English Project 1995-96

Scientific Experimentation on Extra-Sensory Perception

Stéphane Couturier
3rd year Sup'Aéro student

"It is a poor science that dictates conditions to Nature. It is a better one that follows up with it's well adapted controls and conditions."

J.B.Rhine

Scientific experimentation on Extra-sensory perception

Abstract

The objective of this report is to gain a clear picture of how science has apprehended the slippery question of so-called paranormal phenomena, and then to determine the present state of this field of research.
After first defining the exact domain of investigation, a special importance will be attributed to experimental setting, since the validity of the results hinges on a close examination of the conditions under which observations are made. Therefore 3 parameters will be proposed to characterise the relevancy of a psi experiment:
The *true statistics* index, by which one can evaluate an experiment is really described by it's mathematical model,
the *controlling* aspect, which determines whether there might be some element of sensory perception involved,
the *user-friendly* factor, which measures how well an experiment has eliminated unnecessary or detrimental elements.
They will be referred to throughout the evolution of experimentation, along with modifications resulting from the constant criticism on the part of the sceptics. Hence, from the diverse set of experiments by the innovative researcher J.B. Rhine in the early thirties, with Zener cards, this will lead us to the computerised setting involving Random Event Generators in the sixties, and finally to the latest style, the Ganzfeld laboratory conceived by Charles Honorton, which is still running at the moment.
As psi researchers vary their testing conditions, the interpretation of the researcher and of the sceptical will be laid down, and a debate will emerge, eventually leading to a major scientific request for experimentation: the idea of *reproducibility*; the mere question of our capability of setting a repeatable, reliable experiment, for which a positive result would definitely prove psi ability, is finally at stake.

Table of contents

Introduction

Do you believe in paranormal phenomena?

This question, I am sure, arouses immediately a controversy in your mind between instinctive suspicion and some sort of thirst for knowing more about human nature... Now a more down to earth approach to the question suggests that it's terms are surely too wide and vague to be answered outright. In fact, something can easily be noticed in how the existence of such phenomena is traditionally perceived by the layman: on the one hand, there is widespread discredit from people of "scientific" education and on the other hand, public polls keep registering an ever higher percentage of belief in the existence of "scientifically unexplained" phenomena.

And yet, we have to know that for a little more than one century now, communities of scientists have been trying to tackle the existence of such phenomena, with the same reductionist methods as in other scientific disciplines. I believe that, as in so many other subjects, people are not well enough informed to really understand the issues involved. This is why I intend to re-examine some of the pioneering work in this field, in order, if not to answer the question, at least have a clearer idea of where the real questions lie.

PART I: A Methodology for testing Extra-Sensory Perception (ESP)

1) What is Psi, what is ESP?

In most forms of research, new findings can usually be described as modifications of existing terms, but in psychical research, there are drawbacks to this procedure. Some traditional names have unfortunate associations with fraud and superstition; the word clairvoyant, for instance, is liable to stand for fake fortune-tellers at village fairs. Existing terms, moreover, may carry implications about the nature of the findings which may be incorrect. In consequence researchers tend to seek neutral and impersonal names for psychic phenomena. It's first label, *'paranormal cognition'*, which still remains in the popular vocabulary, has found it difficult to sustain an unbiased and scientifically respectable image. Indeed, the implications of the word paranormal are misleading, for there is no reason to suppose that there is anything beyond the normal, in other words, outside the laws of nature, as the prefix 'para' (from the Greek, means "in the proximity of") would suggest.

Psi, as it has finally been called by the scientific community, rather than paranormal, is a term covering the whole family of phenomena and experiences, real or alleged, for which no physical cause has yet been discovered, and which seem to contradict established theories.

Extra-Sensory Perception, or perception without the function of the recognised senses, is a term that has been introduced by J.B. Rhine (see his work later). It splits into perception of the thought or feeling of another (telepathy), and perception of an objective fact or relation (clairvoyance), both without the aid of the known sensory processes. The third ability under ESP that won't be treated here despite the numerous experiments that were dedicated to detect it, is precognition, or the ability to predict an event in the future without any induction from the information acquired by senses.

Three classical examples to illustrate each phenomenon -if it does exist:
I am sitting on a chair, mute and motionless, focusing my mind on the random image of a white horse with a red saddle. My brother, who was supposed to guess what was in my mind, rushes in the room five minutes later, and hands me the drawing of a white horse with a red saddle. He also gives me a sealed, opaque letter, that none of us are supposed to open until the next day. I keep the letter on myself, without looking at it. The day after, we start the same task again, and this time, I have randomly chosen to focus my mind on the image of a house without a roof. My brother comes, and as scheduled, I open the sealed letter; the letter contains the image of a house without a roof. Now the day after, I receive my weekly newspaper still wrapped in plastic. I call my brother up, and I ask him to tell me what is the first sentence beginning after the third line of the fourth article in my newspaper. He writes his answer down, I read it loud, and realise after unwrapping the newspaper, that his answer is correct.
Before I continue, let us calm the audience and specify that this was pure fiction. But it helps visualise the three alleged phenomena that come under the denomination of ESP; the first day, my brother may have shown a *telepathic* ability, the second day, *precognition* from the first day may have been proven, and the third day, he may have experienced *clairvoyance*. Most people would agree - if they trust me on the story...-

that my brother has indeed these 'extra-sensory' abilities. Now why did I say 'may' rather than 'would'? well, because one would have to describe more carefully the circumstances of the experiment to be 100% sure that my brother is not using his senses, fooling me, or simply trying his luck on good hypotheses: Suppose the first day, he tried his luck on an image referring to an obsessive dream that I used to speak out loud while sleeping in my young childhood, a fact that he never told me. This image could have popped up naturally if I was waiting for a "random" image to initiate the experiment. One can imagine other complicated but plausible flukes to explain "rationally" the other amazing results, involving for instance leakage of information from my part to my best friend, or subtle pick-pocket talents from my brother,... Without falling into paranoia, one can feel that testing ESP needs a methodology, a means of evaluating results, and also (remember I said: "if people trust me on the story") a guarantee against flaws from the experimenter himself...

2) The Society for Psychical Research; the first group of scientific investigators

In the wake of the devastating wave of spiritualism that had inflamed the XIXth century, groups of scientists finally decided to build the first structure that would try to rationalise their approach to the following claimed gifts or phenomena: telepathy, clairvoyance, hypnotism and mesmerism, apparitions and haunting, as well as "the diverse unexplained phenomena usually called spiritualism". The Society for Psychical Research (SPR) was founded in 1882 in London by a group of eminent personalities at that time, among whom were the physicists William Barrett, William Crookes and Oliver Lodge, and philosophers Frederic W. H. Myers and Edmund Gurney. Very soon after, in 1885, the American Society for Psychical Research was founded by the Harvard Professor in psychology William James and other psychologists. Their aim was more specifically to study manifestations through a medium.

There were plenty of respected scientists in both of these institutes, and their desire was not to prove the authenticity of spiritualism, but to study and explain the Psi phenomena. Unfortunately, most researchers were only amateurs in the field and, despite their competence in their respective disciplines, it was notably hard to distinguish between what could be fraud and what could be actual phenomena.

3) Importance of experimental setting: a probable example of fraud

In 1882, the British SPR tried to tackle a case of telepathy in Brighton. It involved the journalist Douglas Blackburn and a hypnotist, 19 years old, named G.A. Smith. They were invited to London and submitted to a series of tests, where Smith, the "receiver", eyes folded and ears blocked up, was totally wrapped in several blankets. Blackburn, the "emitter", was going back and forth, freely in the room, while concentrating to mentally transmit a drawing, a figure or any other image that he had been given. The results, once published in the society's journal, were stunning; and if sometimes there happened to be no resemblance between the drawings, most of the time the responses were so close to the model that the experimenters concluded a probable case of telepathy.

Around 30 years later, Blackburn revealed in two articles entitled "confession of a famous medium" and "confession of a telepathist", that he could exchange signals with his partner and was astounded that the SPR staff could not accurately evaluate the facts

of the experiment; he claims: *"describing one of our experiments, they stated that I had not touched him once, where in fact I touched him eight times, which was how our code worked at that time. I have to say that we were not aware of the seriousness of the scientific approach followed by these respectable people who were studying our case. We thought they were some other variety of pseudo-scientist spiritualists of the time, and we wanted to show how they were incompetent."* At another session of the same series, he was staying at a distance from Blackburn, and openly recopying the drawing several times, allegedly for better memorisation. But in the meanwhile, he said he had secretly been able to recopy it onto a cigarette paper. He claims: *"At that time, I was a talented prestidigitator, and it was easy for me, while moving in the room, to put the cigarette paper into the cap of the pen I was using. I then did the signal that indicated that I was ready: my foot under the carpet made me stumble, and the moment thereafter, he shouted: "I got it!", he dodged his hand in and out of the blanket and asked as scheduled: "Where is my pen?" Straight away I put mine down on the table. He took it and the room plunged into a tense silence"* In fact, Smith was taking advantage of a pause and stealthily glancing down at the collected drawing, by the faint light of a phosphorescent stone, previously concealed in the blanket. Five minutes later indeed, he took off the blindfold, removed the blankets and showed a drawing very similar to the original.

Now this magnificent hoax is only Blackburn's version of the facts, which he revealed after the death of the experimenters who thus could not respond to the criticism; what is funny is that Blackburn thought his "colleague" Mr Smith was dead too, but he was not; and he severely responded to these "confessions" in 1911 that the experiment had been conducted seriously, and that it's conclusions were sensible.

One can easily infer from this that **fraud from the percipient** will naturally be a major hindrance for studying psychic phenomena. This more or less reliable testimony also highlights the importance of clearness in an experimental report. At this point, there was an implicit appeal to who could develop a solid basic methodology.

4) Towards the simplest methodology

In the twenties, most of the American universities were very reluctant to conduct research into "Psi", still so much related to the fraudulent image of spiritualism, but some of them were receiving funds from the public who were still interested in spiritualism. The only serious studies conducted in the field did not lead to any positive results. As an example, J.E. Coover, a researcher in Psi working at Stanford University, conducted a series of delicate experiments, where the percipients were supposed to guess the number and the colour of the playing cards that another person was holding. He concluded that no case of telepathy had been detected. At Harvard, L.T. Troland conceived a complicated apparatus to replace the cards, and obtained negative results as well, though it might have been that his actual interest was more to build the device itself than to apply it to Psi research. Both of them gave up their work in this field.

At that time, William McDougall, a British doctor and psychologist, had been promoted as Dean of the Department of Psychology at Harvard. He had acquired sufficient renown in the scientific community to choose his own field of research. He was openly interested in the paranormal and it was actually he himself who changed this term into 'parapsychology', borrowed from the German language. McDougall used the funds that his department was granted, to train a team of scientists who eventually

became quite famous. One of them tried, with little success, to send mental images to a psychologist in Paris; another one, in Cambridge (Massachusetts), conducted telepathic experiments between subjects standing in different rooms.

But among his disciples, the one to be retained was certainly the young researcher, Joseph Banks Rhine; he would soon become one of the respected leaders in Psi experimentation. As, at this point, the reader should feel legitimately suspicious, and as the problem of the reliability of the experimenter in Psi experiments arises, it would appear relevant here to provide elements (as neutral as possible) of the biography of Mr Rhine, in order to be more familiar with the character. Born in 1895, he came from a village in Pennsylvania. At a very young age, during the first world war, he volunteered to join the Marines as a rifleman; on his return, he married a young teacher, named Maria Weckesser, and they both decided to go to the University of Chicago to study botany. Rhine was said to be an irascible and meticulous man. Afterwards, he came to wonder if the scientific methodology that he and his wife were using in their study of plants could be applied to other domains, that so far had been considered mere beliefs. He then felt notably encouraged in reading McDougall's book "Body and Mind", which intended to show that psychical research was necessary to have a global view of human nature. Rhine concluded that it would be *"unforgivable for the scientific community to ignore the testimonies reported all over the world, related to the paranormal - if ever it really existed"*.

In 1926, Rhine gave up botany, and went to Boston with Louisa, in order to join the branch of the American Society for Psychical Research there and start working. After being confronted with a cheating medium in the spiritualist circles, he decided that the only way to test psychic faculties that were puzzling him so much, was controlled experimentation, that could ensure that the same causes were leading to the same results.

In 1927, McDougall created a Department of Psychology at the Duke University of Durham (North Carolina) and asked his young colleague, Rhine, to help him out in his research.

In 1930, Rhine had initiated proceedings to start a new concept of research:
He began to make a series of experiments on subjects, whom he asked to guess the order of the cards in a hand that had been previously shuffled. If the rate of success was clearly over the rate of chance, one could reasonably infer that an unknown factor was at stake; hence, the analysis of results would prove to be an easy task, compared to experiments involving fact-guessing or object-drawing, since cards apparently have a direct link with figures, and thus with **true statistics**. On top of that, it presumably offered easy **control** as well. Finally, the method made the setting quite "**user-friendly**" for both the percipient and the experimenter. He realised very soon that it would be important to steadily improve each of the three interrelated factors mentioned above, since throughout his and his successors' careers, they seemed to govern the quality of an experiment.

5) Three factors to characterise the quality of an experiment

Let the ***true statistics*** factor be defined as the ability of the experiment to approach an ideal probabilistic law that would govern it's outcome without error. Indeed in an experiment, the relationship between the analysis of results and mere probabilities has first to be established, which simply cannot be the case in most 'free choice' (refer to part II, Appendix 2: Psi vocabulary) experiments, where the odds are judged "very unlikely" but cannot be evaluated - this is the case in the simplistic

examples of chapter I.1. Once the probability is calculated, though, one can use statistical techniques to test if the experiment has been significant or not; more details are supplied in part II, Appendix 1. The technique Rhine used by the way, has been validated by the American Institute of Mathematical Statistics. But the correlation between analysis and statistics can then be disturbed by elements that may influence the percipient's mind and thus modify the actual "chance rate" expected, as Rhine could sense in his early experiments. Finally, the tough problem of how to generate random figures will eventually emerge, and the technical tool used to solve that problem will effectively rise up this factor by a great amount...

_The factor of ***control*** is by definition the ability of the experimental setting to hinder possibilities of fraud that would ruin the reliability of the results. Attention will be drawn, on top of deliberate deception, to the more subtle unconscious sources of fraud. Secondly, it is interesting to note that the notion of control, initially conceived to protect an experiment from the percipient, also aims at avoiding as much as possible fraud from the team of experimenters themselves. As control improves along decades, a major concern will eventually arise among the main teams of researchers: to find a setting with efficient control, without the risk of inhibiting any possible Psi ability.

_The ***user-friendly*** factor grossly refers to simplicity in the conception of the setting. It will first have an impact on the feasibility of an experiment, considering the high number of trials required for tangible statistics: it will determine how often the test can be reproduced in a given period, for example (reproducibility in time). It is quite important to keep the testing time short, since the guinea pig happens to be a person who has other obligations in life than to serve the experiment, unlike the seed in a sterilised Plexiglas box which is used for a biology experiment. The *user-friendly* factor is also a guarantee of the reproducibility - in space - of an experiment, which is so important for Psi studies to gain the legitimacy of being a science (see part III). Finally, as the results of a large series of trials proved to fall down at some point, the experimenters claimed that the setting itself could influence the percipient's set of mind, his psychological disposition, and thus the alleged Psi abilities. To these respects, it will be noticed how *user-friendly* experiments, in the sense of psychologically incentive, favour positive results; an analysis which presumably sounds reasonable admitting that if such ability exists, it should be somehow related to the "states of mind" reported in existing psychological theories.

PART II: Rhine's pioneering work (1930-33)

1) Choosing a test with quantifiable results

At first, like his predecessors, Rhine used ordinary playing cards. However, after some trials that brought negative results, he considered that the 52 cards were a set that might favour unconscious habits of choice, or superstitious manners - for instance always mentioning preferred cards and avoiding those that bring bad luck. In other words he thought the *true statistics* index might be bad, which could interfere with the phenomenon he was seeking if it was there, and thus he tried to improve it.

This is the reason why he asked one of his colleagues, Karl Zener, a psychologist specialised in tests of perception, to design a series of 5 cards, each yielding a clear and non confusing drawing. Zener proposed him a series containing a plus sign, a circle, a square, a star, and a curvy line. Rhine named them Zener cards. The basic experiment was highly simple. In a hand of 25 cards containing 5 cards of each symbol, the odds were that the subject was supposed to guess an average of 5 cards properly. It would be a statistical law, which means that the percipient could have 8 hits out of 25 cards, but then over several packs the number of hits would go down to average 5 - unless of course, some other factor came up.

Now as said earlier, Rhine and his team used a statistical technique to analyse all the data they obtained over three years, that represents a little more than 90,000 trials. In order to give consistency to his study, he kept piling up this data over these three years without publishing anything; nobody could therefore reproach him of selecting the good data and thus distorting the statistics. This would indeed be an important flaw, since the nature of the results themselves is statistical. This technique is scanned in Appendix 1, which describes the relevant quantities characterising the results of an experiment. To put it in a nutshell, the final variable X represents a normalised deviation from the mean chance expectancy, which implies that the higher X is, the further off one is from expected chance. Actually, the way X is normalised is that for example, if in an experiment, X=4, the odds against the hypothesis of mere chance are 142 to 1; if X=5, these odds advance to 1,300 to 1; on the other hand if X=1, the odds are 1 to 1 (they are even); if X=0, you are right on the chance expectation figure so odds are very high for chance. With this in hand, you can have a look at the two quantitative tables on Rhine's experiments; the first one displays specific results on the subject Linzmayer, and the second one shows a wider range of results from the subject Pearce.

Table 1

Undifferentiated ESP or Pure Clairvoyance, statistics on Linzmayer

conditions	n° of trials	n° of hits	value of X	Average /25
undifferentiated ESP	360	143	13.9	9.9
Pure Clairvoyance	240	95	11.2	9.9

2) Investigating the experimental setting with Linzmayer as subject

It sounds interesting to enter the core of his work now, since it will give a perspective on his strategy of experimentation. Starting from the beginning of the survey, Rhine and several colleagues were continuously looking for a particularly gifted subject, since their interest was to gather the highest amount of positive results in their stock of experiments. This is not an experimental flaw by the way, as long as they keep, for a given subject, all the data they get from each experiment they choose to conduct in lab. The factor they can legitimately play with, is the change of percipient after an experiment.

2.a) An unconscious source of flaws

After more than a year of seeking in vain, in Spring 1931, a student in economics at Duke University gave a high average in the first sets of trials. His name was Adam J. Linzmayer, and he was an undergraduate student when he began to work with Rhine's team. The second preliminary series he was given by Rhine, produced the very high score of 21 hits in 45 trials. Within these, he got 9 correct calls successively! Let's have a look at the reported experimental conditions:

The pack of cards was systematically shuffled and cut before each trial, and the work was done at a distance, behind screens, with cards left in the unbroken pack until it was called down (this is actually said to be the standard procedure). It is also specified that these trials were made in waking conditions (indicating the percipient's state, as opposed to trance conditions), with undifferentiated Telepathic and Clairvoyant possibilities (TC: refer to Appendix 2: Psi vocabulary): Rhine was holding each card face down, under his hand, after first looking at it.

Notice here, without any malicious intent against Rhine, two potential sources of flaw: the first one is total reliance on a *single experimenter*'s report and this will be fully discussed later. The second one is the possibility of **unconscious physical reactions** (like for instance, a faint nod each time a star pops up) from the experimenter, that could be detected by the percipient, since observer and percipient are in the same room. Now Rhine was seemingly aware of both arguments since he specified against the latter that he *"visualised the figures in such conditions, without verbalisation, as anyone may do with deliberate effort. And the image, not it's name, was in consciousness; so the 'involuntary whispering'* (which is a kind of unconscious physical reaction) *ghost need not*, he thinks, *haunt us at this point"*. He also ensured that during the series of 9 successive hits, Linzmayer kept looking towards the window and not towards him.

2.b) A progressive approach to hinder sensory cues

During the next three days, Rhine came to a total of 600 trials for the purpose of statistical evidence. 360 of the 600 trials were still under the same conditions, i.e. undifferentiated ESP with the observer (experimenter) looking at the back of the card. The other 240 trials were made as pure clairvoyance tests (PC, see Appendix 2: abbreviations for Psi vocabulary), implying that this time, Rhine had not even glanced at the card before calling. The scoring was about the same under the two conditions, as may be seen from Table 1.

Rhine goes on to specify that out of the total of 360 TC tests, 145 were carried out with a motor going that would effectively submerge any conceivable 'involuntary whispering'. And he continues: in 120 of these 145, the cards were screened and

Linzmayer's eyes turned away. These yielded for the 125 a score of 57 hits, which means X=11.2, i.e. a highly significant value, in accordance with the rest of the results (see table 1). Under such conditions, he says, none of the sensory modes of perception were at work. We can notice here how Rhine used to change the setting little by little as he advanced through the pack of trials he was submitting to the subject: there are two reasons he gave for this, and both of them will be at the core of the debates later on.

2.c) 2 reasons for this progressive technique

The first reason would be to minimise the alleged **fatigue effect**, which is the following: he and his colleague had been observing that each percipient who gave significant results for the first sets of trials, would steadily decrease their score after several thousand trials, till eventually it dropped to chance expectation. He interpreted this effect as a state of mental boredom, since in real experimentation, he could understand that guessing a card a thousand times for the mere sake of getting data could cause *a percipient's motivation to drop*. We will see later that the sceptics have expressed another interpretation for this. Anyway this is why a *regular change in conditions* might slow down this fatigue effect, confronting the subject's mind with a different environment each time.

The second reason, which is closer to Rhine's typical strategy, is linked to the so-called **climate of benevolence** between the observer and the subject - later on, a lot of researchers explained the failure of others in their survey because they did not pay attention to this factor. Rhine advises people who would want to repeat these experiments, such things as: appear *friendly* to the percipient, encourage him at any state of the survey, do not force him into changes to the setting; in other words avoid behaving like a censor who stares at his gestures and hints at cheating, whenever results seem good. In brief, act in a relaxed way, and apply subtle changes to the setting to improve *control*, because anything brutal might hinder confidence and chase away capabilities.

This discourse has thereafter provoked a controversy as for the lack of *control* this logic would imply; therefore one of the crucial conditions for this type of experiment might have been to strike the right balance between firm control and humane behaviour... Yet several decades later, control will seem much improved, and it will be interesting to see the significance of Rhine's discourse.

2.d) The downward curve

Linzmayer performed up to 8,724 trials over two years, going through a very diverse range of experimental settings, and on account of the lack of space here to display the other detailed sessions, the previously given analysis of the first 600 trials will prevail as an idea of the techniques employed. This **diversity** of setting by the way scores in Rhine's favour, since if any peculiar flaw were spoiling his experiments without his knowledge, this flaw might have been countered by a particular type of setting, which would cause the number of hits to drop. And yet on his best subjects he obtained significant results for most of his settings. Now towards the end of the whole series of experiments, as explained earlier, the team of observers noted that Linzmayer's results were curving down to chance expectation. Nevertheless, the statistics for the total data still show a strikingly high significance. At the same time, Rhine decided to look for another "gifted" subject.

3) Investigating the experimental setting with Pearce as subject

Mr J.G. Pratt, assistant in their department, was engaged during the year 1932 to help in the necessary prospecting for more good subjects. He carried out 10,035 tests on 15 students with a yield of 144 above chance expectation which is more than 5 times the probable error (X>5); but his main contribution to the overall survey lay in the discovery of Mr Hubert Pearce, a young ministerial student whom Rhine had asked to submit to the tests on learning that his mother was reported to have possessed Psi ability. 2,250 *witnessed* trials were made in clairvoyant card guessing. The yield was 869, which means an average of 9.7 per 25. The huge deviation from chance expectation reached 32.75 times the p.e.... The experiments were interrupted by Pearce's appointment to ministerial service for the summer. Rhine offered him to come back the following school term. He accepted and they could keep on doing packs of trials (usually around 1,000 calls each) with a great variety of experimental conditions.

3.a) The experimenter does not look at the card

The data concerning Pure Clairvoyance Experiments are gathered in table 2 below, including some qualitative elements. These can be analysed right on the table with about the same essential remarks as in the study of Linzmayer; his technique of gradually improved control can be examined there. The most litigious element in the beginning of the setting seems to be the fact that he was *allowing Pearce to touch the top card before calling it.* He prevented this contact on step 3, where the number of hits fell, even if still significant (X>4), but then allowed it again.

Table 2

Clairvoyant perception, conditions guarding against Pearce's sense perception

Serial n° & Condition changing	n° of trials	n° of hits	Dev.	p.e.	Value of X	Average per 25	Remarks
1. general B.T. (see Appendix 2)	5,000	1,834	+834	19.1	43.7	9.2	
2. subject looks away from cards	650	279	+149	6.9	21.6	10.7	not much change in conditions, but 1rst runs low
3. same as 2, plus calling before removing	475	236	+141	5.9	23.9	12.4	little real change; 1rst low
4. same as 3., no contact with cards	275	74	+19	4.5	4.2	6.7	great change; first 4 ran below chance. Last 4 average 8.5
5. same as 3, plus new cards; data on first 3 times used	1,675	626	+291	11.0	26.5	9.3	1rst use runs as high as 3rd use of cards
6. (a)screen concealing cards (B.T.)	300	99	+39	4.7	8.3	8.3	
(b) same, plus P.T., (agent screened)	300	116	+56	4.7	11.9	9.7	began very low
7. D.T., pack left unbroken till end of run	1,625	482	+157	10.9	14.4	7.4	
Totals, P.C. (except 6 (b))	10,300	3,746	+1,686	27.4	61.5	9.1	

To add topics of reflection to the analysis of Linzmayer's results, two main innovative ideas can be mentioned, the second one being a transition to the model of experiment that will be discussed in part III:

3.b) using a professional magician as a witness

Soon after the complete set of experiments on table 2, Rhine tries to deal with the witnessing problem. Indeed, some of the series of trials reported rest upon the good faith of non witnessed observers, and this would soon disturb the critics. Hence, he displayed another amount of data where each in turn, one person from a group of 7 people came in the room to witness the experiment, among whom four university doctors (Dr William McDougall included), and a professional magician, named Wallace Lee. He reported that no one raised any question as to the genuineness of the effectual exclusion of sensory perception of the card symbol. Table 3 shows the scoring before and during the stay of the magician, who was then taking any part or position desired. The visitors all produced a drop in Pearce's scoring. In column B is given results taken from the point of entrance, continuing until there is an upward turning of the score curve that remains above mean chance expectation for at least 2 runs. Column C, then, contains the scores made after the up-turn of the curve described as the stopping point for B, which means after Pearce had become "adjusted" to the visitor's presence. These curve observations would reinforce Rhine's opinion concerning his concept of *climate of benevolence*, whereas sceptics would observe that the results were less impressive because of better *controlling*. Mr Wallace Lee tried a few series himself with only chance average results. He was reported to say frankly that he was convinced. *"It appeared that he was, at least as far as we all are, 'mystified'."* commented Rhine.

Now looking at the quantitative data, Rhine realised that Pierce was scoring relatively low, before and during the entrance of the magician, (he scored 7.4/25 before, and 7.5/25, after the "adjusting" period, during which strictly no ability was showed), compared to the case of the other visitors, and more generally compared to his usual results at that time. Rhine explained it by the fact that Pearce was somewhat ill with tonsillitis on the day the magician was present.

Table 3

Clairvoyance, effect of visitors on Pearce's performances

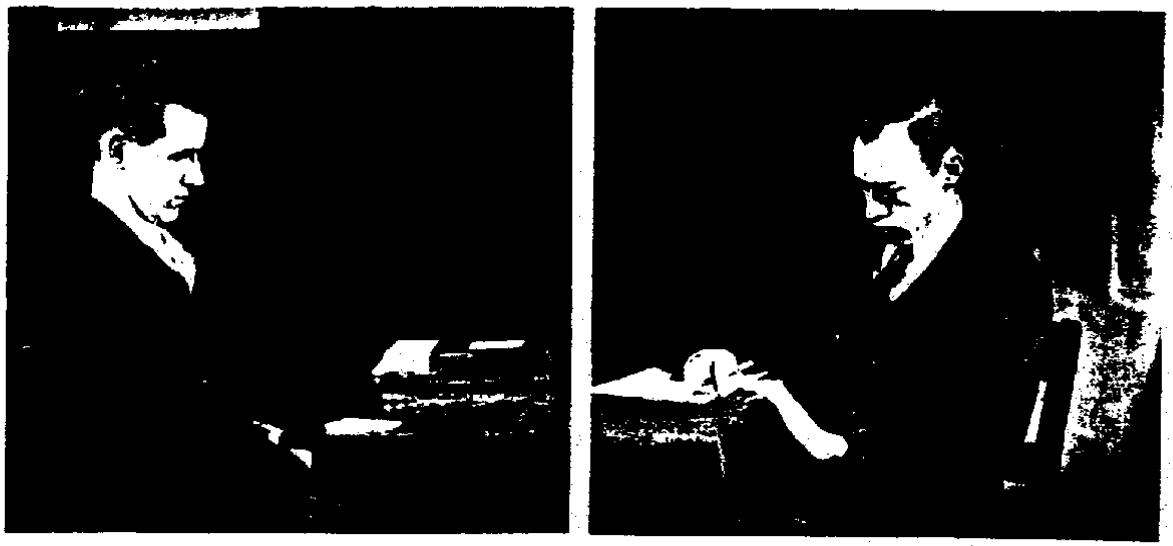
Date 1933	Visitor	Before entrance			Visiting Witness Present					
		Control period A			Control period B			Control period C		
		Trials	Hits	Av/ 25	Trials	Hits	Av/ 25	Trials	Hits	Av/ 25
02/02	Dr W. McDougall	350	132	9.4	125	33	6.6	250	105	10.5
23/02	Wallace Lee, the magician	150	45	7.5	275	55	5.0	125	37	7.4

3.c) A long distance clairvoyance experiment

Towards the end of the "round" with Pierce, Rhine set a new type of experiments, which he labelled long-distance P.C. (B.T.) tests. This time, the experimenter was *out of sight from the percipient*. Mr Pratt, who supervised the experiment, was in a room of the Physics Building of Duke University. Every minute during the running period, he picked up a card taken from a cut and shuffled pack, that lied on the table before him, and put it face down on top of a book. He did not look at it's face. At the beginning of

the same minute, Pearce, in the Duke Library, over 100 yards away, tried to perceive the card then 'exposed' by Pratt, and then wrote his choice down. At the end of each pack, Pratt turned it face up, and wrote down the order of the cards. He then shuffled the whole pack and started again. After each session, Pierce and Pratt slid their respective notes in a sealed envelope and gave them to Rhine, before meeting and comparing data. Pearce was reported to be magnificently successful, after failing for the first few packs, *"as he nearly always does with a new condition procedure"*, said Rhine; The first runs mounted as follows (out of a 25 Zener cards pack): 3, 8, 5, 9, 10, 12, 11, 12, 11, 13, 13, 12. The total 300 at that distance averaged 9,9 per 25; X=12,2 and thus, the chance hypothesis was excluded. In fact, in his B.T. with the same technique, carried out in the same room with the cards but the latter invisible - conditions the same as reported just above, except that Pearce was a few feet from the card instead of 100 yards - he obtained only an average of 7 (this still yielded significant results) in 300 trials. It might have been harder, said Rhine, to sense the observer right in the same room than him off 100 yards away... The researchers concluded that distance did not matter in that mode of perception. By the way, Einstein, who had just been received at Princeton University at that time, emitted his opinion on Psi abilities: he considered it as very unlikely since it was seemingly going against physics basic laws (example here, transmission of energy would not decrease with distance). But anyway the cards were taken to the Duke Medical Building, with over 250 yards between cards and percipient. The results were similar. Rhine considered this type of experiment one of the most impressive proofs for the existence of Psi abilities.

As an objective criticism over this type of experiment, that had been further developed by Rhine, one can remark that sensory cues during the experiment seem definitely wiped out, but also that attention for the *control* factor must be reported this time to the process of collecting the information from both Pearce and Pratt. Working in a lab should be safer, as to maintain information confidential, than having to take pieces of information from two different places; there might be a leakage. This will be an argument in favour of the latest experimental setting in part III, where the idea of keeping both parts out of sight from one another is developed.



Mr J. G. Pratt (left), as he handled the cards in long-distance B.T. clairvoyance with Mr Hubert Pearce (right) as percipient (E.S.P.). Distance, 100 yards and 250 yards. The card, kept face down, lying on the book, is the one 'exposed' at the moment.

4) Emergence of a controversy

4.a) Rhine's conclusions

Linzmayer and Pearce were the two best percipients of the survey and contributed respectively to 8,724 and 17,250 trials out of the 90,000 total. Now according to the team's data, there have been a lot of other percipients who scored above the significant line of X=4 (under 1 chance in 142 that the results are due to chance), and this altogether placed the overall results of the survey well above minimum significant value. Therefore, when he published his survey, the conclusion of his report was outright support for the existence of extra-sensory perception. He comments in his book: *"What alternative to E.S.P. is there left, except to suppose that we are all (a dozen or more are involved) playing a deeply complex game of deception, or else are thoroughly irresponsible and unreliable?"*.

Rhine and his pairs did not know how to explain their results, but they were starting to elaborate a sort of psychology of the mental state they had been observing. They stated that temper affected Psi abilities: scores would rise when the subject was encouraged, according to his *climate of benevolence* conception, and scores would fall when the subject was depressed; Pearce's results, for instance, fell considerably after his girlfriend left him. Performances also declined when the sessions lasted too long. Psi researchers had observed this tendency much earlier than Rhine. But the sceptics were interpreting these declines quite differently, as will be seen in the next chapter.

In 1934, Rhine published his book entitled "Extra-Sensory Perception", and the major effect on the scientific community was that it could provide tangible facts, statistical results, and thus pull Psi research away from mysticism of the XIXth century, that was strongly present in the mind of many scientists. But it would soon draw controversies and strong criticism mainly from traditional psychologists. These controversies are rich in experimental argumentation and will help us visualise the margin of progression of our three factors introduced in part I:

4.b) Overall criticism of Rhine's experimentation

To begin with, some of the critics were bluntly contesting Rhine's **statistics**; In 1937 only, they were reduced to silence after the American Institute of Mathematical Statistics had declared correct the mathematical part of his survey.

The sceptics then observed that a test in which only one experimenter operates and writes down the results without being controlled, cannot guarantee against fraud or mistakes. For instance, cards shuffled by hand, like in the early tests, may not offer, each time, a disposition entirely new and random. This allows the subject to remember their earlier disposition and therefore to enhance their score for each new series. This is what Rhine had called in his book the hypothesis of **rational inference**, a flaw that would merely distort the *true statistics* factor.

On top of that, **unconscious physical reactions** from the experimenter might be detected by the percipient, as could be strongly suspected later on during Soal's tests (see part III). However, one has to notice that this argument is valid for non PC experiments only, since for Pure Clairvoyance, the experimenter has not seen the card.

However, Rhine seemed to get aware of these potential sources of flaw quite rapidly, and introduced more rigorous proceedings. Experimenters were not standing in the same room as the percipient; He asked other researchers to come and double-check

scorings and results, and used an automatic card-shuffling machine. His detractors noticed that the results were a lot less impressive after Rhine had set these different *control* systems. For Rhine, this was simply affecting the factor of *benevolence climate*, and he claimed that significant results were still observed.

Along with time, some events will help to clear up the main claims of the sceptics as for a reliable experimental setting. The need for the elimination of *sensory cues* during the experiment will keep the idea of separating observer and percipient, and lead to the Ganzfeld setting,. The need for controlling the experimenter himself will confront Psi researchers with the *reproducibility* problem. But this is the story of part III...

Appendix 1: Abbreviations for statistical vocabulary

D	Deviation from mean chance expectation (np).
np	Number of trials (*n*) multiplied by the probability of succeeding on each trial (*p*), which gives the Mean Chance Expectation.
p.e.	Probable Error of *np*. This is the deviation from *np* at which the odds are even that it was or was not due to mere chance. The formula that gives p.e. closely enough is $p.e.^2/.67449=npq$ where *q* is the probability of getting a failure. As an example, in the case of $n=1000$ trials, as $p=1/5$ and $q=4/5$, the formula yields p.e.=8.5, which means that if you obtained 208.5 hits out of the 1000 trials (in theory, of course, since 208.5 is not possible), odds would be 1 against 1 that your result would be due to chance.
X	This is the relevant quantity of an experiment; it is the value of the deviation from *np* divided by the p.e.. When the deviation is 4 times the p.e. (X=4) or more, the deviation is regarded as 'significant', i.e. reliably showing a principle beyond 'chance' activity; odds against chance are then 142 to 1.

Appendix 2: Abbreviations for Psi vocabulary

ESP	Extra-Sensory Perception; perception without the function of the recognised senses.
PT	Pure Telepathy; that is, extra-sensory perception of the mental processes of another person. 'Pure' refers to the absence of objective representation of the mental act or image, which might permit of clairvoyance by the percipient.
PC	Pure Clairvoyance; extra-sensory perception of objective facts. 'Pure' refers to the elimination of the telepathy hypothesis from the experimental situation.
BT	Before Touching; clairvoyant card calling, with shuffled and cut pack of 25 cards placed face down before the percipient. He calls the top card and the call is recorded and the card removed. After 5 calls, or after the entire 25, the calls are checked against the inverted pile of called cards. BT-5 represents the condition of checking after every 5 calls; BT-25, after the whole pack.
DT	Down Through; clairvoyant card calling, with the cut pack of cards remaining unopened until after the 25 cards are made; hence, calling "down through", is done without removing the cards called, until the end of the run of 25.
Free Choice	Free Choice Experiments are experiments where the percipient has been given no list of objects in which there is a target to hit, unlike the more convenient method of Forced Choice Experiment.

PART III: Towards the reproducibility criterion

1) Suggestive events for better control

1.a) Soal's conclusions on unconscious physical reactions

In the thirties, Vaudeville telepathists and ordinary people were tested by the British mathematician S. G. Soal. In 1939, he had realised more than 120,000 experiments on 140 individuals from all backgrounds and cultures. The most extraordinary subject of Soal was the illusionist Frederick Marion, born in Prague. He was able to find hidden objects even in an indescribable mess. His rate of success exceeded the average but the scientist, familiarised with the methods of other telepathists, such as muscular reading, was rather suspicious about the authenticity of Marion's gifts. To make sure that he was not exploiting **unconscious physical reactions** from the experimenters, he built a sort of sentry box, in which the observer would stand during the test. The idea was that their movements would be restricted, and most of their body would remain hidden, but they still could keep the necessary vocal and visual contact.

According to Soal, who published a report entitled "Preliminary Studies of a Vaudeville Telepathist", Marion lost his faculty of discovering hidden objects when experimenters were in the box. This strongly suggested to the researchers either to operate experiments at a distance, or to ensure that the experimenter would not have access to any information on the target of the guess (thus in case of non PC experiment, for example, an agent would be introduced, out of sight of the percipient).

1.b) Psi research legitimacy and the reproducibility problem

Above all, the main criticism that scientists addressed to Rhine was that his work could not be reproduced. In other fields of research, experimental results were considered uncertain as long as they had not been reproduced by independent researchers. The **non reproducibility** of ESP experiments has remained the major obstacle to the recognition of parapsychology as a scientific discipline.

In 1969, ESP researchers finally obtained an official attachment to a scientific institution, since the Parapsychological Association, founded 12 years earlier by Rhine and others, acquired the status of society affiliated to the American Association for the Advancement of Science. If the hypothesis of ESP and other paranormal phenomena finally had not been proven, parapsychology had at least gained it's legitimacy as a domain of scientific research.

An easy way of showing the need for reproducibility in science is to examine another side of the *control* factor, so essential to the reliability of an experiment: control over the experimenters themselves. The only way found to avoid mistakes or fraud to take over the result of an experiment is to get it conducted by several other teams in the same conditions (in reality, as close as possible).

For sure to the dismay of most Psi researchers, and Rhine first, this rule would soon be backed up by an affair, that happened in Rhine's own laboratory. In 1973, Rhine enrolled a young physician, named Walter J. Levy, to ensure his succession at the head of the Foundation for Research on the Nature of Man, created by Rhine in 1962

with private funds. Rhine was then 70 years old and was looking for young people to carry on his work. Levy seemed the ideal choice, because he was inventive and a hard worker, although some of his colleagues had judged him ambitious and scheming. Anyway, in 1974, Levy was doing experiments on rats (Rhine's team turned to seeking Psi abilities among animals; one of the advantages is that fraud hypothesis from the percipient was then ruled out...), plugging electrodes in their cerebral pleasure centres, to see if they could themselves finally influence the electrodes, to stimulate the area for sexual pleasure. At the beginning, he seemed to get positive results, but two of his collaborators, James Davis and James Kennedy, suspected him of fiddling his data. They decided to set a trap for him. On the 11th of June, they plugged an extra set of cables to connect Levy's computer to a computer in another room, in order to follow up the whole experiment. Davis hid in Levy's lab. Effectively, Davis saw Levy fraudulently introduce positive data in his computer, whereas the data recorded by Kennedy's computer revealed only chance expectation. They informed Rhine, who asked Levy for an explanation. The latter admitted the deception, and Rhine dismissed him on the spot. This was a double disappointment for Rhine because he had to publicly announce the circumstances of the dismissal, so that other researchers would not be misled by any false data already published by Levy. One can easily imagine the effect of such an event on the image of Psi research among traditional scientists thereafter...

But the real consequence for this was that if Psi researchers wanted to get rid of this intrinsic weakness from the *control* factor, they definitely had to find a kind of experiment that could be reproduced easily. The **reproducibility challenge**, more than a challenge related to Psi research, was a problem essentially linked with **scientific experimentation** on the human mind. The climate in which they had to take up this challenge after this event was of course very discouraging; attacks came from scientists who refused to give any consideration to progress in this field, an attitude which can seem unreasonable since Psi researchers had not even conceived a proper experimental process yet.

2) New experimental tools and new ideas

2.a) Random Event Generators boost up the *true statistics* factor

Anyway, at the time of Rhine's misfortune, methods had enormously evolved (recall computerised records), improving not only the *true statistics* factor, but also the *control* factor, and to some respects the *user-friendly* factor as well. Instead of Zener cards, used until the sixties, and manual experimentation, test scorings were now directly computerised, which ruled out a large amount of mistakes and modifications found in hand-written reports. The *true statistics* factor also improved thanks to special devices called 'Random Event Generators' (REG). This was the next step after Rhine's initiative of choosing 5 distinguishable patterns, and it would also greatly improve the *control* factor. They were especially useful for treating PT types of experiments, since statisticians had established that the human being had trouble expressing a choice at random; for instance, when individuals are asked to choose a figure between 1 and 10, most of them choose 7.

One of the first REGs conceived to test precognition was built by physicist Helmut Schmidt, who was at that time doing research in Boeing's laboratories. The

essential part of the REG was a piece of radioactive strontium 90, an unstable element which emits electrons when it decays. The decaying process, according to particle physics theory, occurs at a precise probability rate. This apparatus would then light on and off, in a perfectly fortuitous order, one in four bulbs coloured differently, and the subject would have to guess which one would switch on. Over a long series of tests on Schmidt's device, an average of 26,7% of good answers have been reported. The number of tests was sufficiently high to achieve significance, and such results are known to have puzzled a lot of scientists from traditional disciplines.

2.b) But people need a more "user-friendly" experiment...

It seems however that other types of more *user-friendly* settings were more popular, aiming at enhancing at the same time the percipient's "dispositions for imagination": Indeed, it had been noticed on these basic REG experiments, that a "gifted" percipient would not keep high scores for a long time, allegedly because of too much repetitiveness that would easily cause scorings to drop... And thus a complete series of experiments would require to gather too many percipients. Seemingly, Rhine's **fatigue effect** was still treated very seriously by optimistic experimenters, and very suspiciously by their detractors.

A lot of more "open" experiments on very distant clairvoyance were thus conducted in the seventies by respectable people; subjects were supposed to describe precisely what the agent (an additive person who is not the experimenter) was doing a thousand miles away. Some tests took place on a submarine... and they were shown to give surprising results: a lot could be argued about them but this report will remain focused on experiments with **quantifiable statistics**. This leads us finally to a type of test that has shown great hope for meeting the requirements by the sceptics and also by the researchers: the Ganzfeld experiment.

3) The Ganzfeld Experiment: a good compromise?

In 1971, the Psi-researcher Charles Honorton introduced the notion of Ganzfeld (or "full field", in German), some sort of deep dreaming state, from which Psi abilities might come up. The tendency of parapsychology was indeed to associate the phenomenon with some altered state of consciousness. Anyway he conceived the corresponding experimental setting, described below, that is truly interesting regarding our three factors.

The evolution of Honorton's Ganzfeld work is an example in itself of how parapsychology has steadily changed in response to criticism. experiments began in the mid-1970s and were so popular that by 1990, nearly 70 studies had been reported. Undoubtedly the main reason for this interest was the high rate of success in terms of scoring claimed for this method. Before having a qualitative look at a large sample of such experiments, it seems important to note, like in Rhine's case, that Honorton has always been considered, to the majority of the sceptics, as an honest investigator, and if his results have been contested, it was essentially interpreted as the product of an unforeseen flaw, rather than a deliberate deception.

3.a) General description of the setting

The basic idea behind the study was that during a trial, only the inner experiences of the percipient might have some relation to the target. Hence the next step was to isolate the subject from ordinary sensory input as much as possible:

Isolated inside a steel-lined cubicle with walls a foot thick, the subject relaxes on a comfortable bed and describes his/her imagery. A homogeneous visual field is created by placing halves of ping-pong balls over the subject's eyes and a more or less homogeneous auditory input is achieved by presenting the subject with white noise delivered through earphones. The main senses are thus provided with a constant and shapeless input which in general results in diminished contact of the subject with the external environment and increased **sensitivity for internal processes**. Both the reduction of sensory 'noise' and a shift of attention to internal experience are supposed to be an incentive to Psi. Meanwhile an isolated 'sender' (an agent who is not the experimenter) is looking at a target picture chosen randomly from a set of possible targets in another soundproofed room. After the half hour or so is over, the subject has to try to match up the imagery experienced with each of the pictures he is presented and choose which was the target.

When in the Ganzfeld, the percipient usually gives a more or less continuous verbal report about his ongoing thoughts and feelings. These reports are often written down by an experimenter who is either situated in the percipient's room or has auditory contact with the percipient by means of an intercom system. Occasionally, percipients have also been required to provide verbal state reports during the Ganzfeld. It can be noticed first that the *user-friendly* aspect, although improved due to "mental comfort", has lost a lot in **time consuming**.

3.b) A view of the statistical aspects

Often the target pool (all targets available for use in the study) was split up into target sets, each set containing an equal number of target pictures (usually four, a number that is not too high and thus does not need so many trials to validate results slightly above significance level). Selecting a target for a trial therefore often involved first selecting a target set, and then selecting from that target set one picture as the target. Most target pools consisted of pictures or slides. Pictures in a given target set were selected as different as possible. A notable and interesting exception to this process is the BTP (Binary Target Pool), a target pool without fixed target sets in which each target is characterised by a unique combination of 10 aspects. Hence the total pool consisted of 2^{10} (1024) targets. Applying the BTP makes it possible to treat a free-response trial as if it were a forced-choice test of 10 trials, by requiring the percipient to code the thoughts as regards containing or not containing the same aspect. Hence one free-response trial can be considered as 10 trials (one for each aspect), with a chance expectation of 1/2 for each, which considerably improves the *user-friendly* factor of the study regarding the time spent and the boredom aspect.

3.c) The judging process

At the conclusion of a session, the subject would be shown all four images from the group and asked to pick the one that seemed to best match the imagery experienced in the chamber. In a step that was later seen as controversial, this judging process was aided by the researcher, who pointed out correspondences that the subject might otherwise miss. Honorton felt it essential that the experimenter help the subject, who often emerged somewhat drowsy or disoriented from the experience. To his mind, the results were not invalidated by this intervention because the experimenter had no idea which film clips had been used (only the sender had the information).

If such an experiment were ruled by chance, subjects should pick the correct target only one in four times - a hit 'rate' of 25 per cent. In his early experiments, Honorton was disappointed to find that, despite a few seemingly impressive matches, scoring was not significantly above chance levels with the static photograph targets (45 hits in 165 trials). But with the film and TV clips - a much richer source of target imagery, Honorton argued - the hit rate was about 40 per cent (77 hits in a total of 190 trials). The chance of this being a statistical fluke was just two in a million.

4) An analysis of the early experiments (1974-1980)

There has been an overview of all the published studies from the first one in 1974, to 1980, taken from a Research Letter in the Journal of the Society for Psychical Research. This study was made by an independent (did not participate in any of the recorded tests) Psi researcher, Sybo A. Schouten from the University of Utrecht, in Holland. His aim in that letter was to give an objective criticism on the use of the laboratory, and especially locate all the potential weaknesses that one should be aware of, when designing a experiment. According to an analysis made by Susan Blackmore in 1980, 25 of the 50 published and unpublished studies carried out since the first was published in 1974 yielded a significant result: an unprecedented rate of success of 50%... The data for this study is gathered in an inconveniently large table expanding on several pages, and preceded by a list giving the reference for each of the 34 experiments, and a long index explaining each abbreviated variable. For the purpose of keeping the report easy to read, only the qualitative remarks will be here displayed, and not the exhaustive data.

4.a) Discussion of the *true statistics* factor

The author of this study presented his analysis of the potential sources of error, which were thereupon taken by the sceptics as a powerful criticism of the early experiments. The factor which is most striking when considering the data, and which is due to the time it takes for a single trial, is the **small number of trials** per experiment (around 70 trials for each experiment). As described above, studies are usually evaluated as Forced-Choice studies (see part II, Appendix 2: abbreviated Psi vocabulary, at *Free Choice*). From forced-choice studies, researchers know that one has to be careful when evaluating an experiment of, for instance, one run (25 trials, p=1/5). The distribution is skewed, the continuity correction exerts a relatively large influence, and especially the randomisation of the targets becomes a problem considering the always present response bias (what goes against the *true statistics*): Since targets in studies are much richer in content than for instance Zener cards, a bias due to a hypothetical unconscious

a priori preference for particular targets is stronger (remember Rhine's reasons for choosing Zener cards). This type of bias is called zero-order response bias, and it could be reduced by balancing the targets over the experiment. However, this would also lead to some serious problems on the side of the *user-friendly* factor, e.g. that either subjects and experimenters should not be given feedback as regards the success of each trial, which tends to make the experiment rather dull for all participants. Indeed, if feedback is given, subjects and experimenters learn during the course of the experiment which pictures have not yet been used as targets and hence the targets for the remaining trials have a higher probability of being correctly identified.

To facilitate the judging to be made after each session, an experimenter has to write down verbal reports from the receiver. However, rarely will the complete utterances of the subject have been written down, they will be shortened and if the subject speaks fast, a selection might take place again. Consequently, the thoughts so recorded by the experimenter can bias the results, especially when content analyses are to be carried out, if the experimenter who takes these notes is familiar with the target set or, in the case of a small one, with the entire target pool. Then he might be biased to pay more attention and rather to note elements which look familiar and fit some pictures in the target set and to ignore elements which are clearly unrelated to any of them. Assuming that the experimenter does not know the actual target *this will not influence the probability of achieving a hit*, but it might lead to a closer correspondence to other pictures suggestive to displacement effects.

4.b) The ghost of sensory leakage is back!

Another fluke which critics qualified as a very gross flaw factor later on, would occur during the judging process: when utilising for the judging the original target set including the target which had been handled by the agent ('sender'), the possibility arose of an effect of **sensory cues**. Pictures could be marked (inadvertently or intentionally) by fingerprints, creases, or other signs, slides having been placed in projectors could still be warm, etc.... Using duplicate sets avoids these problems, but even then using the same codes on both sets to identify the pictures offers an opportunity for malpractice.

5) The latest experimental work on Ganzfeld (1983-1992)

5.a) Improvement of the setting

The early work with Ganzfeld chambers, for which some negative aspects have been briefly exposed above, ended in a stalemate between parapsychologists, who felt they were getting positive results, and sceptics, who saw problems with every experiment. But in 1986 things changed. The chief protagonists from each side, Honorton representing the parapsychologists and Ray Hyman, a psychologist at the University of Oregon and a member of a sceptics' organisation called the Committee for the Scientific Investigation of Claims of the Paranormal (CSICOP), were about to publish another round of claim and counterclaim. However, by chance, they met at a conference and over lunch realised that their real argument should be not about the theoretical existence of parapsychology but about how to do research that was methodologically rigorous. Out of this dialogue came the automated experiments that Honorton used at his New Jersey laboratory. One of the most important advances was to use a computer and videotape system to select the targets and record data. This

reduced the possibility of both deliberate fraud and inadvertent 'sensory leakage' such as fingerprints on photographs. The automation of data collection made it harder for an experimenter to fudge results by abandoning trials that were not going well or altering the tally of hits at a later date. Honorton also went to great lengths to protect against possible collusion between subject and sender. The subject was isolated in an acoustic chamber with walls 30 centimetres thick. The sender's room was less secure but still had four inches of acoustic padding. Copper screening was used as a precaution against concealed radios and the only doors to the cubicles opened onto the experimenter's room. As well as these physical safeguards for better *control*, Honorton made methodological changes to protect against deception. Rather than using 'star' subjects - people noted for public performances of psychic abilities - as many previous studies had done, Honorton used 241 subjects most of whom contributed only a single session. He also used a total of eight experimenters to conduct the sessions.

5.b) Honorton's results

Honorton published the fruits of his work in 1990 and some more detailed analyses in 1992. The results had their curiosities. For instance, two of the experimenters, Mario Varvoglis and Patricia Derr, produced results that were only fractionally above chance levels while the other six experimenters obtained results that appeared more positive. Over the seven year study, the series produced an average hit rate of 34 per cent compared with the 25 per cent that would be expected by chance, which, considering the amount of trials, was a very significant number, and considering the number of participants, attained a high level of credibility too.

5.c) Replication: the ultimate judge?

The reaction of sceptics since the publication of Honorton's results has been mixed. Hyman, the strongest critic of previous work, says he is reserving judgement until there is an **independent replication** of the results; remember this is only one piece of kit and one lab. The strength of the design was that it was supposed to produce effects that would be straightforward to replicate. Hyman does concede that the methodological improvements Honorton made provides sceptics with their stiffest test to date: *"There are a lot of minor things I would quibble with but Honorton met most of the objections I had."* Yet the study has not changed his views. He expects time will show that Honorton's results can be explained by some hidden experimental flaw just like so many other parapsychology claims before them. It is to be noted that other critics like Nicholas Humphrey, a psychologist currently working on a four-year Psi study at the University of Cambridge, are more radical towards that survey: *"The experiments are interesting and somewhat better done than others before them, but they certainly are not foolproof or conclusive... nobody need take them seriously until there are replications".*

Humphrey backs his judgement with the same kind of argument as Schouten for earlier experiments; he notably estimates that sensory cues on the tape used are still possible, and that the first order response bias (see explanation above) is still potentially distorting the *true statistics*, because he considers that researchers somehow know which target segment is playing, and deliberately, or subconsciously, they do encourage subjects to select the correct clip during the judging process. It is obviously suggested here that one should eliminate the helping part of the judging process, and let the percipient choose alone. Honorton, with his psychological approach of Ganzfeld,

considered this part as very important, especially to test a percipient who will accomplish only one trial for the whole study. His approach is comparable to that of Rhine with his gradual control theory;

In fact, the real question might lie as follows: have researchers been totally misled by too much consideration of their "Psi emergence conditions", or on the contrary, if the phenomenon does exist, have they allowed for it's detection by careful data picking, in such a way that if the percipient loses this support (which is the case if they try their abilities under sceptics' laboratory conditions; for example at the lab where $100,000 is offered to anyone successfully undertaking the experiment proposed), his ability would be more or less inhibited?

6) Perspectives

To conclude, the general standpoint on the paranormal is not as rigid on both sides as it had been in the early eighties; the Parapsychological Association considers the existence itself of Psi ability is still only a hypothesis. Likewise, Honorton's mission had been to put the disreputable field of parapsychology on a firm scientific footing, and to design telepathy experiments that could be rigorously interpreted; Susan Blackmore, a noted debunker of psychic claims from the CSICOP, comments favourably on the recent experiments realised by Honorton and his colleagues on the auto-Ganzfeld: the 11 independent sets of experiments had been published in 1990 to respond to the remarks of the critics, before any attempt of replication. After the closure of his New Jersey laboratory in 1994, due to a lack of funds, Honorton joined the parapsychology team at the University of Edinburgh, lead by Robert Morris, and planned a further round of investigation. But then Honorton died towards the end of 1994 of a long-standing heart condition.

Morris, who has run the University of Edinburgh's parapsychology unit since its foundation, represents the new reassuring face of Psi research. The first surprise for any visitor to Morris' laboratories is that they are part of the university's psychology department. He says he is fully integrated there and his postgraduate students get their qualifications in psychology. He is also insistent that the unit splits its time evenly between parapsychology experiments and the study of the psychology of magic and deception: *"Unless you understand the techniques of deception or the power of coincidence, you can't begin to do proper work in this area."* Morris invited Honorton to Edinburgh to recreate his automated Ganzfeld set up after his New Jersey laboratory closed. Honorton's sudden death left the project in pieces, but other experimenters have now been recruited to recreate it themselves.

Attention is really turning towards attempts to replicate Honorton's results. As well as the Edinburgh research, Ganzfeld systems are being set up in the US by Daryl Bem at Cornell University, New York, and Richard Broughton at the Institute of Parapsychology in North Carolina, and in the Netherlands by Hans Gerding at the Parapsychology Institute in Utrecht. The Edinburgh team plans to start with a faithful copy of Honorton's system, even down to using the same pool of film clips. If they get a positive result, they will vary conditions to try to eliminate some of the more controversial steps, such as the researcher's participation in the judging process. As noted above, a heavy workload is involved; an experiment will have to reach 200 trials before it begins to produce statistically reliable results and this represents at least 400 hours of laboratory time. It is thus unlikely that any of the four groups will have much

to report before the end of this year. In the meantime, somewhere in the clash between sceptics and Psi researchers must lie a fascinating thesis for someone on the sociology of science and the construction of knowledge...

7) My personal impressions on all of this...

As far as I am concerned, I would find it very suitable if both parties could work constructively together on these replications; for example some sceptical people could themselves replace the experimentation staff in the Psi research labs sometimes, and the illusionists enter the process of being tested in the latest experimental settings, so that they can comment on the possible flaws to track down. And on the side of the Psi researchers, it would be ideal to get a more '*neutral*' atmosphere, as defined by Pr Morris: "*We have to have an atmosphere where people are not afraid to report positive results, but equally importantly, not afraid to report flat nothing.*"

This might help to solve this reproducibility problem. Hence, I have not seen myself any argumentation that would prove the *impossibility* of solving it. Therefore I see no reason - other than purely economical... - for science not to seek in that direction, and I encourage the approach, which Pr Morris and other experimenters today seem to employ.

Conclusion

It finally seems that we are going towards the right setting/methodology, and that the ultimate step might be the replication of this setting/methodology, in order to apply the **reproducibility** criterion. Among the main obstacles that render this prospect a lot less optimistic, there is still this battle between the *control* factor and the *user-friendly* factor that has not ceased: will it be possible for both parties to agree on a setting where controlling and adopting "percipient-friendly" experiments at the same time meet higher standards ? If so, let us replicate it! In case of really high *control* factor, will further studies still render significant statistics? If not, what will one conclude in the case of a poor *user-friendly* factor? The disappearance of experimental flaws or the disappearance of abilities due to bad conditions? Anyway, and whatever the opinion one may have on the final interpretation, I definitely think that the future development and replication of such an experimental setting is a scientifically valuable decision and an interesting experimental challenge in the history of science.

Bibliography

This list displays the main sources of information used for the report. Their confrontation suggested me to set the three simple factors (*true statistics*, *control* and *user-friendly*) to analyse the scientific experimentation of extra-sensory perception.

Books:

"Extra-Sensory Perception", by J.B.Rhine (1934)

"The Sixth Sense, an inquiry into Extra-Sensory Perception", by Rosalind Heywood (1965)

"Les Pouvoirs de l'Esprit", edited by Timelife (1988)

Articles from periodics:

Research Letter written by Sybo A. Schouten, from the Journal of the Society for Psychical Research, in 1980

"Against the rise of pseudo-science", written by John McCrone in The New Scientist, 11th of July 1992

Articles written by Mrs Blackmore and Mr Costa de Beauregard in the special issue about Parapsychology, Sciences et Avenir June 1995

Along with the presentation of the bibliography, I would like to address sincere acknowledgements to three people voluntarily working at the "Groupe d'Etudes Expérimentales des Phénomènes Parapsychologiques" (G.E.E.P.P.) in the University of Le Mirail; they were very helpful in providing some of the research material (issues of Archive Journals from the most respected institutions working in that domain), but more importantly, who could give me the opportunity to sense the human climate in such a peculiar field of research. These people are Yves Lignon, a maths professor who initiated the lab, Francois Bouleux, and Pierre Macias, two of his collaborators, both computer specialists.
I also appreciated a lot the remarks of my tutor for this project, Mr Rupert Salmon, who helped me with the language, and also shared interesting thoughts on the subject.

References

To be aware of the current experimental work done in this field, I would suggest the following publications:

Books on Experimental Parapsychology:

"Parapsychology: the Controversial Science", by R. Broughton, edited by Ballantine Books, New York (1991)

"Foundations of Parapsychology: Exploring the Boundaries of Human Capability", by H.L. Edge, R.L. Morris, R.L. Rush & J. Palmer, edited by Routledge & Kegan Paul in London (1986)

"An Introduction to Parapsychology", by H.J. Irwin, edited by McFarland in London (1989)

"Advances in Parapsychological Research" (volume 6), edited by McFarland in London (1991)

Journals of Parapsychology:

European Journal of Parapsychology, published once yearly by the Koestler Chair of Parapsychology, University of Edinburgh, Edinburgh EH8 9JZ, UK

Journal of Parapsychology, published quarterly by the Parapsychology Press, Box 6847, College Station, Durham, N. C. 27708, USA

Journal of the Society for Psychical Research, published quarterly by the Society for Psychical Research, 49 Marloes Road, London, W8 6LA, UK

Journal of the American Society for Psychical Research, published quarterly by the Society, 5 West 73rd Street, New York, N-Y 10023, USA

The Skeptical Inquirer, Official Journal of the Committee for the Scientific Investigation of Claims of the Paranormal. Published quarterly. Box 229, Central Park Station, Buffalo, New York 14215, USA