

**ON THE NATURE OF ANOMALOUS PHENOMENA:
ANOTHER REALITY BETWEEN THE WORLD OF SUBJECTIVE CONSCIOUSNESS AND THE
OBJECTIVE WORLD OF PHYSICS?**

Dick J. Bierman
Utrecht University

Abstract

Cumulating evidence suggests that anomalous correlations occur between mental (conscious and non-conscious) states and apparently unrelated physical or mental phenomena at a distance in space and time. In spite of the fact that the evidence is very strong, these correlations are difficult to replicate. Several examples are given of 'failures' to empirically replicate¹ original anomalies. It is speculated that this failure to replicate, rather than indicating that the original findings are due to statistical flukes or errors, suggests that when consciousness interacts with matter, an underlying reality arises. This reality is somewhere in between the purely objective shareable reality and the purely subjective reality of one's individual consciousness.

Efforts to 'push' anomalous phenomena observed in this intermediate reality into the objective one apparently destroy the phenomena. Possible explanations within a physical and within a system theoretical model are discussed. The physical model is based upon an analysis of the role of consciousness in the so-called 'Measurement Problem' in Quantum Physics. Based upon these discussions a new systematic experimental approach for the study of anomalous phenomena is suggested

¹ We use the term 'failure to replicate' here loosely. I.e. if an original effectsize shows a significant regression to zero we call this a failure to replicate the original finding although the elements of the database might not differ in a statistically significant way.

Introduction

History is replete with accounts of so-called psychic phenomena like telepathy or precognition. The use of oracles during Greek civilization is just one example to illustrate that belief in the reality of these phenomena was widespread for millennia. However with the advent of modern western science and especially with the success of physics and derived disciplines in explaining and controlling natural phenomena, belief in, and serious study of such 'unexplainable' phenomena has generally been viewed as superstitious or even as anti-scientific behaviour.

The findings of experimental psychology seem to support this view. Human memory is extremely fallible as has been demonstrated by research on the false-memory syndrome (Loftus, 1997). As suggested long ago by William James by his statement: ... *whilst part of what we perceive comes through our senses from the object before us, another part (and maybe the larger part) always comes out of our head* (William James, 1890), perception is now generally seen as a constructive process and perceptual experience only partly reflects objective reality (Gordon, 1996). Thus human accounts of spontaneous or natural events are not acceptable as objective scientific evidence.

Nevertheless, during the last 100 years, a limited number of scholars have persisted in serious study of the alleged paranormal. Confronted with the shortcomings of personal accounts and the fact that deception and fraud were unavoidable outside of a well-controlled environment, most serious work was done within experimental laboratories.

It is fair to say that not many scientists share Dennett's opinion that consciousness 'has been explained' already. The growth of interest in what arguably is the most complex scientific topic has brought about a renewed interest in psi phenomena that, if real, would pose strong constraints on any comprehensive theory of mind. To illustrate this point, Dennett's Multiple Drafts model (Dennett, 1991) is not capable of explaining precognition because precognition is a violation of temporal causal principles while Dennett's model is purely causal. Hence Dennett's promise that he would "commit suicide if paranormal phenomena turned out to be real" (personal communication).

Dennett's confidence notwithstanding, decades of cumulating laboratory evidence strongly suggests that real correlations exist between mental states and randomized events that are distant in space and time, some even in the future. The laboratory findings are generally weaker than those reported from the field and because of the experimental paradigms used, they are basically statistical in nature. In other words they need a large number of trials to reach statistical significance. What is worse is that it is difficult to replicate the phenomena even when allowing for their intrinsic statistical character. Because of the claimed transcendental nature of the phenomena the experimenter is an intrinsic part of the experiment and therefore replication from one experimenter to another is not expected to be easily accomplished. Replication by independent scientists is of course the requirement for a phenomenon to be considered objective and real. Therefore the controversy about the reality of these phenomena continues in spite of the cumulating evidence. In this chapter we will try to formulate a theoretical position which assimilates both points of view.

The Database of Experimental Parapsychology

In a review article in *Statistical Science*, Utts argued that for four different parapsychological paradigms there is overwhelming meta-analytic evidence for the occurrence of anomalous correlations between either mental states of separated people, or mental states and some objective material state (Utts, 1991).

The first paradigm concerns the so-called Ganzfeld telepathy experiments. In these double blind experiments one participant (*the Receiver*) is brought to a slightly altered state of consciousness (hypnagogic state) and reports his/her experiences while at a remote location another participant (*the Sender*) is looking at a picture or video-clip (the target) randomly picked out of a pool of many such pictures or clips. After half an hour *the Receiver* is shown a copy of the actual target along with three decoys, in random order, and is requested to pick the actual target. The probability for the participant picking the actual target is thus 1/4. Utts reports two meta-analysis with mean effect sizes of 0.28 (experiments from 1970-1985) and 0.20 (for the 11 automated series performed between 1983-1990). These effect sizes are quite unlikely to occur by chance as can be seen in the column 'associated p-value' of Table 1.

The second paradigm that Utts reviews concerns card guessing experiments. A meta-analysis of all precognitive card-guessing experiments from 1935-1987 shows a mean effect size of 0.033. Although this effect size is much smaller than the ones reported for the Ganzfeld paradigm the statistical significance is impressive due to the large number of trials involved.

The third and fourth paradigm for which meta-analyses were done are of the so-called mind-over-matter kind. Participants are asked to 'influence' the behavior of electronic Random Number generators (based on quantum mechanical noise). The mean effect size in the 597 studies reviewed was only 0.00032 but the

total result was still quite significant. Finally 148 studies designed to test whether or not a conscious intention to influence the outcome of tossing dice, biases the outcome distribution of dice throwing were reviewed. Here the over-all effectsize was 0.0122. A review of her findings is presented in the first 5 rows of table I.

<i>Paradigm</i>	<i>Mean effect size</i>	<i>Associated p-value</i>	<i>Original publication</i>
Mind over Matter Dice	0.0122	$\sim 10^{-14}$ ($z=7.617$)	Radin & Ferrari (1991)
Ganzfeld telepathy (1972-1985)	0.28	$3.37 * 10^{-11}$	Rosenthal 1986
Ganzfeld telepathy automatic (~1990)	0.20	$5 * 10^{-5}$	Bem & Honorton 1994
Card Guessing (precognition)	0.033	$6.3 * 10^{-25}$	Honorton & Ferrari (1989)
Mind over Matter RNG's	0.00032	$\sim 5 * 10^{-5}$ ($z = 4.1$)	Radin & Nelson (1989)
Mind over Matter Biological systems	0.33	$2.85 * 10^{-14}$	Braud & Schlitz

Table I: A review of important meta-analytic results used to argue the reality of psi-phenomena

An omission in Utts' review of 'convincing' parapsychological meta-analyses concerns a research program by the Mind-Science laboratory that lasted from the late seventies to the early nineties (Braud & Schlitz, 1991). The major research effort was concentrated on the effect of distant attention on remote biological systems and the over-all results were again highly significant (see last row table 1).

It should be stressed that, according to Utts, normal 'explanations' for these significant findings like selective publication or methodological errors were ruled out. Thus, from these findings, the case for the reality of so-called psi phenomena looks very strong. However a closer look at the data reveals an intriguing sequential structure known as the 'decline effect'.

The decline effect

Originally the label 'decline effect' was used to describe decreasing performance of subjects when doing a sequence of trials. Actually the decline in performance was generally followed by a rebound in performance when the end of the run of trials approached. Rhine used the position effects that were established *post hoc* as an argument for the reality of the psi effect (Rhine, 1969). The position effects then were attributed to intervening psychological variables like motivation and boredom. However it soon became clear that apart from sequential effects within a run of trials it was quite difficult to exactly 'replicate' (see footnote 1) the results of original studies: a sequence of similar studies tended to show a decline of effect size from study to study.

Dice throwing studies

As a first example we consider the meta-analytical results of Radin & Ferrari (Radin & Ferrari, 1991) concerning dice-throwing experiments. These experiments started in the thirties based upon anecdotal evidence that some people were consistently more lucky when throwing dice than others. Using dice throwing devices Rhine and other researchers tried to test the idea that these individual differences were due to difference in psychokinetic abilities. A regression plot of the normalized (z) score of each of these studies versus year reveals a steady decline with a regression coefficient of -0.053 ($p=0.0020$)².

² One outlier was removed. This does not affect the regression statistics at all.

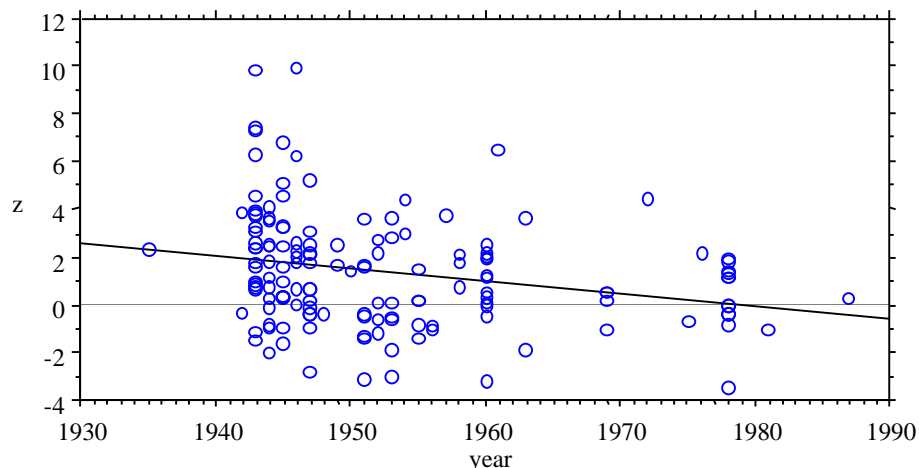


Figure 1: Regression plot of normalized scores versus year of all dice throwing studies.

Ganzfeld studies

The first Ganzfeld ESP studies were done in the early seventies and continue to date. These experiments were used for the first meta-analysis ever done on psi experiments (Honorton, 1985). This meta-analysis eventually resulted in a joint report of a well-known sceptic of the paranormal and a major psi researcher in the field of ganzfeld studies (Hyman & Honorton, 1986). In the report both authors concluded that the overall results differed from chance expectation, but there remained a difference of opinion about the interpretation of this difference. Some recommendations were made to standardize procedures and eliminate potential methodological weaknesses. In 1994 a second review was published in a main stream journal concerning a series of studies that had followed closely the prescriptions described in the joint report (Bem & Honorton, 1994). Although these results did globally confirm the earlier meta-analysis, a closer inspection yielded some findings that put the claimed replicability into question. The earlier 1985 meta-analysis concerned ganzfeld experiments with static targets having a mean scoring rate of about 33%. However the 1994 data were a mixture of static and dynamic targets also yielding a mean scoring rate of about 33%. When using the results for static targets alone the scoring rate in the 1994 analyses dropped to a non significant 28% level. Thus there was an indication of a decline in scoring rate when the analyses were restricted to comparable experiments.

Also the 1994 analysis was performed on the results of only one research-group although many more experiments had been done by other researchers in the period from 1985. A regression plot (figure 2) of ALL studies from the early 1970's to 1994 revealed a steady and significant decline in effectsize with a regression coefficient of -0.027 ($p=0.022$) (see Bierman, 1995)³.

³ The analysis is not sensitive for the removal of three outliers. In fact the analysis is conservative because part of the data produced by PRL using dynamic targets (which runs counter to the decline effect) is included twice in the database because it is unclear which auto-ganzfeld data from the final publication (Honorton et al, 1990) had been reported before.

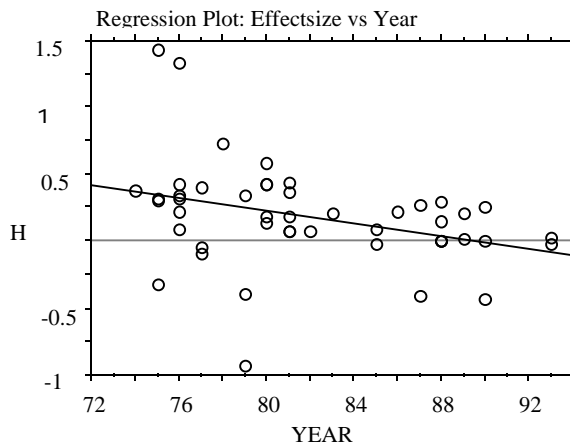


Figure 2: Regression plot of the effectsizes versus year of all Ganzfeld experiments in the period 1974-1994.

Not only was there a steady decline in scoring rate but also the variance in the experimental outcomes declined considerably. There is strong difference in variance between the pre-1985 studies and the post-1985 studies ($F= 4.903, df1=27, df2=14, p = 0.018$). Adding the recent data from 1994 to 1997 decreases the regression coefficient to -0.017 but increases the statistical significance of this decline considerably ($p = 0.0013$)

A rather spectacular example of the decline in effectsize in ganzfeld experiments happened when in 1995 two independent groups, one from Durham NC, the other from Amsterdam, published the data of the first part of their respective ganzfeld series (Broughton & Alexander, 1995; Bierman, 1995). The over-all hit rates were: 33% ($N=100$) and 38.2% ($N=76$). One year later the series were finished with the following astonishing figures for the second part: Durham 13.7% ($N= 51$) and Amsterdam: 15.6% ($N=32$) (Broughton & Alexander, 1996; Wezelman & Bierman, 1997). Thus results of the first and second part of both series differed significantly between the years while within the same year the groups replicated each other as if some outside factor in 1996 had influenced both groups to go from hitting into missing (fig. 3)

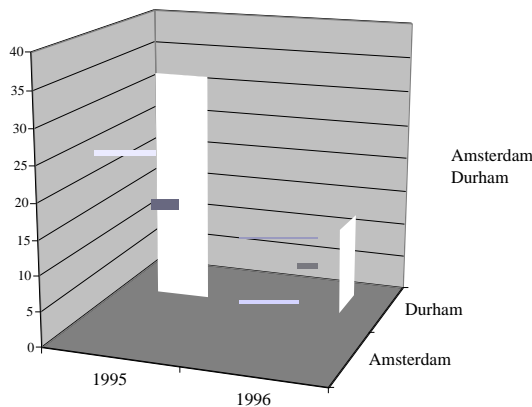


Figure 3: Decline in scoring rate in Ganzfeld ESP experiments in two independent laboratories in two consecutive years.

Recently a new meta-analysis on Ganzfeld experiments up to 1998 was reported concluding that the post 1985 data do not show any effect (Milton & Wiseman, 1999). Although this conclusion is quite dependent on the choice of criteria to include studies in the meta-analysis there is no doubt that when adding the data from 1994 to 1998 the decline effect still is present. When including the more recent and significant data that were omitted in the Milton & Wiseman analysis the decline is dampened considerably. One could even interpret these data to represent a recovery from the decline effect.

PK effects on RNG's

Experiments where true Random Number Generators have been used as a target in psychokinesis experiments are very interesting because this paradigm is generally completely automated and therefore:

1) the experimenter is less involved in doing the experiment, very often participants are doing the experiments on their own. And

2) the quality of the studies is more stable than for any other paradigm.

A regression plot for the database of about 600 studies reported by Radin and Nelson (Radin & Nelson, 1989) doesn't show a clear decline. However this database is a mixture of data, the majority contributed by the PEAR group and by a single experimenter (Helmut Schmidt) using different procedures than the other groups. In the Princeton data and for a part in the Schmidt data, we have three possible target directions namely high aim, low aim and no-aim. The other researchers, although also using the methodologically sound target counterbalancing procedure, do not split the results in this way. If we exclude the data obtained with the first procedure we are left with 264 studies. Plotting these reveals a sequential effect that differs from the straight decline effects reported above (fig. 4).

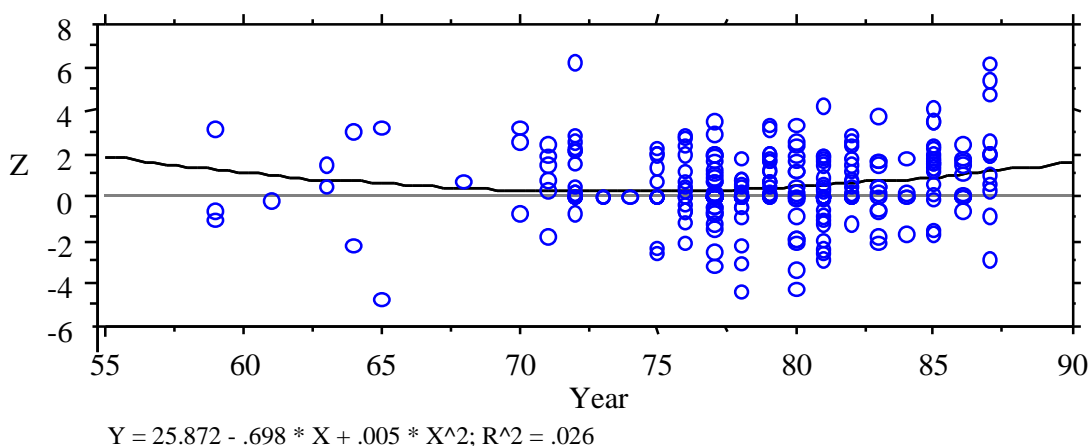


Figure 4: Polynomial regression plot of all RNG-PK study results under exclusion of PEAR and Schmidt data

Rather than just declining, the effect size seems to recover after a period of about 15 years. Both polynomial regression coefficients are significant ($p < 0.03$). What we might see here is the onset of a kind of oscillatory behaviour which fits well with the models we will discuss in the theoretical section. Note by the way that in spite of a period of about 10 years where no effects were found, the total residual database provides strong support for the existence of a true anomaly (mean z-score is 0.541, $N=264$, $t=5.34$, $p < 0.00001$).

When we analyze the data for the high aim and low aim separately we find regression coefficients indicating an incline with time for the low aim and decline in time for the high aim in the RNG experiments of Helmut Schmidt and in the PK-RNG data of the PEAR group. However none of these values are significant.

Recent major attempts though to replicate the findings of the PEAR RNG-PK work in two independent laboratories as well as by the PEAR group themselves, resulted in a failure. Detailed data are not yet public so it remains to be seen if this is another example of the 'decline effect' (personal communication).

PK effects on biological systems

In the research program of the Mind Science foundation many biological target systems were used (Braud & Schlitz, 1991). We will restrict ourselves here to a series of identical studies that was replicated often enough to check for a potential decline. In these studies a remote 'influencer' tried to increase or decrease the arousal of an 'influencee' during randomly assigned periods. The dependent variable was the electro-dermal response of the 'influencee'. Of course the 'influencee' was blind with respect to the influencing periods. A regression plot of the results of these studies versus their sequence number (the experimental years were not available in the publication) shows again a steady decline with a regression coefficient of -0.034 ($p = 0.043$).

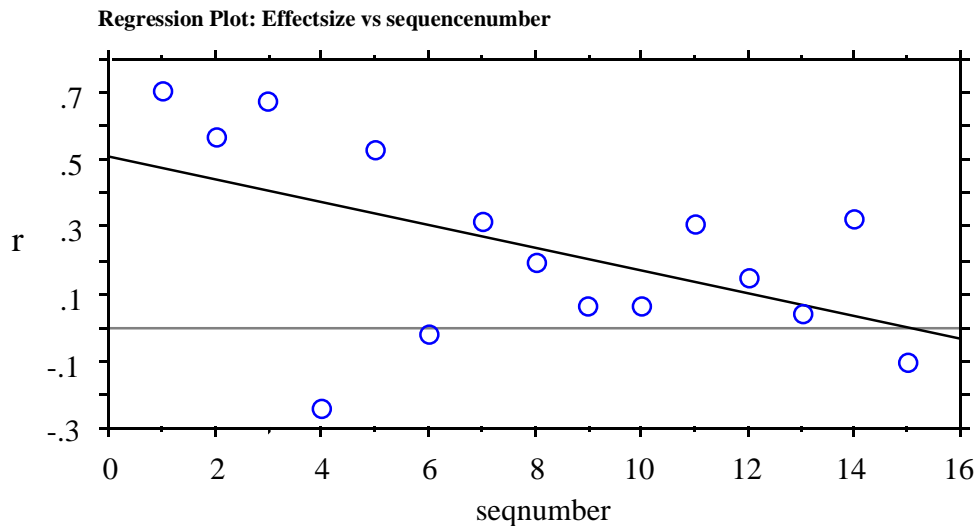


Figure 5: A regression plot of the effect sizes found in all studies on the effect of distant attention on arousal

Also the effects in experiments with other biological target systems (like orientation of fish, blood pressure and human ideomotor activity) with the exception of the mammal locomotion experiments show a tendency to decline over time though the number of experiments is too small to do a proper statistical evaluation.

Process oriented findings: Psi correlations

In Utt's review article it is claimed that not only the evidence for the existence of psi is overwhelming but also that some intrinsic correlations of psi scores with external independent variables have been replicated. As an example the correlation of performance in a psi task with the personality variable 'extraversion' is mentioned. This weak correlation of about 0.20 was originally established by examining the relevant free-response studies between 1945 and 1983 (Honorton, Ferrari & Bem, 1992) and was later 'confirmed' in the Auto Ganzfeld database where a slightly smaller correlation between extraversion and psi-score of 0.18 was found. However in the replication attempts of the Utrecht, Durham and Edinburgh group in the mid nineties the correlations declined even further from $r=0.15$ in the Utrecht replication via virtually nothing in the Durham replication to even a negative relationship in the Edinburgh replication. (Van Kampen et al, 1994; Broughton & Alexander, 1996; Morris et al, 1995)

One of the most cited correlations in the parapsychological literature is the correlation of performance in a psi task with the score on the Defense Mechanism Test (DMT). A close inspection of all the experiments which tried to assess this correlation reveals again a strong decline in the correlation coefficient (see figure 6, database from Haraldsson & Houtkooper, 1995). The regression coefficient is -0.034 ($p=0.0002$)

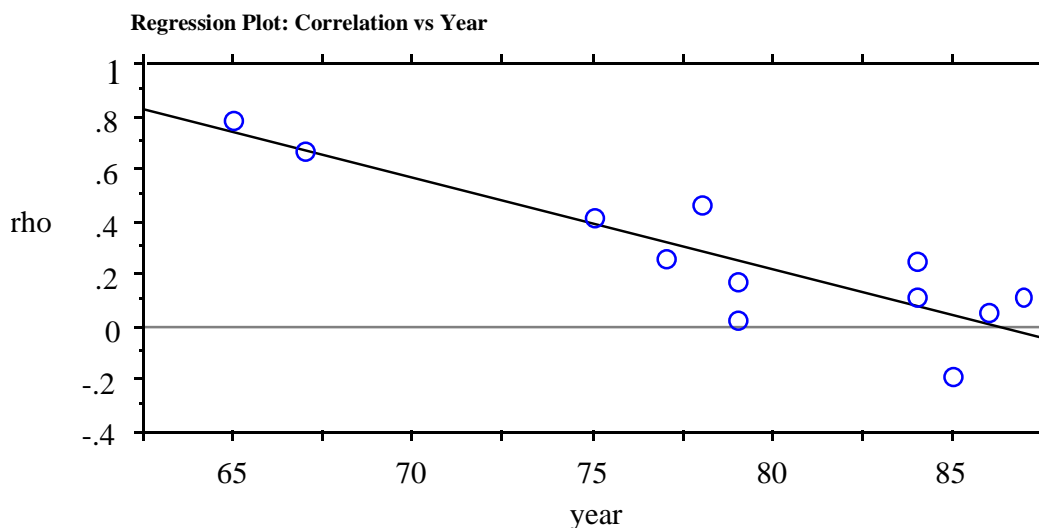


Figure 6: A regression plot of the correlations found in all studies exploring the relation between performance in a psi task and the DMT

Card guessing

So far we have discussed all databases from table I except the precognitive card-guessing database. Table I is generally taken by parapsychologists as the best evidence for the existence of genuine anomalies. However, besides the undeniable cumulative evidence for anomalies, we have also established in all cases significant sequential effects. Thus it seems that original effect sizes can not be replicated in the weak sense explained in footnote 1. In spite of my many efforts to trace the card-guessing database it appears that it has disappeared and it is just too large of a job to reconstruct the database from scratch. But, based on observations in the other databases, we would expect that this database too would exhibit decline effects. It appears that these sequential effects may be a rule in the study of psi phenomena. However, some exceptions have been claimed by individual researchers.

Paradigms with 'stable' effect sizes: Remote viewing

It has been claimed by May and others that remote viewing (clairvoyance) experiments do not exhibit a decline (Utts, 1996).

However this is only true for the remote viewing experiments that were run in or just outside of the framework of the (US government formerly) classified "Stargate" program. By contrast, the effect sizes in the remote viewing database of the Princeton PEAR group declined over years with a regression of -0.085 ($p=0.0006$; Spottiswoode, 1993). It may be that the difference in decline effects for the same paradigm, but between the two research labs, points us into the direction of an explanation.

Possible Explanations of Decline Effects

This section starts with two possible explanations that deny that there is any real anomaly to be accounted for. I don't share this opinion but leave it to the reader to decide.

Statistical artefact, selective publication

Regression to the mean has been proposed as an explanation for decline effects. Thus the first experiment is published because it found some significant effect by chance. Further replications will of course converge the result toward the true null-effect value. However, this explanation doesn't fit the data because the mean of the 'unselected' study 2 and further on, should show a mean null effect size. In almost all the examples given in the previous chapter, this is not the case. The analysis of across-investigator PK-RNG studies shows an incline after about 15-20 years of decline. This of course does not fit with a 'regression to the mean' hypothesis.

In one experimental parapsychological journal the official publication policy is that the research plan and design have to be submitted before the experiment starts, and that acceptance is on the basis of this design

rather than on the basis of the results. Thus selective reporting is prohibited. However it is uncertain how often this official policy is actually applied.

Starting around 1970, all parapsychological journals began to accept articles with null-results while before that date only significant studies tended to be published. If we look at the subset of the dice throwing studies that were done after 1970, for instance, we find that there is no over-all effect left. Thus, it could be argued that the change in publication policy artificially induced a decline. However, if we just focus on the dice experiments from 1970 to the present, we find that the regression coefficient is even higher (-0.219 for this subset of 17 studies while it was -0.053 for the whole data set; interestingly, when we take the subset before 1970 the regression coefficient is -0.097 ($p = 0.0045$)). This suggests that a break-down in the two subsets reveals different declines, the first one less strong, possibly due to selective effects of the publication policy in that period.

Increasing study quality

Another explanation of decline, like the previous one mostly put forward by sceptics, is that the earlier studies are of lower quality, resulting in artificially large effects. Later, with increasing quality the spurious effects would disappear. Several authors have argued (for instance Kennedy, 1938) that an increase in the experimental rigor of ESP research had resulted in a corresponding decline in ESP results, suggesting that extra chance ESP scores were due solely to loose experimental conditions.

This explanation, which indeed could apply to early card-guessing & dice experiments, does not apply to the automated psychokinesis experiments with RNG's as targets. The decline effects reported above for PK-RNG experiments are the weakest. Thus the quality hypothesis seems to fit the data at first sight. However, as already mentioned, recent replication attempts of the PEAR PK-RNG-studies using the same methods and hence the same quality as in the original PEAR studies, failed. Furthermore the remaining PK-RNG database after 1965 shows an incline rather than a decline.

Utts argued that in most meta-analyses no dependence of effect size versus study quality could be established. However, rather than supporting the idea that no correlation exists, it can also be argued that some relation might exist because the study quality score is determined using the written report rather than an assessment of quality based on a close inspection of the actual experiments. In low effect size studies the authors generally limit their report in such a way that a later meta-analyzer rates the quality lower than it actually was (e.g. authors tend not to report extensive randomness tests or safety measures when the outcome of the study was nonsignificant). Thus, the quality scores for low effect size studies may be higher than the figures used in the analyses cited by Utts. It can be seen that if one corrects for this potential reporting artifact, that this would result in a negative correlation between study quality and effect size.

From proof-oriented to process-oriented research

One of the 'explanations' put forward to account for declines is that during the last 20 years the focus on proof-oriented research has shifted toward process-oriented research. In the latter type of research one tries to derive underlying processes by either correlational analyses or manipulations of experimental variables. The latter could include manipulations deliberately introduced to obtain smaller effects. Thus the over-all effect size would be smaller. However, in practice most studies are either of the type using secondary correlational analyses or they involve manipulations intended to increase effect-size, not to decrease effect size. Actually this explanation cannot account easily for ANY of the declines described above, with the possible exception of some of the recent Ganzfeld experiments where in order to study the effect of the emotional content of target clips some extremely boring target-clips were used.

Psychological 'explanations'

Human performance in many types of tasks decline with time. Generally these declines can be attributed to a gradual change in physical or psychological variables that are relevant for the performance under study. Physically the body might become tired throughout a session; mentally, the subject might become bored with the task.

Thus within session/subject decline effects in psi tasks could easily be 'explained' by assuming that the subjects became bored with the task. However declines have also been reported in psi-tasks with fish (Braud, 1976). It may be doubted that mental fatigue is a good explanation for the decline from day 1 to day 2 of experiments involving fish, although similar declines had been found in conventional behavioral research using the same fish (Baenninger, 1966). Nevertheless, the explanation that within-session decline effects are due to fatigue seems the most natural one.

However between-experiment declines of effect sizes cannot be explained by fatigue on the part of the participating subjects because generally the subjects are different from experiment to experiment. If the replication attempts are conducted by the same experimenter it still seems possible to 'explain' them by assuming that the experimenter is getting bored and 'communicating' this non-verbally to the subjects. In

the parapsychological literature experimenter effects have been reported extensively. Not only of the subjects-influencing kind described by Rosenthal (Rosenthal, 1969) but also of a more subtle experimenter psi-influencing kind (Kennedy & Taddonio, 1976). Thus the attribution of between-experiment decline to the experimenters becoming more and more bored with the experiments is not as far-fetched as it seems. Note though, that for the classified remote viewing series that used the same experimenters and subjects over a period of more than 15 years there seems to be no decline. If ever boredom should have struck then it should have been within these series.

The third kind of declines, declines between-experiments and between-experimenters (between-labs), is most difficult to 'explain' by some gradually changing psychological variable. Of course the excitement of trying out something completely new is lost upon a replication, and the replicating experimenters are often novices with regard to the experiment that they try to replicate and their psychological set therefore does not differ much from the experimenter that did the original study.

Physical explanations

The striking finding by Spottiswoode of a correlation between cosmological time (LST) and performance in (free response) psi tasks has renewed interest in the possibility of some physical background variables playing a role in the psi process. Earlier, weak correlations with geomagnetic activity had been reported (Persinger & Krippner, 1989) but the LST-psi effect showed a dramatic 380% increase of the effect when the galaxy disappears behind the horizon (Spottiswoode, 1997)

Thus declines could possibly be 'explained' by some gradually changing physical variable. One argument against this explanation is that most physical (including cosmological, like sunspot frequency) variables have cyclic components with a period smaller than the period that parapsychological experimentation has been going on. Most of our databases do show a more or less steady decline but the PK-RNG database suggests that a cyclic component of about 30 years might be present providing some support for the influence of a cyclic background variable. If different psi phenomena are affected in the same (or in opposite) ways by the (cosmological) background variables one should be able to see some correlation between the long term sequential effects in different databases. The databases do not overlap enough to see if this is the case.

Von Lucadou's System theoretical 'Explanation'

In a provocative paper von Lucadou (1990) claims that for all complex information processing systems, including human consciousness, a quantum-like description is required. I.e. the formalism to describe these systems has basically the same structure as Quantum Physics. As a direct consequence these systems should, under specific conditions, also show some of the remarkable aspects of quantum systems. Most notably parts of the systems might show EPR-like correlations. Within the framework of quantum physics it has been shown that these 'non-local' correlations cannot be used to transfer bits of information. If the experiment is prepared in such a way that it is possible to do a measurement which could give information about specific parts of the system these non-local correlations will not occur. The actual potentiality of a measurement of classical information transfer using non-local correlations will destroy the 'carrier' of this information. So, according to von Lucadou, there is no way to construct a public experiment in which apparent psi information can be **used**. On the other hand by using a triple blind set up it must be possible to show that more correlations arise in the system than can arise by chance. However, specific correlations tend to decline and other correlations unexpectedly will pop up in the system. Thus the system theoretical framework of von Lucadou's offers a natural explanation of the rather elusive character of laboratory experiments while simultaneously not prohibiting strong anomalous correlations to occur spontaneously in the field.

Walker's Observational 'Explanation'

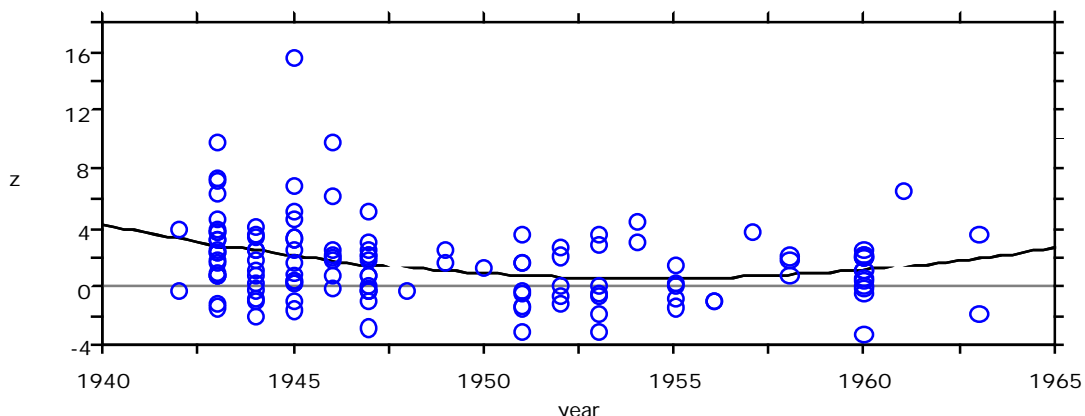
In the early seventies Walker (1975) extended Wigner's interpretation of the measurement problem. According to Wigner, the idea that measurement had a special status and therefore was responsible for the transition of a quantum system from its superposed state into a singular observed state (a.k.a. the collapse), was false (Wigner, 1967). He therefore speculated that it was the interaction with consciousness, or rather the interaction with some measuring device, that was responsible for the collapse of the state-vector. At that time, the measurement problem had become more popular due to the formulation of Bell's theorem a few years before and the growing body of evidence that physical hidden variables did not exist. In other words, it became a more popular idea that the collapse of the physical system was a real physical transition and not just a gain in knowledge.

Walker assumed that the observational process (i.e. the interaction of the physical system with consciousness) was not a unidirectional one but that in the process information could flow from the observer into the observed system. This information then could bias the distributions predicted by quantum

physical calculations and hence anomalous correlations could arise between the observed system and the mental (intentional) state of the observer. The theory predicted that it wouldn't matter if in psychokinesis experiments the data were first stored before the subject tried to influence those into a later specified target direction (so called PK on prerecorded targets). This rather contra-intuitive prediction was then confirmed. The theory also unified all psi phenomena by including the random processes in the subject's brain as a possible target for 'retro-active' PK. Upon feedback (or observation) brain states could retro-actively be biased so that they would 'correlate' to observed feedback. A fundamental problem in the theory was how to describe the process of multiple observation. Several flavors of the Observational theories have been developed that deal with that problem (Houtkooper, 1977, 1987; Schmidt, 1975, 1978). In a number of those, further observation of the data still has an 'influence' on the observed outcome. Thus, it was not only the subjects in an experiment who were relevant for the outcome but also experimenters, people who later analyzed or checked the data and even readers of the final article. The so-called "checker effect" had already been experimentally established (Weiner & Zingrone, 1989), but remained a mystery within the already mysterious world of the paranormal until Walker's formulation of the observational theory.

The fact that uncontrolled multiple and future observers of an experimental outcome might have an effect on the outcome could also 'explain' the difficulty to replicate these effects because it is obvious that after an initial success the number of potential future observers for a second experiment increases dramatically. One even could argue that such an effect would be strongest after a meta-analysis or another widely cited publication. Houtkooper (Houtkooper, ????) argued that there is ample evidence for such a special strong effect after publications of meta-analyses. He labeled this the MAD (Meta-analysis Destruction) effect. The fact that the Durham and Amsterdam Ganzfeld series declined so strongly between 1995 and 1996 might also be interpreted in this light. The 1995 outcomes of the experiments were used by Utts in 1996 in a widely published paper dealing with the evaluation of the 'until then' classified StarGate psi research program. It should be noted however that if one would like to interpret this decline in this way one should adhere to the so-called potential observer theory (Hartwell,1977; Millar & Hartwell, 1978) in which not the actual observers of data are tied to the outcome but the potentiality of becoming an observer is the important factor.

In an article specifically dealing with the (non) replicability issue, Walker (1983) wrote: "...The time has passed when we may reasonably expect to accidentally discover some simple psychological test that will separate out individuals who can assure repeatability, or when we may, without understanding of the underlying physical principles, prescribe experimental procedures to circumvent the current difficulties...." He then continues to give a formal treatment of models that include the effect of 'other observers' but assumes a constant experimental success rate rather than one in which the success rate is dependent on a factor related to the number of (potential) observers of the outcome and their world view. If one corrects his equations accordingly the solution shows an oscillatory behaviour for the success rate. We have found one clear indication for such a cyclic behaviour in the PK-RNG data (see fig. 4). The only database that covers a similarly long period is the PK-Dice database. Visual inspection of figure 1 suggests that maybe in that database we have a full cycle of decline, incline followed by the start of another decline. If we leave out the few data points after 1965 which might be related to the start of the next decline phase the resulting datapoints show a sequential effect quite comparable to the PK-RNG sequential effect.

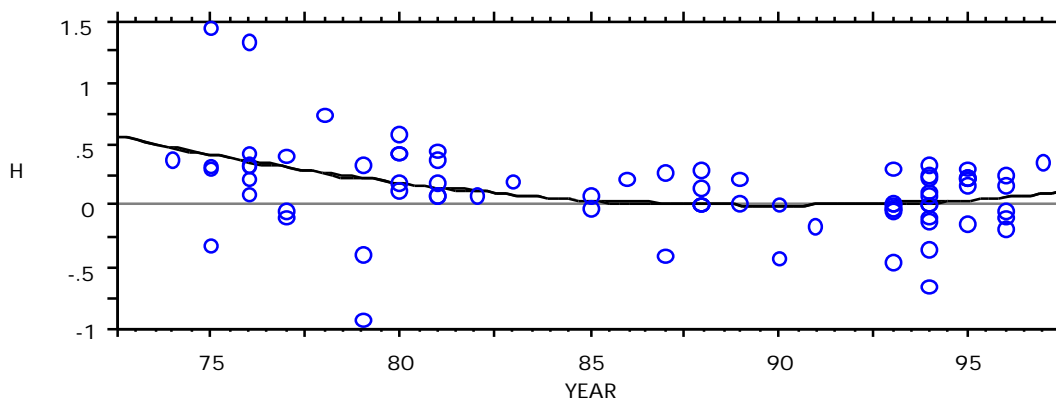


$$Y = 67862.69 - 69.449 * X + .018 * X^2; R^2 = .089$$

Figure 7: The Dice throwing experiments analyzed for a non-linear effect between 1940 and 1965

Both regression coefficients are significant at the $p < 0.025$ level.

A similar analysis for the Ganzfeld data including the more recent studies and excluding 2 Ganzfeld studies that were completely atypical because auditory targets rather than visual targets were used, results in a similar plot, suggestive of a recovery from the decline after a period of about 17 years (Fig.8).



$$Y = 15.148 - .337 * X + .002 * X^2; R^2 = .154$$

Figure 8: The Ganzfeld experiments analyzed for a non-linear effect between 1972 and present.

Both regression coefficients are marginally significant.

A new position about reality

Anomalous phenomena like telepathy, psychokinesis and precognition have long been seen as 'just an extension' of people's normal sensory capabilities. Hence labels like Extra Sensory Perception (ESP) and more recently 'anomalous cognition'. In this view the phenomena are discussed in terms of information transfer, bandwidth and signal to noise ratio. The low signal to noise ratio is then seen as the culprit of the difficulties to replicate and the research can be characterized as a kind of engineering approach to the properties of the information transfer channel. Such an approach does not (at first sight) require a fundamental change of world view. Even the apparent retrocausal aspects of some (if not all) of the phenomena are not necessarily incompatible with a mechanistic world view, as Newtonian physics is basically time-symmetric.

However, while explanations like those mentioned above cannot be ruled out, I speculate that the difficulty to replicate is not just due to a low signal-to-noise ratio but has a deeper reason, a reason that suggested by the theoretical approaches by von Lucadou and Walker. Both theoretical frameworks account for psi phenomena by either using directly or metaphorically non-local EPR-like correlations.

Von Lucadou argues that no signal can be transferred using these correlations and any situation where such a signal *could* be encoded in the psi-induced correlations would tend to destroy the correlations itself. This logically explains why in exploratory situations these correlations seem to appear easily while in confirmatory replication attempts (where in principle given the knowledge from the first experiment one could use some scheme to encode a signal) it is difficult to replicate. That some correlations are still found in some of the later replication attempts could be explained by the fact that in order for a parapsychological experiment to be valid the information that supposedly is going to be 'transferred' has to be randomized. Thus the experimenter generally is not 'free' to select this information so that we cannot really speak of intended information transfer but merely of correlations between target information and behavior of the percipient. In von Lucadou's terms there is only a signal if meaningful information is transferred, i.e. a signal that can result in some meaningful action. If a system is closed in such a way that these actions have no effect upon the outside world then it still would be possible to have anomalous correlations within the system. Von Lucadou used the concept of organizational closure from the domain of self-organizing systems to characterize systems that have this property of having no meaningful action on the outside world.

Thus, according to both von Lucadou and Walker, some closed-system correlations between mental states and other mental or physical states which are EPR-like could be real. However, when one tries to move these phenomena into the realm of the purely objective, i.e. make them sharable by everybody, they disappear, the local reality disappears.

I have called this local reality 'tribal' reality (Bierman, 1999), because it nicely accounts for the disappearance of the power of the shamans. Originally, mental states of the shaman could indeed correlate in an EPR like way with mental or bodily states of clients. However when the tribe was broken open by *Coca Cola* and *MacDonald's* the 'powers' of the shaman got lost and objective science was unable to establish any reality to the stories from the past. The parapsychological research community can also be seen as a kind of a tribe in this respect.

Conclusion

The account given here for the difficulty to replicate paranormal phenomena requires a radical change in world view. We cannot expect science, or society, to change a time-honored (at least in the West) world view on the basis of data that appear to be extremely difficult to replicate. Is there a way out of this vicious circle?

According to Walker, a slowly changing world view would lead to gradually easier replications so that psi phenomena in the end would become part of the objective world (Walker, 1983).

However there might be another option which has been overlooked. If the replication problems are due to the fact that any situation in which signals could be deduced from them prohibit the correlations from occurring, then we should realize that we are talking about **classical** bits of information. There is no restriction on the transfer of quantum bits. Thus one could try to set up experimental systems that are a macroscopic analogy of quantum teleportation systems.

How should we conceptualize what constitutes a macroscopic analogy of a quantum bit? According to Hameroff & Penrose (ref) non-conscious states are superposed (non-collapsed) states and therefore 'contain' quantum bits. This suggests that correlations between non-conscious states of two subjects at a distance in time and space might become EPR-like correlated, especially if there has been a strong interaction between the two subjects in the past. This would argue for experiments with twins. Of course measurement of non-conscious states is difficult but it could be done using psychophysiological variables or brain imaging techniques. In fact some experiments have been reported in the parapsychological literature that come close to the experimental approach suggested here, and they were extremely successful, but not part of a theoretically driven research program and hence they had no follow up (Grinberg-Zylberbaum et al, 1992). It remains to be seen if these correlations could be used (i.e. result in some meaningful action) or if the correlations will disappear if one tries to "use" the system in some way.

A quote from the star subject in the US Governmental StarGate program, Joseph W. McMoneagle:

.....I've observed that elusiveness more times than I care to count. I agree that it would seem to be an element itself that requires study. There is always a tendency to search for order in the stream, when it may be a function of disorder being observed. Or perhaps a relaxation in the expectation for order.....

References

- Baenninger, R. (1966). Waning of Aggressive motivation in *Betta Splendens*. *Psychonomic Science*, **5**, 207-208.
- Bem, D.J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin*, **115**, 4-18.
- Bierman, D.J. (1995). The Amsterdam Ganzfeld series III & IV: Target Clip emotionality, effect sizes and openness. *Proceedings of the 38th Annual P.A. Convention*, 27-37.
- Bierman, D.J. (1999) Is non-replication a fundamental characteristic of psi? *Proceedings of the 42nd Annual P.A. Convention*, 417.
- Braud, W.G., (1976). Psychokinesis in aggressive and non aggressive fish with mirror presentation feedback for hits: Some preliminary experiments. *Journal-of-Parapsychology*, **40 -4**, 296-307.
- Braud, W.G. & Schlitz, M.J. (1991). Consciousness Interactions with Remote Biological Systems: Anomalous Intentionality Effects. *Subtle Energies*, **2-1**, 1-46.
- Broughton, R. & Alexander, C. (1995). AutoGanzfeld II: the first 100 sessions. *Proceedings of the 38th Annual P.A. Convention*, 53-61.
- Broughton, R. & Alexander, C. (1996). AutoGanzfeld II: An attempted replication of the PRL Ganzfeld research. *Proceedings of the 39th Annual P.A. Convention*, 45-56.
- Dennett, D.C. (1991). *Consciousness Explained*. *Penquin Books*, London. ISBN 0-7139-9037-6.
- Gordon, I.E., (1996). Empiricism: perception as a constructive process. In: *Theories of Visual Perception*, 154-179. John Wiley & Sons, Chisester.
- Grinberg-Zylberbaum, J.; Delaflor, M., Sanchez Arellano, M.E., Guevara, M.A., & Perez, M. (1992). Human Communication and the Electrophysiological Activity of the Brain. *Subtle Energies*, **3 -3**, 25-43

- Hameroff, S.R., and Penrose, R. (1996) Conscious events as orchestrated spacetime selections. *Journal of Consciousness Studies*, **3-1**, 36-53.
- Haraldsson, E. & Houtkooper, J.M. (1995). Meta-analysis of Ten Experiments on Perceptual Defensiveness and ESP: ESP Scoring Patterns, Experimenter and Decline Effects. *Journal of Parapsychology*, **59**, 251-271.
- Hartwell, J.W. (1977). A bound for the observational theories of psi. *European Journal of Parapsychology*, **2-1**, 19-28.
- Honorton, C. (1985). Meta-analysis of psi ganzfeld research: A response to Hyman. *Journal of Parapsychology*, **49**, 51-91.
- Honorton, C., Ferrari, D. C., & Bem, D. J. (1992). Extraversion and ESP performance: Meta-analysis and a new confirmation. In L. A. Henkel & G. R. Schmeidler (Eds.), *Research in parapsychology 1990* (pp. 35-38). Metuchen, NJ: Scarecrow Press.
- Houtkooper, J. M. (1977). A comment on Schmidt's mathematical model of psi. *European Journal of Parapsychology*, **2 (1)**, 15-18
- Houtkooper, (1983) Observational Theory: A research programme for Paranormal phenomena. Thesis University Utrecht, dec. 1983, 57-64.
- Hyman, R. & Honorton, C. (1986). Joint communiqué: the psi ganzfeld controversy. *Journal of Parapsychology*, **50**, 351-364.
- James, W. (1890). cit. Gordon, 1996.
- Van Kampen, D., Bierman, D.J. & Wezelman, R. (1994). Personality and psi: Unravelling relations between extraversion, agreeableness and openness to experience with Ganzfeld performance. In: D.J. Bierman (Ed.) *Proceedings of the 37th PA Convention*, University of Amsterdam, 175-181.
- Kennedy, J.L. (1938). The visual cues from the backs of ESP cards, *Journal of Psychology*, **6**, 149-153.
- Kennedy, J.E. & Taddonio, J.L. (1976). Experimenter effects in parapsychological research. *Journal of Parapsychology*, **40**, 1-33.
- Loftus, E.F. (1997) Creating False memories, *Scientific American*, September 1997, **277-3**, 70-75.
- Millar, B. & Hartwell, J. (1978). Dealing with divergence. In W. C. Roll (Ed.), *Research in parapsychology, 1978*. Metuchen, NJ.: Scarecrow Press, 1979.
- Milton, J. & Wiseman, R. (1999). Does psi exist? Lack of replication of an anomalous process of information transfer. *Psychological Bulletin*, **125**, 387-391.
- Morris, R.L., Dalton, K., Delanoy, D.L. & Watt, C. (1996). Comparison of sender/no sender condition in the Ganzfeld. *Proceedings of the 38th Annual P.A. Convention*, 244-259.
- Persinger, M.A. and Krippner, S. (1989). Dream ESP experiments and geomagnetic activity. *Journal of the American SPR*, **83**, 101-116.
- Radin, D.I. & Ferrari, D.C. (1991). Evidence of consciousness on the fall of dice: a meta-analysis. *Journal of Scientific Exploration*, **5**, 61-83
- Radin, D.I. & Nelson, R.D. (1989). Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics*, **19**, 1499-1514.
- Rhine, .B. (1969). Position Effects in Psi Test Results, *Journal of Parapsychology*, **33**, 136-157.
- Rosenthal, R. (1969). Interpersonal Expectancies: The effects of the experimenter's hypothesis. In: R. Rosenthal & R. Rosnow (Eds.) *Artifacts in Behavioral Research*, Academic Press, New York.
- Spottiswoode, S.J.P. (1993) Personal communication on (SMTP/TCP) e-mail (message_id's: <9304160128.AA19945@jsasoc.com> of april 16 1993 and <9304132041.AA16375@jsasoc.com> of april 13 1993)
- Spottiswoode, J. (1997). Geomagnetic fluctuations and Free Response Anomalous Cognition: A new Understanding. *Journal of Parapsychology*, **61-1**, 3-12.
- Utts, J. (1991). Replication and Meta-Analysis in Parapsychology. *Statistical Science*, **6-4**, 363-403.
- Utts, J. (1996) An assessment of the evidence for psychic functioning, *Journal of Scientific Exploration*, **10-1**, 3-30.
- von Lucadou, W. (1990). Some remarks on the problem of repeatability of experiments dealing with complex systems, *NATO Advanced Study Institute, Information Dynamics*, Irsee, FRG, June 15-26, 1990
- Walker, E.H. (1975). Foundations of parapsychological and parapsychological phenomena. In Laura Oteri (Ed.), *Quantum physics and parapsychology*. New York: Parapsychology Foundation.
- Walker, E.H. (1983). The repeatability problem in Parapsychology. In: Betty Shapin and Lisette Cly (Eds), *Procs. Of an International Conference held in San Antonio, Texas, oct. 28-29*, Parapsychology Foundation, New York, N.Y., 98-143.

- Weiner, D.H. & Zingrone, N.L. (1989). In the eye of the beholder: further research on the "checker effect". *Journal of Parapsychology*, **53**, 203-233.
- Schmidt, H. (1975). Toward a mathematical theory of psi. *Journal of the American Society for Psychical Research*, **69**, 301-319.
- Schmidt, H. (1978) A remark on the "divergence problem". *European Journal of Parapsychology*, **2**, 163-166.
- Wigner, E.P. (1967). *Symmetries and Reflections*. Bloomington: Indiana University Press.
- Wezelman, R. & Bierman, D.J. (1997). Process Oriented Ganzfeld Research in Amsterdam: Series IVb, emotionality of target material, series V and series VI: judging procedure & altered states of consciousness, *Proceedings of the 40th P.A. Convention*, 477-491.