between the presence of procedural safeguards and effect size and hence no suggestion that methodological problems had played any strong and obvious role in the overall effects; nor did any of the moderator variables examined show any significant relationships to effect size. However, the small number of studies in the database would be expected to provide relatively low statistical power for detecting relationships between effect size and either type of variable.

Contrast between groups in effect size alone, regardless of the statistical significance of the difference, was used to identify promising variables to examine or exploit in future research. This yielded the following optimal procedures for precognition studies: trial by trial feedback, target selection within seconds of the participant's response and individual (rather than en bloc) presentation of target material. For clairvoyance the optimal procedures appeared to be: indicating the target by motor response rather than verbal or written response, trial-by-trial feedback and individual presentation of target material. Two of the variables were shared by both types of ESP task; response mode was the only variable that might possibly separate out precognition and clairvoyance. However, in the absence of any significant differences between the relevant contrast groups, further research will be necessary before these procedures can be taken seriously as effect-size optimisers. Suggestions for future research are discussed.

¹ The first author was responsible for designing and carrying out the meta-analysis, with the second author acting as a consultant. Bob Morris contributed many hours discussing coding decisions and coding descriptions as one of the initial coders. We greatly appreciate the valuable time spent on this task. The quality of this meta-analysis has been much improved by the efforts of the independent coders - Caroline Watt, Dagmar van der Neut, Kathy Dalton and Paul Stevens. To them we are also indebted. We are also grateful for the input from two anonymous referees. Finally, this research would not have been possible without the financial support of the first author by the Fundação Bial.

INTRODUCTION

Parapsychologists have long been interested in the question of whether clairvoyance and precognition are really different phenomena. It has been suggested that apparent precognition might actually consist of a person sampling the present environment via real-time ESP and extrapolating from the information to make an informed prediction about future events (see, e.g., Mundle, 1978). If this hypothesis is correct, then clairvoyance studies would be expected to result in higher effect sizes than precognition studies because of the extra calculational step involved in the latter type of ESP task. Conversely, it has been proposed that what appears to be a clairvoyance task may not involve real-time information acquisition but rather consists of precognition of the feedback about the target's identity that either the experimenter or the participant later receives (e.g., Carington, 1945). In this case, the effect sizes of the two types of ESP task would be expected to be the same and would be expected to be influenced by the same moderator variables, all other things being equal. However, only two papers to date (Tart, 1983; Milton, 1998) have reported comparisons between groups of such studies.

Tart (1983) presented an unusual meta-analysis of 85 forced-choice ESP studies. He discarded all studies that were not statistically significant at an alpha of .05 and also excluded all but the most successful subset of data from any study that broke down the overall result into individual participant data or individual condition or run data. He found that real-time ESP (clairvoyance or telepathy) studies outperformed the precognitive studies ($p < 5 \times 10^{-4}$). However, the reverse held in Milton's (1998) meta-analysis of 78 free-response ESP studies where the mean effect size ($z/\sqrt{N}^{1/2}$) for the 6 precognition studies (0.34) was higher than that of both the telepathy studies (0.18) and the clairvoyance studies (0.08).

Although the results from Tart's (1983) and Milton's (1998) papers appear somewhat contradictory, both results have problems of interpretation. Milton's (1998) one-way ANOVA to test the effect size differences between precognition, clairvoyance and telepathy was non-significant. The number of precognitive studies is too small to be representative of the latent population of such studies or to give an accurate estimate of their true effect size, which may be much lower than the observed figure. The confidence intervals around the estimated effect sizes of the three types of ESP overlapped considerably, making it easily possible that the true effect size of the population of precognition studies is below that of the real-time ESP studies rather than above as the point estimates suggest. Tart's (1983) result is also inconclusive, but for different reasons. In this paper the study selection criterion may have introduced a bias. If real-time and precognitive studies had the same effect size but precognitive studies tended to have a larger number of trials than real-time ones, the higher level of statistical power in the precognition studies would have enabled more precognition studies with lower effect sizes to reach the .05 alpha cutoff. This in turn would have meant a decrease in the average effect size for the precognition studies. Tart did not report study sizes and this artefact may be a strong possibility, given that precognition trials can be easier to run than clairvoyance trials and hence easier to collect in large numbers. Moreover, if more real-time than precognition studies reported breakdowns of data into subsets, Tart's policy of including only the "peak performance" data from each study would have similarly inflated real-time effect sizes relative to precognitive effect sizes.

The current meta-analysis aimed to address more directly the question of whether there is any difference between clairvoyance and precognition by comparing clairvoyance and precognition effect sizes in pairs of forced-choice studies in which both types of ESP were examined. The

study pairs were expected to use similar procedures when testing the two types of ESP, with the same experimenters and perhaps even the same participants, making any difference in effect size between the two types of ESP task easier to detect. A secondary aim of the meta-analysis was to see whether or not the optimal conditions for the two types of ESP appeared to differ.

METHOD

Selection of Studies

It was decided in advance to include only studies published between 1935 and 1997 inclusive in: *European Journal of Parapsychology* (and its forerunner, the University of Utrecht Parapsychology Laboratory's *Research Letter*); *Journal of the American Society for Psychical Research*; *Journal of Parapsychology*; *Journal of the Society for Psychical Research* and *Research in Parapsychology*. Although not decided upon in advance, *Proceedings of the Annual Convention of the Parapsychological Association* were consulted for any additional details for papers reported in *Research in Parapsychology*. The first author scanned all papers that looked as though they might contain relevant studies. This is generally considered to be a more effective method of study retrieval than relying upon search engines or inspecting titles or abstracts only (Glass, McGaw & Smith, 1981).

The criteria for determining which studies were eligible for the meta-analysis were designed to include studies that tested both precognition and clairvoyance and to exclude studies where it was unclear that the appropriate comparison could be made.

Studies were included in the meta-analysis if they met the following criteria:

- Both clairvoyance and precognition were tested in trials reported in the same paper, with the authors clearly making a conceptual distinction between targets selected at the time of the participant's guess and those selected after the participant's guess.
- Papers reported only clairvoyance or precognition trials but stated that the procedure was chosen either to be similar to the procedure used in a previous paper studying the other type of ESP or for the purposes of direct comparison with the other type of ESP examined in the previous paper.

A number of exclusion criteria were also prespecified, but additional criteria (listed below in square parenthesis) were added whenever a new study presented an unanticipated challenge.

- Participants aimed for psi-missing under both clairvoyance and precognition conditions.
- [Outcomes were merely reported as an aside with no details about procedure.]
- [Outcomes required correction for a stacking effect and did not provide data with which to make the correction.]
- [Possible PK effects on ESP success were examined (e.g., checker effects)]
- Either clairvoyance or precognition trials took place only as the occasional run not intended to serve as comparison trials; or if they apparently had been intended for comparison but comprised fewer than 100 trials.
- A paper reported a number of ESP procedures tested in a variety of ways without

indicating any particular clairvoyance trials as comparable to any specific precognition trials and the number of trials for either ESP type exceeded the other by more than 80% (a larger ratio being assumed to indicate that the study was not conceptually comparing precognition against clairvoyance sufficiently well to be included).

- Clairvoyance and precognition data were not reported separately and thus an effect size for each could not be calculated.
- [Some of the data reported in a study had already been included in a previous study.]

Study Definition

The criteria for defining what constituted a study were designed to group precognition and clairvoyance trials into coherent study pairs. By prespecifying how to group the data into studies, it was hoped that any potential bias that might arise in choosing how to divide the data in any given paper would be avoided. The following criteria for considering a set of trials to be part of a study pair were used:

- Clairvoyance and precognition conditions reported in the same paper were considered a study pair, as were precognition and clairvoyance studies reported in separate published papers as long as it was clear that their authors had intended them to be compared.
- When either precognition or clairvoyance trials were presented as separate experiments in a single paper only because the trials were carried out at different dates without any change in procedure, they were treated as a single study.
- When reports included a number of studies from a number of experimenters, experimenters were held constant in both conditions whenever possible, thus giving separate studies per experimenter. Thus where there were reports of studies from a number of investigators in a single paper, this paper could represent a number of studies for the meta-analysis.
- If a study manipulated moderator variables of interest in the present meta-analysis under both clairvoyance and precognition conditions, each moderator category constituted a separate study pair for the meta-analysis. For example, if participants made verbal responses in some trials and written responses in others, that would constitute a study pair with verbal calls and a study pair with written calls.

Coding

The coding criteria are listed in Appendix 1. Once the first author had coded all the studies, the coding criteria were given to the second author, who independently coded the studies in five papers selected from the database. The authors then discussed ambiguities in the coding criteria, textual ambiguities in the papers and any instances of human error. Coding criteria and the codes assigned to individual studies were changed where appropriate, all such decisions being made by the first author. The first author also reviewed the codes assigned to all other studies in the database, making changes where appropriate to ensure that they had been coded consistently with any amended coding criteria.

Clairvoyance (N = 22)	Precognition (N = 22)
Safeguard (%)	Safeguard (%)

	yes	unk.	no	<i>r</i>	yes	unk.	no	<i>r</i>
Participant access to target prevented during trial	77	18	5	-0.13	-	-	-	-
Experimenter blind	45	50	5	0.04	-	-	-	-
Participant access to re-cords prevented after trial	73	23	5	-0.19	-	-	-	-
Number of trials in study prespecified	59	27	14	-0.03	59	27	14	0.08
Adequate randomisation	64	5	32	-0.20	73	9	18	-0.29
Checking of target records	36	14	50	0.23	45	14	41	0.41
Checking of calls	27	59	14	0.02	32	55	14	0.13
Checking of hits	41	18	41	0.07	45	23	32	0.16
Random sequence entry point selection made blind to calls	-	-	-	-	45	23	32	-0.01
Average				-0.02				0.08

Table 2: Percentage of studies with procedural safeguards present, unknown and absent, and correlations between presence of procedural safeguards and effect size

Variable	Study pair	Clairvoyance	Precognition
Study pair procedural similarity	58	-	-
Within vs between subjects design	94	-	-
Participant knowledge of trial type	65	-	-
Experimenter knowledge of trial type	84	-	-
Interspersion of precog. and clair. trials	68	-	-
Participant population type	-	81	74
Response modality	-	84	84
Expectation of success mindset	-	61	84
Target type	-	94	97
Target presentation mode	-	81	81
Timing of experimenter's feedback	-	81	81
Participant feedback of each target	-	55	81
Participant told score	-	97	100
Timing of participants' feedback	-	90	87
Individual vs group testing	-	94	90
Participant contact with experimenter	-	100	100
Predicted direction of results	-	61	61
Target-participant distance	-	77	-
Precognitive timespan	-	-	74
Complexity of target selection method	-	-	81
Average	-	81	84

Table 3: Percentage agreement between first author's and blind coding of studies for each moderator variable

Next, the first author gave Robert Morris (RLM) the revised coding criteria and five other papers from the database. The papers had been photocopied with the results sections blanked out and any reference to the study's outcome excised as far as possible so that RLM would be blind as to the studies' results as well as to the first author's coding of them. The first author

and RLM then compared their codings, discussing any criterion definition problems, textual ambiguities and human error until they had mutually agreed how each variable in each study should be coded. The first author then checked the coding of all other papers against any amended coding criteria resulting from her discussions with RLM.

All papers containing studies in the database were then prepared for blind coding in the same manner as before and shared out amongst five independent coders with a background in parapsychology, including the second author¹, who all used the final coding criteria (see Appendix 1). All studies in the database were therefore coded at least twice - once non-blind by the first author and once blind by another coder. About a third of the studies had received an additional coding by either the second author or RLM.

Coding reliability. After the five independent coders had coded the studies in the database, the first author thoroughly checked all disagreements between her coding and that of the five independent coders. As a result, five of her previous coding decisions were altered on the basis that they represented human error on her part. Analyses of the reliability of the coding of the whole database were then performed. For the moderator variables it is only possible to report percentage agreement between coders rather than any correlational measure because the codes are categorical rather than ordinal. The generally acceptable standard for inter-coder percentage agreement is 80% (Hunt, 1997) and on average, both clairvoyance and precognition studies achieved agreements slightly above this figure, with a range from 55% to 100%. The percentage intercoder agreements for each variable are presented in Table 3.

Variable	Clairvoyance		Precognition	
		%		%
	<i>r</i>	agree- ment	<i>r</i>	agree- ment
Participant access to target prevented during trial	0.46	90	-	-
Experimenter blind	0.75	84	-	-
Participant access to records prevented after trial	0.23	71	-	-
Number of trials in study prespecified	0.68	74	0.70	77
Adequate randomisation	0.64	74	0.66	74
Checking of target records	0.41	65	0.46	55
Checking of calls	0.47	52	0.45	52
Checking of hits	0.51	68	0.68	77
Random sequence entry point selection made blind to calls	-	-	0.70	77
Average	0.52	72	0.61	69

Table 4: Percentage agreement and correlation between first author's and blind coding of studies for each procedural safeguard variable

Table 4 shows the extent of the agreement in coding between the first author and the blind coders for each procedural safeguard. The average percentage agreement for the procedural safeguard variables was somewhat lower than for the other moderator variables at 71%, with a range of 52% to 90%. Because the safeguard variables were ordinal, with a three-point scale (safeguard not present, not known if safeguard present, safeguard present), it was possible to

calculate post-hoc correlations between coders' judgements. Rosenthal (1991) argues that a correlation is a better indicator of coding reliability than percentage agreement, because a high percentage agreement could merely show that two judges share the same bias while coding (for example, of believing that optional stopping is generally unlikely). When the percentage agreements are considered for both moderator and procedural safeguard variables together, the overall percentage agreement is an acceptable 78%.

The first author's coding (corrected in the light of any indication of human error on her part if indicated by the blind judges, as discussed below) was used throughout as the data for the meta-analysis.

Determining database samples for analysis

Before the database was assembled, it had originally been intended to restrict moderator variable analyses to a subset of the whole database. This subset would include only those studies in which participants were aiming for high scoring (as opposed to chance or below-chance scoring) in both precognition and clairvoyance trials; otherwise, differences in participants' goals across studies would have rendered any comparison between precognition and clairvoyance effect sizes extremely difficult. This subset was to be called the "high aim sample". This latter sample was then itself to have a subset consisting only of study pairs in which the clairvoyance and precognition trials had been coded as having at least partially similar procedures (see Appendix 1, Moderator Variables). Although this latter group, the "comparable studies" sample, would clearly be the best sample from which to draw conclusions, it was expected to be considerably smaller than the high aim sample and was only to be examined as a check on the validity of conclusions drawn from the high aim sample. However, the comparable studies sample was larger than expected at 22 study pairs and only 7 study pairs smaller than the high aim sample. It was therefore decided after the coding was completed but prior to undertaking any analyses that the comparable studies sample would be the sample used for the meta-analysis. Unless otherwise stated, all further analyses in this paper are conducted on the comparable studies sample only.

RESULTS

All analyses were preplanned unless otherwise stated.

Cumulated outcome

The summary data for the comparable studies sample are presented in Table 5. The same data for the whole database and for the high aim sample are also presented for the sake of completeness. All analyses presented in the table were preplanned. In all three databases, both the unweighted and weighted Stouffer z s show a highly statistically significant cumulation for the precognition studies. For the clairvoyance studies, in the comparable studies sample and the high aim sample, both types of Stouffer z are statistically significant, although to a lesser degree than for the precognition studies, presumably due to the larger number of trials involved in the precognition studies. Neither the unweighted nor weighted Stouffer z was significant for clairvoyance studies in the whole sample.

Comparable	High aim	Whole
-------------------	-----------------	--------------

	studies sample (N = 22 pairs)		sample (N = 29 pairs)		database (N = 31 pairs)	
	Clairv.	Precog.	Clairv.	Precog.	Clairv.	Precog.
Total number of trials	108,995	190,100	154,530	214,160	182,596	271,710
Mean no. trials/study	4,954	8,641	5,329	7,385	5,883	8,765
Median no. trials/study	1,497	1,512	1,760	1,760	1,840	1,840
Unweighted Stouffer z (p)	2.68 (0.004)	4.15 (2x10 ⁻⁵)	1.84 (0.03)	4.21 (1x10 ⁻⁵)	1.04 (0.15)	3.99 (3x10 ⁻⁵)
Quality-Weighted Stouffer z (p)	2.81 (0.002)	4.78 (9x10 ⁻⁷)	2.18 (0.01)	4.73 (1x10 ⁻⁶)	1.58 (0.06)	4.66 (2x10 ⁻⁶)
Mean effect size (SD)	0.009 (0.046)	0.010 (0.050)	0.007 (0.045)	0.013 (0.043)	0.005 (0.040)	0.010 (0.040)
t_{diff} between mean effect sizes	$p(2-t) = .65,$ $t = 2.08, 21 df$		$p(2-t) = .48,$ $t = 0.71, 28 df$		$p(2-t) = .4,$ $t = 0.85, 30 df$	

Table 5: Summary data for the comparable studies sample, high aim sample and whole database

Comparing precognition and clairvoyance effect sizes

Surprisingly, the findings from Table 5 indicate that when forced-choice experiments are conducted under similar conditions, the effect sizes for precognition (0.010) and clairvoyance (0.009) are nearly identical. Neither the whole sample nor the high aim sample indicate a different picture, with the t -tests showing no significant difference between the effect sizes within any of the three databases. There appear to be no distributional peculiarities about the differences in precognitive and clairvoyant effect sizes within individual study pairs in the comparable studies database. Out of the twenty-two studies only four revealed individually a significant difference (using post-hoc chi-square analyses) between clairvoyance and precognition, two studies coming out in favour of precognition (Thouless, 1949; Freeman, 1962) and two studies coming out in favour of clairvoyance (Buzby, 1968; Nash 1982).

Larger studies would be expected to give more accurate estimates of effect size, given their lower variance, so N-weighted effect sizes were calculated post-hoc. Again, the two mean effect sizes for clairvoyance and precognition were remarkably similar to each other at 0.030 (clairvoyance) and 0.034 (precognition).

It is clear that this database shows no evidence for thinking that either clairvoyance or precognition obtains better results. Before moving on to a discussion of whether the two types of task appear to involve different moderator variables, we will briefly discuss whether study quality and coding reliability appear to be important issues in this meta-analysis.

Procedural safeguards and effect sizes

Studies were assigned points according to whether they had reported procedural safeguards, with a larger number of points indicating higher quality. Clairvoyance studies could obtain between 0 and 16 points, with the actual range assigned being from 1 to 16 (mean = 10.6,

median = 10). Precognition studies could gain a minimum of 0 points and a maximum of 12, and the actual range covered this spread (mean = 7.5, median = 7). Table 2 shows the percentage of studies for which the individual safeguards were known to be present, of uncertain status and known to be absent and the correlation coefficients between the presence of individual procedural safeguards and effect sizes in the studies. Negative correlations indicate a tendency for studies with fewer safeguards to have higher effect sizes. No significant negative correlations were obtained, indicating that if study quality did play a role in the meta-analysis, it was not an obvious one. However, the 22 study pairs offer relatively low statistical power for analyses of this kind so the absence of evidence for a relationship between effect size and the presence of procedural safeguards should not be taken as evidence of absence.

Moderator variables and effect size

Because it was not known in advance what the characteristics of the database would be and, therefore, which analyses would or would not be possible, the precise statistical analyses on the moderator variables were not preplanned. No analysis was conducted when a comparison group contained five or fewer experiments. The results of the moderator variable analyses are summarised in Table 5. No statistically significant effects were found.

Optimal conditions

In their meta-analysis of forced-choice precognition studies, Honorton and Ferrari (1989) identified post-hoc what conditions appeared to be optimal for effect size. Too few studies in the current database used all of the conditions that they identified as optimal (selected subjects, individual testing, trial-by-trial feedback) for a preplanned contrast with studies meeting their definition of suboptimal studies (unselected subjects, no feedback, group testing). Nevertheless, it may be useful to identify which conditions the present meta-analysis suggests as being psi-conducive for precognition and clairvoyance studies.

Unlike the Honorton and Ferrari meta-analysis, none of the moderator variable analyses in the present meta-analysis were statistically significant. It is possible, though not certain, that the lack of significant relationships could have been due to low statistical power in the present meta-analysis, which contains only 22 study pairs. This being the case, it is reasonable to use the contrast in effect sizes within each type of ESP task, rather than statistical significance, as a suggestive guide to potentially interesting variables for future research. Using this approach, for the precognition studies, an effect size of roughly 0.040 was chosen post-hoc as a criterion of indicating a possibly psi-conducive procedure, this effect size being approximately the mean effect size achieved by the highest-scoring contrast groups of precognition studies (as shown in Table 6).

Variable	Clairvoyance				Precognition			
	N	ES	t_{diff}	p	N	ES	t_{diff}	p
Trials interspersed	7	-0.010	-1.74	.10	7	0.040	1.29	.22
Trials separated into experiments	11	0.024			11	0.010		
Special and selected participants	7	0.014	0.30	.77	7	0.017	0.18	.86

Unselected participants	15	0.007			15	0.013		
Motor response	7	0.021	0.80	.44	-	-	-	-
Verbal or written response	15	0.004			-	-		
Participants encouraged to succeed	9	0.010	0.17	.86	7	0.010	-0.54	.59
Participants not encouraged	13	0.007			15	0.020		
Individual target presentation	6	0.018	0.73	.48	6	0.039	1.30	.21
<i>En bloc</i> target presentation	15	0.002			15	0.009		
Participant feedback at end of run	6	0.004	-0.59	.57	7	0.012	-1.13	.28
Participant feedback trial-by-trial	6	0.021			6	0.042		
Experimenter feedback at end of study	-	-	-	-	6	-	-0.55	.59
Experimenter feedback after each run	-	-			8	0.005		
Participant touches target	9	-0.001	-0.84	.41	-	-	-	-
Participant does not touch target	13	0.016			-	-		
Target selected seconds after guess	-	-	-	-	6	0.042	1.69	.11
Target selected longer after guess	-	-			15	0.004		

All p-values are two-tailed.

Table 6: Moderator variable analyses for clairvoyance and precognition studies

There were three variables with a mean effect size of this magnitude (individual presentation of targets, trial-by-trial feedback, target selection within seconds). Notably, these three optimum procedures for precognition tend to occur together; trial by trial feedback involves individual presentation of the targets and in forced-choice precognition experiments trial by trial feedback will often entail that targets are selected soon after the guess. Therefore, even if further research indicates that the relatively high mean effect sizes in these groups are not artefactual, it will be important to determine whether any effects are due to only one or two of the variables rather than all three.

For the clairvoyance studies, a mean effect size of roughly 0.020 was chosen as the criterion for selecting variables that might be psi-conducive on a similar basis as the criterion selection for the precognitive studies. Four possibly psi-conducive procedures were identified by this means (motor response, individual target presentation, trial-by-trial feedback and no physical contact with target material). This group of optimum clairvoyance procedures differs from the precognitive one only by the inclusion of motor response (precognitive studies necessarily having no participant contact with the target material at the time of the trial). As with the potentially psi-conducive variables identified for precognition studies, however, these variables

are only tentatively suggested as worthy of further study because of the lack of statistical significance of the comparison between mean effect sizes in the contrast groups.

CONCLUSIONS

This database provides no evidence to support the idea that clairvoyance works better than precognition. When tested under similar conditions - and for the most part with the same participants - the resultant effect sizes for clairvoyance and precognition appear to be remarkably alike. All but two of the study pairs contained exactly comparable studies. Nevertheless the results from this meta-analysis are generalizable only in relation to the criteria used for determining studies to be comparable in this particular meta-analysis. It is possible that other meta-analyses using different coding criteria may produce different results. However, in the meantime, the burden of proof rests with those who would argue for a difference between effect sizes under real-time and future ESP.

The database for this meta-analysis provided no strong evidence for any potentially psi-conducive moderator variables or for differences between precognition and clairvoyance in terms of which variables are important. It is possible that the lack of findings was due to low statistical power, given that the database had relatively few studies and had only a very small overall effect.

The results from this meta-analysis suggest that theories about parapsychological phenomena should not rest on assumptions that require a difference in precognition and clairvoyance effect sizes. The fact that the precognition studies in the present database involved, on average, more trials than the clairvoyance studies suggests that this difference in sample size might be true of precognition and clairvoyance studies in general, lending support to the criticism of Tart's (1983) analysis that we offered in the introduction of this paper. If nothing else, this meta-analysis illustrates just how far parapsychology still has to go in understanding the phenomena it investigates.

REFERENCES

General references

- Braud, W.G. & Schlitz, M. (1991). Consciousness and interactions with remote biological systems: Anomalous intentionality effects. *Subtle Energies*, **2**, 1-46.
- Carington, W. (1945). *Telepathy: An Outline of its Facts, Theory and Implications*. London: Methuen. P.91
- Glass, G.V., McGaw, B., & Smith, M.L. (1981). *Meta-analysis in Social Research*. London: Sage Publications.
- Honorton, C., & Ferrari, D.C. (1989). "Future telling": A meta-analysis of forced-choice precognition experiments, 1935-1987. *Journal of Parapsychology*, **53**, 281-309.
- Hunt, M. (1997). *How Science takes Stock: The Story of Meta-analysis*. New York: Russell Sage Foundation.
- Milton, J. (1998). Meta-analysis of free-response ESP studies without altered states of consciousness. *Journal of Parapsychology*. In press.

- Mundle, C.W.K. (1978). Does the concept of precognition make sense? In Jan Ludwig (Ed.), *Philosophy and Parapsychology*. New York: Prometheus.
- Rosenthal, R. (1991). *Meta-analytic Procedures for Social Research*. London: Sage Publications.
- Tart, C.T. (1983). Information acquisition rates in forced-choice ESP experiments: Precognition does not work as well as present-time ESP. *Journal of the American Society for Psychical Research*, **77**, 293-311.
- Wolf, F.M. (1986). *Meta-analysis: Quantitative Methods for Research Synthesis*. London: Sage Publications.

Chronological Bibliography of Studies in Meta-analysis

- Stuart, C.E. (1941). An analysis to determine a test predictive of extra-chance scoring in card-calling tests. *Journal of Parapsychology*, **5**, 99-137.
- Rhine, J.B. (1941). Terminal Salience in ESP performance. *Journal of Parapsychology*, **5**, 183-244.
- Rhine, J.B. (1942). Evidence of precognition in the covariation of salience ratios. *Journal of Parapsychology*, **6**, 111-143.
- Humphrey, B.M. & Pratt, J.G. (1941). A comparison of five ESP test procedures. *Journal of Parapsychology*, **5**, 267-292.
- McMahan, E.A. & Rhine, J.B. (1947). A second Zagreb-Durham ESP experiment. *Journal of Parapsychology*, **11**, 244-253.
- Thouless, R.H. (1949). A comparative study of performance in three psi tasks. *Journal of Parapsychology*, **4**, 263-273.
- Anderson, M. & Gregory, E. (1959). A two-year program of tests for clairvoyance and precognition with a class of public school. *Journal of Parapsychology*, **23**, 149-177.
- Freeman, J.A. (1962). An experiment in precognition. *Journal of Parapsychology*, **26**, 123-130.
- Osis, K. & Fahler, J. (1965). Space and time variables in ESP. *Journal of the American Society for Psychical Research*, **59**, 130-145.
- Buzby, D.E. (1967). Subject attitude and score variance in ESP tests. *Journal of Parapsychology*, **31**, 43-50.
- Buzby, D.E. (1968). Precognition and clairvoyance as related to the draw-a-man test. *Journal of Parapsychology*, **32**, 244-247.
- Nash, C.S. & Nash, C.B. (1968). Effect of target selection, field dependence, and body concept on ESP performance. *Journal of Parapsychology*, **32**, 248-257.
- Schmidt, H. (1969). Precognition of a quantum process. *Journal of Parapsychology*, **33**, 99-108.
- Schmidt, H. (1969). Clairvoyance tests with a machine. *Journal of Parapsychology*, **33**, 300-306.
- Haraldsson, E. (1978). ESP and the defence mechanism test (DMT). A further validation. *European Journal of Parapsychology*, **2**, 104-114.
- Haraldsson, E. & Johnson, M.U. (1979). ESP and the defence mechanism test (DMT) Icelandic study no. III. A case of experimenter effect? *European Journal of Parapsychology*, **3**, 11-20.
- Winkelman, M. (1981). The effect of formal education on extrasensory abilities: The Ozolco study. *Journal of Parapsychology*, **45**, 321-336.

- Nash, C.B. (1982). ESP of present and future targets. *Journal of the Society for Psychical Research*, **51**, 374-377.
- Crandall, J.E. & Hite, D.D. (1983). Psi-missing and displacement: Evidence for improperly focused psi? *Journal of the American Society for Psychical Research*, **77**, 209-228.
- Tedder, W. (1984). Computer-based long distance ESP: An exploratory examination. In Rhea A White & Richard S Broughton (Eds) *Research in Parapsychology 1983*, Metuchen, NJ: Scarecrow Press. 100-101
- Johnson, M. & Haraldsson, E. (1984). The defence mechanism test as a predictor of ESP scores. *Journal of Parapsychology*, **48**, 185-200.
- Haraldsson, E. & Johnson, M. (1985). The defence mechanism test (DMT) as a predictor of ESP performance: Icelandic Studies VI and VII. *Research in Parapsychology 1985*. London: Scarecrow Press. 43-44.
- McDonough, B.E., Warren, C.A. & Don, N.S. (1990). Precognition and real-time ESP performance with an exceptional subject: A replication attempt. *Journal of Parapsychology*, **54**, 141-150.
- Don, N.S., McDonough, B.E. & Warren, C.A. (1992). Psi testing of a controversial psychic under controlled conditions. *Journal of Parapsychology*, **56**, 87-96.
- Haraldsson, E. (1993). Are religiosity and belief in an afterlife better predictors of ESP performance than belief in psychic phenomena? *Journal of Parapsychology*, **57**, 259-73.

APPENDIX 1: CODING CRITERIA FOR THE META-ANALYSIS

Moderator variables

Journal: [EJP/Research Letter]/[JASPR/Proceedings]/[JP]/[JSPR/Proceedings]/[RIP]

Study pair procedural similarity: Are the studies' procedures directly comparable? [No]/[Partially]/[Yes]

Within vs between subjects design: [Between subjects design (or mixed)]/[within subjects design (i.e., same subjects in both experiments)]

Participant knowledge of trial type: [Subject knows only that clairvoyance and precognition will be interspersed (either in runs or trials) and mixed; they won't know which trials are which at time of guess]/[Subject knows when it is a precognitive trial and when it is a clairvoyance one]/[Subject knows studies only as a precognition study]/[Subject knows studies only as a clairvoyance study]/[Separate studies with different subjects]/[Unspecified, ambiguous or other]

Experimenter knowledge of trial type: Does the experimenter know what condition (clairvoyance or precognition) is operating? [Experimenter knows only that will be interspersed and mixed; experimenter need not be present during trials]/[Experimenter knows when precognition and when clairvoyance if interspersed]/[Separate experiments at different times or experiments radically different in procedure]/[Other or unknown]

Interspersion of precognition and clairvoyance trials: [Precognition and clairvoyance

interspersed in runs or blocks of runs]/[Precognition and clairvoyance interspersed in trials]/[Unknown or other]/[Precognition and clairvoyance separate, not interspersed}

Participant population type: [Special subject]/[Special population]/[Selected]/[Experimenter]/[Volunteer (paid or unpaid)]/[Children (under 18)]/[Students (college or university)]/[Animals]/[Mixed/unspecified]

Response modality: [Open matching (key card face up)]/[Blind matching (keycard face down)]/[Matching, unspecified or mixed (or precognitive matching)]/[Select]/[Call, verbal]/[Call, written]/[Call, unspecified or mixed]/[Mixed or unknown]

Expectation of success mindset: [Yes]/[No]

Target type: [Symbol]/[Colour]/[Picture]/[Word]/[Number]/[Other or unknown]/[Mixed]

Target presentation mode: [One at time]/[DT & variants]/[Sheet layout (include RNT)]/[Mixed, other or unspecified]

Timing of experimenter's feedback: [At end of whole study or experiment]/[At end of subject's participation]/[At end of run or blocks of runs]/[Trial by trial]/[Mixed]/[Unknown, other or ambiguous]

Participant feedback of each target: [No]/[Yes]/[Mixed or possibly yes]/[Unknown]

Participant told score: [No]/[Yes]/[Mixed]/[Unknown]

Timing of participants' feedback: [Never]/[Delayed (include postal results sent back)]/[On completion of study or experiment or series]/[On completing their participation]/[At end of run or blocks of runs]/[At end of each trial]/[Unknown]/[Mixed]

Individual vs group testing: [Individually (include pairs)]/[In group]/[Mixed]/[Unknown or ambiguous]

Participant contact with experimenter: [Subjects had direct contact with experimenter]/[Subjects had only indirect contact with experimenter (e.g., mail)]/[Subjects had no contact with experimenter]

Predicted direction of results: [Chance]/[Psi hitting]/[Psi missing]/[Unknown, too ambiguous or unspecified]

Target-participant distance: [None (ie subject touches target)]/[Same room/vacinity as subject]/[Same building]/[Same city]/[Same country]/[Abroad]/[Mixed]

Precognitive timespan: [Seconds or less]/[Minute or more]/[Hour or more]/[Day or more]/[Week or more]/[Month or more]/[Unknown, mixed or ambiguous as to when target selection took place (eg if you don't know whether it was between runs or at the end of the experiment)]

Complexity of target selection method: Complexity of target selection, precognition studies only. [Hand shuffling, die casting etc.]/[Use of mechanical shuffler etc. with rotations etc. decided by random or prespecified means]/[Entry point (in RNT, deck of cards etc.) decided by hand shuffling; or use of PRNG, true RNG or mechanical shuffler]/[Entry point decided by die casting etc plus complex calculation]/[Use of stock market figures, weather etc.]

Procedural safeguard variables

Clairvoyance only

Participant access to target prevented during trial: Possible access to target whilst guessing. [Subject has access to target material by direct touch]/[Subject may have access - e.g., target in not explicitly tamper-proof container or not explicitly sealed opaque envelope or target could be guessed at or there was poor supervision or screening is poor or DT and cards not in box]/[Subject had no access to target material - screening was solid (eg sealed opaque envelope) or target was in different room. This includes screen-touch matching, other forms of screening or adequate precautions were taken]

Experimenter blind: [Receiver's experimenter not blind to target]/[Experimenter possibly not blind - eg DT or ambiguous. Includes cases where cards stated to be in box, because here experimenter may have put cards in box]/[Experimenter blind or not present]

Participant access to records prevented after trial: Possible access to/modification of records after guessing. [Participant could access records]/[Participant possibly able to access records (for example, if S could distract experimenter)]/[Participant not able to access records. Includes use of opaque, sealed envelopes when participants supervised. Includes computer studies, although for special subjects the computer should be explicitly stated to be tamper proof]

Both clairvoyance and precognition

Number of trials in study prespecified: I.e. is optional stopping a possibility? [No. of trials not explicitly prespecified and number of trials uneven (i.e., not ending in 0 nor exactly divisible by p of hit)]/[Mixed (ie some use even nos of trials and some don't) or prespecified but not carried out as prespecified or appear to be prespecified because of no of trials even across conditions]/[Explicitly prespecified no. of trials or classroom testing where number of subjects cannot be prespecified, but prespecified that those present will take a prespecified number of trials each]

Adequate randomisation: [Hand-shuffling, die casting without banking board or with piped dice]/[Die casting with banking board and non-pipped dice, P/RNG without any check of its randomness; mechanical shuffler with no check of its randomness]/[Actual target order checked for randomness use of Random Number Table (RNT); empirical cross-check (NOT "control" shuffle); P/RNG or mechanical shuffler with check for the randomness of its running either before and/or after the experiment or on target set itself. Or extensively checked mechanical shuffler]

Checking of target records: [No duplicate recording of target or unsure but seems unlikely that even if there were a duplicate recording that the two target records will have been cross-

checked for either experimenter or subject fraud]/[Duplicate, or automatic independent recording of target]

Checking of calls: [No duplicate recording of S's calls]/[Observed recording of S's calls or unsure but looks likely that there were duplicate records]/[Duplicate or automatic (includes photos) recording of S's guesses]

Checking of hits: [Hits not checked independently by two separate people; unsure but unlikely]/[Recording of hits observed or experimenter does double check; unsure but likely]/[Independent or automatic checking of hits]

Precognition only

Random sequence entry point selection made blind to calls: [Person deciding entry point to randomisation or person recording the random numbers not blind to calls]/[Unknown]/[Person deciding entry point is blind to calls (inc. computer)]

CHARACTERIZATION OF STATES OF CONSCIOUSNESS BASED ON GLOBAL DESCRIPTORS OF BRAIN ELECTRICAL DYNAMICS

Jiri Wackermann

*Institut für Grenzgebiete der Psychologie
und Psychohygiene, Freiburg i.B., Germany*

ABSTRACT

The spatiotemporal dynamics of the electrical field of the brain reflects variations in the global functional state of the brain. Functional states of the brain are related to states of consciousness in their normal or pathological variety. Thus, an objective characterization of states of consciousness should involve quantitative assessment of the dynamics of brain's electrical fields.

Traditionally, electrical activity of the brain (EEG) is analyzed by methods operating on the frequency domain (FFT). This approach, rooted in signal engineering and clinical empiricism, dominated the research of brain functions for decades and focused onto frequency spectra of local activities, neglecting the spatial aspect of brain fields.

An alternative approach, established by 'Zurich school', has focused onto space domain; each snapshot of brain's field is taken in its entirety as a unit of analysis. These methods provide quantitative description of brain field topographies, going beyond naive 'brain mapping' techniques. The spatial analysis can be extended by the state-space approach that assesses brain's electrical dynamics by means of geometrical properties of state-space trajectories. Thus, an epoch of several seconds of whole-scalp EEG can be reduced to few 'macroscopic' descriptors, quantifying the global functional state of the brain.

A system of three state variables has been proposed (Wackermann, 1996): Σ (integral power), Φ (generalized frequency), and Ω (measure of spatial synchronization). These quantities can be taken as coordinates of a 3-dimensional 'macrostate' space in which variety of brain functional states can be visualized and identified. In contrast to current trends towards refinement of spatial and temporal resolution, this approach intends to provide a 'holistic', phenomenological, macroscopic description of brain as a whole.

The usefulness of the approach sketched above was proven in studies of sleep EEG (Szelenberger et al., 1996) and vigilance fluctuations (Wackermann & Szelenberger, 1996). The spatial complexity itself occurred to be a sensitive indicator in studies of brain state 'shifts' induced by sensory input, drug administration, or a psychiatric disturbance [4-7] (Kondakor et al., 1997a, 1997b; Saito et al., 1998; Yagyu et al., 1996). Thus we can expect that global description would be useful also for a quantitative characterization of altered states of consciousness.

REFERENCES

- Kondakor I. et al. (1997a). Multichannel EEG fields during and without visual input: frequency domain model source locations and dimensional complexities. *Neurosci. Lett.* 226:49-52
- Kondakor I. et al. (1997b). Global Omega-complexity indicates change of 47-channel EEG after single doses of piracetam. Presented at 8th World Congress of ISBET, Zurich, March 6-8.
- Saito N. et al. (1998). Global, regional and local measures of complexity of multichannel EEG in acute, neuroleptic-naive, first-break schizophrenics. *Biol. Psychiatry* 43:794-802.
- Szelenberger W. et al. (1996). Analysis of complexity of EEG during sleep. *Acta Neurobiol. Exp.* 56:165-169.

- Wackermann J. (1996). Beyond mapping: estimating complexity of multichannel EEG recordings. *Acta Neurobiol. Exp.* 56:197-208.
- Wackermann J. & Szelenberger W. (1996). Towards a quantitative phenomenology of brain functional states. Presented at the 9th Biennial IPEG Meeting, Prague, September 12-14,.
- Yagyu T. et al. (1996). Diazepam effect on global dimensional complexity and Omega complexity of multichannel EEG. Presented at the 9th Biennial IPEG Meeting, Prague, September 12-14.