

**Is some of the evidence for ostensible
precognition indicative of
Darwinian adaptation to retrocausal
influences?**

Louie Savva

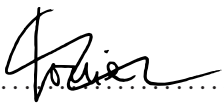
A dissertation submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
of the
University of London.

Department of Psychology
Goldsmiths College

Declaration

I, Louie Savva, declare that this thesis, entitled *Is some of the evidence for ostensible precognition indicative of Darwinian adaptation to retrocausal influences?* and the work presented herein are my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly whilst a candidate for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

Signed: 

Abstract

Parapsychologists continue to report experimental results that imply information from the future can be utilised in the present in a way that defies conventional understanding. Although treating such results as though they are the outcome of a sensory process is common, much less so is consideration as to what psi might be for. This thesis attempts to examine some of the evidence for precognition from the perspective that it might be indicative of Darwinian adaptation to retrocausal influences. The aims of the research were two-fold. The main objective was to replicate and extend previously published, statistically significant parapsychological research, concentrating on precognition for a variety of reasons. The secondary objective was to develop and test a synthesis of evolutionary theory with a needs-based theoretical conception of precognition (namely that precognition is a form of death-avoidance). Primarily this was achieved by testing the effectiveness of using fear-based stimuli (and comparing the appropriate fear groups), though in the final experiment this involved direct testing of an insect species (where the consequence of failure was death) along with a selective breeding element.

Ten empirical investigations are presented which together provide scant evidence for any kind of precognitive functioning. Moreover such results as were obtained highlight both the difficulty in conducting parapsychological research (ensuring results are not normal influences misconstrued as paranormal), and the responsibility of researchers to not make strong claims on the basis of weak or uncertain evidential findings. Given the lack of paranormal findings it was difficult to evaluate the evolutionary perspective (and it is left for future researchers to determine whether such a pathway is a complete dead end). Much remains mysterious and yet the continued rejection of the reality of precognition by the conventional scientific community seems reasonable given the evidence.

Acknowledgements

I would like to thank certain individuals for their help in producing this thesis, though no endorsement of its content is implied.

Thanks especially to Professor Christopher French for his supervision. Rob Davis, Steve Yesson, Maurice Douglas and Richard Billingham for their technical assistance. Special thanks go to Dr Aarti Sawhney. Bernie Lehmann. Liz Smith. Thanks also to the students who volunteered their assistance. Dr Debbie Custance. Dr Susan Finch. Other staff, fellow postgraduates and students in the psychology department and the wider community of Goldsmiths College. Dr Devin Terhune. Dr Matthew Smith. Visitors and staff at the Museum of the Unknown and the Fortean Times Unconvention. The staff and members of the Society for Psychical Research. The users of StackOverflow.com and related sites. Professor J. B. Gilmore. Dr Gert Camfferman. Thanks also to David Mandell.

Special thanks must also be extended to Professor Alan Pickering, Dr David Luke, and Dr Susan Blackmore for their constructive comments received as part of the examination process.

I would also like to recognise and thank The Bial Foundation, Portugal for awarding funding supporting some of this research and to the Parapsychology Foundation, USA, for awarding me the Eileen J. Garrett Scholarship in 2001.

Contents

1	Introduction	10
1.1	An Introduction to the Problem of Precognition	10
1.2	The Putative Evidence for Ostensible Precognition	13
1.2.1	The Sporadic Cases	13
1.2.2	Empirical Tests of Precognition	26
1.2.3	Questions Raised by the Evidence for Precognition	31
1.3	Psi as a Darwinian Adaptation	38
1.3.1	Wonder-Animals and Anpsi	38
1.3.2	Altered States: Both natural and induced	42
1.3.3	Evolution and Psi	47
1.4	Working Hypothesis	57
2	Retrocausal Influences in Electrodermal Activity	58
2.1	Presentiment I	58
2.1.1	Introduction	58
2.1.2	Method	65
2.1.3	Results	68
2.1.4	Discussion	70
2.2	Presentiment II	71
2.2.1	Introduction	71
2.2.2	Method	72
2.2.3	Results	73
2.2.4	Discussion	75
3	Retrocausal Influences in Stroop-Based Tasks	78
3.1	Time-Reversed Interference I	78
3.1.1	Introduction	78
3.1.2	Method	86
3.1.3	Results	87
3.1.4	Discussion	97
3.2	Time-Reversed Interference II	98
3.2.1	Introduction	98
3.2.2	Method	99
3.2.3	Results	100
3.2.4	Discussion	107
3.3	Time-Reversed Interference III	108
3.3.1	Introduction	108
3.3.2	Method	110
3.3.3	Results	112
3.3.4	Discussion	115

4	Testing Psi-Mediated Timing with a Simple Computer-Based Task	118
4.1	Psi-Timing I	118
4.1.1	Introduction	118
4.1.2	Method	122
4.1.3	Results	123
4.1.4	Discussion	126
4.2	Psi-Timing II	129
4.2.1	Introduction	129
4.2.2	Method	129
4.2.3	Results	130
4.2.4	Discussion	131
4.3	Psi-Timing III	133
4.3.1	Introduction	133
4.3.2	Method	133
4.3.3	Results	133
4.3.4	Discussion	134
5	An Indirect Test of an Exceptional Precognitive Claim	138
5.1	Do Independent Judges Agree with the Putative Correspondence?	138
5.1.1	Introduction	138
5.1.2	Method	152
5.1.3	Results	153
5.1.4	Discussion	158
6	Insect Death-Avoidance	168
6.1	Anpsi test and Selective Breeding with Laboratory Populations of <i>Tribolium</i>	168
6.1.1	Introduction	168
6.1.2	Method	175
6.1.3	Results	180
6.1.4	Discussion	181
7	Conclusion	187
7.1	What has been learnt?	187
7.2	Future Directions	193
7.3	Final Comments	195
	References	197
	Author Index	212
A	TRI Appendix	215
A.1	Summary of Questionnaire Data	215
B	DM Appendix	216
B.1	Booklet Instructions	216
B.2	Rating Task	217
B.3	Images	224
B.4	Supplementary Evidence	232

List of Figures

2.1	Diagram of procedure and example physiological response	67
2.2	Line plot of the mean change in EDA	74
3.1	Graphical representation of null and TRI hypotheses	80
3.2	Graphical representation of TRI procedure.	87
3.3	Interaction between responder type and colour congruence for stimulus 2	88
3.4	Values of ΔZ_{RT1} & ΔZ_{RT2} by colour for TRI I	97
3.5	Values of ΔZ_{RT1} & ΔZ_{RT2} by colour for TRI II	107
3.6	Significant Stroop influence at Stimulus 2	114
4.1	Combined density plot of hits for Psi-Timing I	124
4.2	Combined density plot of hits for the control trials	126
4.3	Two combined density plots showing displacement for the control trials	127
4.4	Plots to investigate potential periodicity effects for the control trials	128
4.5	Displacement hit plot & periodicity plot for Psi-Timing II	132
4.6	Combined density plot of hits for Psi-Timing II & III	134
4.7	Displacement hit plot & periodicity plot for Psi-Timing III	135
4.8	CI plot for all Psi-Timing & Control Trials	136
5.1	“Concorde Crash” painting	148
5.2	CI plot showing the mean difference between the event ratings	156
5.3	Scatterplot showing correlation between event ratings	157
5.4	Comparison of Heathrow mortar attack picture and DM’s image	160
6.1	Diagram of testing apparatus & an example image showing inside the testing chamber	177
6.2	Graphical representation of insect testing procedure	179
B.1	Picture 7: “Heathrow mortar attack”	225
B.2	Picture 9: “Dead fish in River Irfon, Wales”	226
B.3	Picture 15: “Indian earthquake”	227
B.4	Picture 28: “Rackhams knife attack”	228
B.5	Picture 33: “Twin Towers sketch”	229
B.6	Picture 35: “Twin Towers colour painting”	230
B.7	Picture 37: “Two miners killed”	231
B.8	Two example bank photographs	232
B.9	Note written after the Neftegorsk earthquake	233
B.10	Note written after the Tokyo terrorist attacks	234

List of Tables

1.1	One historical proposed distinction of parapsychological phenomena	12
1.2	Percentage of events of precognitions	36
2.1	Summary statistics for Presentiment I	68
2.2	Summary statistics for Presentiment II	73
3.1	Summary statistics for TRI I	88
3.2	Klintman analysis for TRI I	90
3.3	Camfferman analysis for TRI I	91
3.4	Frequency of slow & fast RTs for Camfferman analysis for TRI I	93
3.5	Frequency of slow & fast RTs for all trials for TRI I	94
3.6	Frequency of slow & fast RTs (congruence split) for all trials pooled for TRI I	94
3.7	Number of trials across colour for TRI I Radin analysis	97
3.8	Summary statistics for TRI II	101
3.9	Klintman analysis for TRI II	102
3.10	Camfferman analysis for TRI II	103
3.11	Frequency of slow and fast RTs for Camfferman analysis for TRI II	104
3.12	Frequency of slow & fast RTs for all trials for TRI II	105
3.13	Frequency of slow & fast RTs (congruence split) for all trials pooled for TRI II	105
3.14	Number of trials across colour for TRI II Radin analysis	107
3.15	Hypothesised TRI III RT2 results	110
3.16	Summary statistics for TRI III	113
3.17	ANOVA results for RT1 for TRI III	113
3.18	[ANOVA results for RT2 for TRI III	113
3.19	Camfferman analysis for TRI III	114
5.1	Frequency of responses to the question, "Have you ever heard of David Mandell?" 154	
5.2	Frequency of responses to the question, "Do you think it is possible for dreams to provide a paranormal glimpse of the future?"	154
5.3	Frequency of responses to the question, "Do you think all such ostensibly precognitive dreams can be explained in terms of concepts already accepted by modern science?"	154
5.4	Frequency of responses to the question, "Have you ever had a dream that appeared to foretell the future & if yes, how many?"	154
5.5	Summary results from the Australian Sheep-Goat Scale	154
5.6	The seven pictures where a significant difference between ratings for DM and ALT event was found	155
6.1	Results for all insect trials	180
6.2	Two tables showing frequency of the side chosen and side insect was on	180

A.1	Summary results for the Spielberger State-Trait Anxiety Questionnaire (a)	215
A.2	Summary results for the Spielberger State-Trait Anxiety Questionnaire (b)	215
A.3	Summary results for the Spielberger Trait Scale	215
A.4	Summary results for the Beck Depression Inventory	215

Introduction

1.1 An Introduction to the Problem of Precognition

vulnerant omnes, ultima neceat

—TRADITIONAL

I propose to consider the question, “Can humans and other animals perceive the future?” The specific emphasis of this investigation relates not to the future as a concept, that is, how people predict what is going to happen to them as part of their daily lives (what Suddendorf & Corballis, 2007 have recently referred to as *mental time travel*), but is an investigation into the question as to whether some of the evidence for ostensible precognition is indicative of Darwinian adaptation to retrocausal influences. Simply put, do organisms make use of information that originates from the temporal future?

It is a peculiar fact that the word “time” is the most common noun in the English language (e.g. Leech, Rayson, & Wilson, 2001) and yet time is an utter mystery, both as a philosophical and scientific concept. St. Augustine (354–430) wrote,

What, then, is time? If nobody asks me, I know well enough what it is provided that nobody asks me; but if I am asked what it is and try to explain, I am baffled. (Augustine, 2008, xi, 17)

A modern-day perspective adds little. Take, for example, the opinion of probability expert, Warren Weaver,

Science has probably never measured anything as frequently or as accurately as intervals of *time*; but I am not aware that this persistence has at all increased an understanding of what “time” really is. (Weaver, 1963, p. 36)

If time is a mystery, perhaps there are temporal phenomena that are as yet not fully recognised by science. The fundamental aim of the research outlined in this thesis is to confirm previously published statistically significant accounts of ostensible precognition. Without evidence for the elusive quarry being sought, understanding the phenomenon of precognition will be a great deal harder. If we achieve evidence supporting the hypothesis of precognition, the secondary aims are clear. To assess the evidence gathered from an evolutionary perspective and to develop a synthesis of parapsychology and evolution (an evolutionary parapsychology if you will). Although the nature of that synthesis will depend very much on first accumulating evidence for precognition, the basic hypothesis can be simply stated by analogy. Just as a spider sits in a web and detects the direction of prey through the vibrations that it senses, that organisms sit in a temporal web and that information about future states can be used to make beneficial decisions in the present.

The basic hypothesis is then that precognition is in essence a perceptual phenomenon. A relatively unknown perceptual phenomenon can provide a counterpoint. Discovered in the 19th century, *Haidinger's Brush* (Haidinger, 1844) is an optical phenomenon caused by the polarisation of light and a quirk of the structure of the human eye. Although not consistent across all people and the effect is subtle, hence not being discovered before 1844, it has been shown that in the presence of polarised light, small coloured patterns can be observed in the optical field, such that the patterning can inform the direction of polarisation. Described as a figure of eight or a propeller shape, the pattern of Haidinger's Brush is not independent of the eye and cannot be photographed. Von Frisch's work in the 1960s has shown that bees readily make use of a related optical pattern found in the heavens during the day, and perhaps humans have used or do use the polarisation of light in the struggle for existence (for example our ancestors may have used Haidinger's brush as a means of navigating on cloudy days). Thus, in an attempt to reduce the contentiousness of the hypothesis, the preliminary position with which this investigation starts is that, like Haidinger's Brush, precognition is a perceptual ability that can be both subtle and useful.

Already a problem presents itself, in that there are a wide range of terms and jargon used both in the public and scientific domains, often with some confusion surrounding individual use. As such, it seems pertinent to define some of the major terms briefly, and then further on they may be fleshed out, where appropriate. The term *paranormal* is perhaps one of the most general of all that will be used. Although somewhat context-specific, it is generally used to refer to any area that falls outside of conventional scientific scope. The phrase "seeing into the future" describes a range of phenomena which have been identified throughout human history. *Premonitions* and *prophecy* are much older terms for what will be called, throughout most of this thesis, *precognition*. Precognition is most simply defined as the paranormal acquisition of information from the future. For example, a dream of a train crash which subsequently occurs *could* be due to precognition. Importantly, precognition implies a non-inferential process or at the very least a paranormally based inferential process.

An understanding of the historical development of parapsychology helps to comprehend the prevalent terminology. The beginning of parapsychology is often traced to the formation of the Society for Psychical Research in London in 1882 by Henry Sidgwick, Richard Hodgson, F. W. H. Myers, Edmund Gurney and others, to investigate survival of life-after-death, among other questions. (The American SPR was founded in 1885 by William James at Harvard.) Established as a scientific society whose aim was to investigate the phenomena associated with Spiritualism, it is today widely acknowledged that the influential founders of the society expected to establish the reality of life-after-death and as such were looking to confirm what they already believed (Inglis, 1977; Webb, 1974). A statement regarding the aims of the society was drawn up,

From the recorded testimony of many competent witnesses, past and present, including observations accurately made by scientific men of eminence of various countries, there appears to be, among much illusion and deception, an important body of remarkable phenomena which are *prima facie* inexplicable on any generally recognized hypothesis, and which, if incontestably established, would be of the highest possible value. (Lodge, 1909, p. 9)

The success of psychical research was not as great as its promise and, despite a host of eminent Victorian intellectuals applying their skills to the questions, little credible science was advanced before the evolution of psychical research into the more rigid and scientific

Table 1.1: One historical proposed distinction of parapsychological phenomena

Psi Type	Past	Present	Future
Mental (ESP)	Retrocognition	Telepathy Clairvoyance Apparent Precognition	True Precognition
Physical (PK)	RetroPK	MacroPK MicroPK	PrePK

discipline of parapsychology. At Duke University, the father of modern parapsychology, plant physiologist turned experimental parapsychologist, J. B. Rhine, and his wife Louisa, developed the scientific investigation of the putative extrasensory perceptual abilities of humans.¹ It was here that terms such as *extrasensory perception* (ESP), and *psychokinesis* (PK) were coined.

Table 1.1 shows a general historical proposed categorisation of parapsychological phenomena. Somewhat unhappy with the value-laden terms that promulgated in the literature, Thouless and Wiesner (1947) were the first to suggest that the 23rd letter of the Greek alphabet “psi” (Ψ , not short for psychic) be used to denote paranormal phenomena.² *Psi* is the term favoured to cover *all* phenomena in the table. The first major distinction is between the wholly mental (extrasensory perception) and the wholly physical (psychokinesis). ESP is generally divided into telepathy, which involves information transfer between minds, and clairvoyance, where the mind appears to have discovered information about something independently of another person knowing it. (It is frequently noted that ESP is defined *negatively*, and that researchers attempt to eliminate the influence of *normal* sources of information, in the hope that any remaining evidence of information transfer *is* evidence for ESP.) Psychokinesis (which we will generally ignore given the specific emphasis on precognition) can itself be sub-divided into large-scale demonstrations, dubbed macroPK (levitating a piece of furniture, or bending a spoon) and small-scale demonstrations, dubbed microPK (influencing radioactive breakdown of substances to bias random number generators). It should be noted that this particular distinction is far less common than that used for the mental phenomena.

The final distinction is made according to the temporal aspect of the phenomena: whether it is proposed to work in the past, the present, or the future. It should be noted that apparent precognition refers to those occasions where information gathered in the present could be misconstrued as having been got from the future. For example if, in a precognition experiment, the participant is to guess the order of a pack of playing cards, it would be apparent precognition if the pack is shuffled before the participant guesses, and a test of true precognition if the pack is shuffled *after* the participant has made their guesses. Obviously the former might be construed as a test of clairvoyance, rather than precognition. It should also be noted that tests of true precognition and retrocognition can be tests of telepathy or clairvoyance. The same point is true for retroPK in the physical side of the table, which might involve either microPK or macroPK. (Although the term prePK has been used to denote tests of future psychokinesis, such tests are rare in parapsychology and this is an illustrative neologism, rather than a standard parapsychological term.) Finally, it should be noted that some have suggested that precognition is best explained by psychokinesis, that is, rather than know the future before

¹Martin Gardner reports that J. B. Rhine left botany for parapsychology after attending a lecture on Spiritualism by Conan Doyle. Rhine has described Gardner as a ‘professional denigrator’.

²Thouless and Wiesner (1947) also advocated the use of the terms psi gamma (Ψ_γ) and psi kappa (Ψ_κ) to differentiate the wholly mental and physical division of psi (ESP and PK respectively), though only *psi* has established itself as a commonly used term within the field.

it has happened, some argue that the more parsimonious explanation is that the future is created from present expectations (Rhine, 1977). We return to some of these issues in a later section.

There is one reason for concentrating primarily on precognition. Though it will become apparent that there is a range of evidence which makes precognition a compelling research question anyway, there is also the fact that precognition is perhaps the tightest of all the parapsychological phenomena. Because often the “guess” is recorded before the target has even been generated, with the correct methodology you can potentially develop an empirical regime that reaches a high standard of stringent control. This is obviously important if we discover any evidence for precognition during the empirical phase. If there are not other possible reasons for a deviation from chance, then the possibility must be that evidence for precognition has been found.

Finally, it is worth indicating the general direction that this introductory chapter will take. Firstly, the putative evidence for ostensible precognition will be examined. This includes a variety of anecdotal collections and experimental evidence. This evidence raises a number of questions which are advanced, and potential answers examined. After this examination of the case for precognition, the chapter changes tack, and one of the central questions of this thesis is raised. That is, whether there is any evidence that psi has evolved as a consequence of Darwinian adaptation. This question is tackled in a three-fold manner: firstly some of the evidence for animal psi is examined, and this is followed by a discussion of altered states of consciousness. Both sets of evidence argue that psi has a biological basis and is shared amongst different species. Finally, some evolutionary theories of psi are critically examined, with the final summation that there is some good evidence, both for the existence of precognition as a veridical phenomenon, but also that such a phenomenon may well have evolved through Darwinian adaptation to retrocausal influences. This final conclusion sets up the rest of the experimental chapters, where attempts are made to build on these conclusions, and to determine if they are correct or helpful when investigating psi, and specifically precognition.

1.2 The Putative Evidence for Ostensible Precognition

1.2.1 The Sporadic Cases

The sporadic cases of precognition (called the “non-quotitative cases” by Broad, 1968) and the associated spontaneous case collections, represent the very bedrock of parapsychology and it is important to understand how very influential the collected anecdotal experiences of a wide range of men, women and children were in the development of the area. The formal start of collecting cases of apparent paranormal experiences can be traced to the founding of the Society for Psychological Research in London in 1882.³

The SPR’s first great contribution to the fledgling field of parapsychology was the publication of *Phantasms of the Living* (Gurney, Myers, & Podmore, 1918), based on an extensive search and collection of real time ESP experiences, which specifically *excluded* any premonitions or

³An earlier draft included a small section on the historical evidence for precognition from ancient times, of which there is a great deal. Cicero’s (1887) *De Divinatione* is an excellent historical account of a Roman’s opinion of the supernatural, and one of oldest remaining sceptical treatises. Dodds (1971) provides a review of much of the evidence that would have been presented, along with Inglis (1977), who provides a book-length history of the paranormal from ancient times until the founding of the SPR.

precognitions. Myers noted that the basis for *Phantasms of the Living* was to establish that communication could occur free from the limitations of space, before there could be any discussion of temporally displaced experiences. *Phantasms of the Living* included some 359 cases of telepathy and those were the best drawn from a much larger pool. The general conclusions were positive, with the authors believing the phenomena well demonstrated.

Precognition was not the emphasis of *Phantasms* and yet it does contain an argument which is worth considering: the so-called faggot hypothesis, which has no doubt influenced many of the case-collections and the positive assessments of those who have investigated them. This hypothesis essentially boils down to the idea that if rejection of the account as having a paranormal explanation, itself requires an even more unlikely explanation (for example collusion between many trusted individuals), then it adds to the cumulative case being made for the reality of those paranormal experiences (a faggot of cases). No case alone, however strong, will provide the indisputable evidence that paranormal claims are evidence of a previously ignored ability of our species. But perhaps, given enough impressive elements, there is an argument to be made that a strong faggot of cases could be of some considerable worth.

Two years after the publication of *Phantasms of the Living*, Mrs Henry Sidgwick, agreeing the case had been made for contemporaneous phenomena, published the first survey of precognitive cases from the files of the society. Sidgwick (1888) was seen as a continuation of the work begun in *Phantasms*, where she reviewed some 38 first-hand cases of premonitions, of which twenty-four were dreams. This is a much smaller number than were available to compile *Phantasms*. Sidgwick began her review by adding the caveat that despite the cases she reviewed being the very best, none were impressive enough to judge the case for precognition made. This was a very prudent beginning, although Sidgwick is much more positive about contemporaneous ESP. Sidgwick also raised the initial concern that over a fifth of the cases she reviewed could easily be examples of clairvoyance, that is, that the events could have been logically inferred if all of the circumstances leading up to it could be known.

Sidgwick (1888) identified some potential factors to bear in mind when examining the spontaneous cases for precognition. These included the fact that some premonitions may bring about their own fulfilment, so that a dream of an accident may cause apprehension in the percipient, who might subsequently have an accident caused to some degree by the 'premonition'. Sidgwick also raised the concern that whilst the evidence for contemporaneous ESP has a sense of urgency or immediacy about it, premonitions rarely specify an exact time for fulfilment, which leaves the experience open to interpretation.

Sidgwick (1888) claimed death was a common element to the experiences that she investigated and that it was a widely held belief that an apparition of a living person was a warning of a disaster or imminent death. Other signs or premonitions surrounding the death of a loved one are found within Sidgwick's collection. Along with the reports of black dogs appearing before a death, there is a lot of folklore which associates screams as portents of death (the banshee), or the unusual flocking of birds. Sidgwick did not argue for the reality of any supernatural events associated with the portents of death, but suggested that such things might represent a shared symbol or hallucination which has become associated with this aspect.

With regard to mediumistic prophecy (precognitive information provided by a spiritualist medium), Sidgwick (1888) reported that it is difficult to assess mediumistic communication because of the unreliable nature of memory, arguing that positive content will be remembered

and any incorrect statements quickly forgotten. This then reduces the worth of any mediumistic prophecy. Dreams too are seen as unreliable to some degree and Sidgwick raised the very reasonable normal explanation that there are so many dreams dreamt every night that some of them must come true. However, Sidgwick argued that there are a number of cases which represent a more striking occurrence than merely being a dream which coincidentally matches a future event. These more specific premonitions are at least suggestive of paranormal functioning. Sidgwick also argued that there is a notable weakness in the dream evidence, in that human memory for dreams is less vivid and trustworthy than that for waking experiences and that much interpretation can be read back into dreams so that they have much more detail than was initially experienced as part of the dream. This source of weakness is to some degree diminished by recording the dream prior to its fulfilment, or at the very least telling it to another person, and Sidgwick presented a number of cases which fit this particular standard.

Sidgwick (1888) also identified another weakness in the dream reports. Namely that remembering that a dream has been experienced is very difficult to be certain of, and it is possible that somebody could have a dream where their house catches fire every night, then forget it, and only remember it the one time that their house really does catch fire. If all of the dreams had been known about, a judgement could be made about whether the dream premonition of the house catching fire represented a unique event or was one of many dreams the percipient had had.

Although uncommon in Sidgwick (1888), some of the most striking premonitions often involve action *preventing* a potentially fatal future event (as distinct from portents of natural death). Much more will be discussed later about this particular aspect of the precognition literature but it is worth highlighting the seemingly fortuitous nature of such reports, if they occurred exactly as described. There is an obvious question that has been asked in relation to these kinds of events. If the initial event depicted in the dream does not occur (e.g. an injury caused by an accident), what provokes this particular sequence of events, where the disaster is averted? This is not a question which has been answered successfully. Sidgwick noted that everyone must die and therefore a mere dream of a death has no particular evidential value. It is the details which are important, dates especially so.

A common type of premonition involved dreams which correctly identified the winner of a race, usually involving horses. Sidgwick (1888) claimed that this is not particularly surprising given that such races are probably frequently a subject of dreams amongst those who have an interest in them and therefore it is not unexpected that some dreams come true. Such dreams are found in later collections and it has been suggested indicate that desire may play a role in the precognitive process. A further class of premonitory dreams which Sidgwick identified are the so-called symbolic dreams. There are two classes of symbolic dreams: those that are evidently not real but have an underlying theme which suggests the premonition and those that are completely symbolic in form. The first type of symbolic dream is a kind of distorted reality and is interpreted thus. The purely symbolic dream follows a convention which is imposed by tradition. Sidgwick argued that it is not important what kind of symbolism is employed, merely that it is recorded before the event which it is claimed it pre-dates. Obviously assuming any dream contains symbolism is a matter of opinion and it seems very easy to read into something as odd as a dream, symbolism which can be used to identify it as a premonition of a later event. As such, symbolic dreams lack much of the specificity which is seen with the other accounts and so represent the weakest evidential category for the spontaneous cases. Broad (1935) notes that the prevalence of dream reports in Sidgwick's collection does represent

a 'serious weakness'.

Sidgwick (1888) also noted the frailty of human memory and the weakness of testimony. In fact, she concluded that, given the evidence and the many and varied potential normal influencing factors, the possibility of a paranormal precognitive ability has nowhere near the standard of evidence required to be accepted, even as a 'working hypothesis', by the scientific world. However, there are some cases which are very impressive and it is those that suggest further explanation is required. Finally, Sidgwick concluded that evidence showed that people rightly or wrongly believed that premonitions happened and happened to them.

Saltmarsh (1934) followed up Sidgwick (1888) with his own survey of some 349 precognitive cases that had thus far been published by the SPR. Saltmarsh regarded the first question to be asked to be one of whether there is any evidence for true, ostensibly paranormal, premonitions. He categorised each of the 349 cases he could find in the SPR annals, according to their quality and the nature of the experience, be it dream or waking impression. Saltmarsh noted that a great many of the reports involved some degree of dissociation which he suggested implied that premonitions involve a subliminal part of the mind. As Sidgwick did, Saltmarsh noted a prevalence of dream reports within the collection and that the dreams provided some of the best cases. Saltmarsh sorts the cases into impressive and ordinary, dismissing the poorest accounts outright. Some 99 out of the 281 remaining cases were associated with death: with 72 natural and 27 accidental. Of this selection, a far higher number of death accounts were associated with hallucinations than any other class of report. In regard to dream symbolism, Saltmarsh remarked at a distinct lack of such reports given the strong influence he supposed it played in the creation of dreams. Saltmarsh claimed that the review was undertaken to discover potential hypotheses for any supposed precognitive phenomena. He identified four potential explanations for some of the accounts, (a) telepathy, (b) auto-suggestion, (c) subconscious knowledge and (d) hyperaesthesia.

The telepathic explanation is very similar to the clairvoyance hypothesis as proposed by Sidgwick (1888), whereby someone with knowledge of a forthcoming event telepathically conveys the information to another party. He concluded that the telepathic hypothesis certainly explains some of the accounts, although not others. Saltmarsh (1934) defined auto-suggestion as a bringing about of circumstances through a certain thought or action and gave the example of a person hallucinating that they will die at a specific hour and subsequently dying at that future time. He suggested that it was certainly plausible that it was the hallucination which brought about the expectation of death, which caused an increase in anxiety, which precipitated the death. The subconscious knowledge hypothesis is again similar to the clairvoyance hypothesis except that normal information is gleaned on a subliminal level, which informs the premonition. This hypothesis is based on the assumption that the subliminal mind somehow has greater levels of inference than the supraliminal mind, whatever those terms might actually mean. Examples of subconscious knowledge might include a premonition of a person's future physical state. For example, a person might dream that they are diagnosed with cancer and such an experience might be caused by the subliminal inferences gleaned from internal homeostatic monitors. The last of Saltmarsh's hypotheses, hyperaesthesia, refers to an acute sensitivity to the surrounding environment. Specifically, the premonition is based on normal information which has been acquired through an unusual sensitivity to a stimulus. The example is given of a woman hallucinating a fire in a house nearby, and then minutes later consciously registering the first sounds and smells of a fire.

Thus Saltmarsh (1934) identified the potential factors which might influence conclusions

about premonitions and is left with some ninety-four cases which seemingly require further explanation than those already suggested. Saltmarsh noted that there is one other major explanation which might account for any or all of the cases under review. Chance coincidence is held up as a majority view of the uninstructed and to those that have not made an unbiased assessment of the evidence. Like Aristotle, Saltmarsh considered whether, given the fact that billions dream every night, it is surprising that some dreams seem to match an event some time at a future date. However, Saltmarsh dismissed this chance coincidence explanation, given that millions may dream every night, but only a small number are actually remembered and that a smaller number still might be seen as having paranormal content. Saltmarsh called on the specificity of the premonitions as a demonstration of the paranormal content of the accounts. For him, it was not so much that people have dreams which seem to be about the future which is impressive, but that there are too many details and facts which just cannot be explained with the chance coincidence explanation. This is essentially the 'faggot hypothesis' described above.

Saltmarsh (1934) proposed that a central fact, common to all cases of precognition, is that when observing any event at some point it must enter consciousness. A cognitive relationship is set up between the event and the mind. In normal recognition, this is fixed within the normal temporal direction, but in precognition this relationship is between the mind and the future. This particular theory is deterministic and Saltmarsh argued that the future exists now and is in some way knowable by consciousness. Saltmarsh, however, did not argue that the future is entirely predetermined, but that some events can be precognised. Importantly, he suggested that just as some events can be predetermined, that it may, in fact, be an indispensable requirement for a precognition that the event is completely determined. Only those portions of the future which are entirely fixed could be precognised. Saltmarsh (1934, p. 74) made the pertinent conclusion that "it is clear that the crux of the whole thing is Time."

Saltmarsh (1934) distinguished between two types of time; the time of mathematics and physics, and so-called psychological time, the time we experience for example, passing in a dream. Saltmarsh argued that psychological time has a kind of dimension or extension. The present moment does not exist for no-time, but has some kind of duration. Perception and consciousness also have some kind of duration, otherwise, he argued, there could be no consciousness. If consciousness did not have duration, it would be separated or punctuated by gaps in consciousness, which we would surely feel. The idea that the psychological present moment has duration is called by Saltmarsh the specious present and is based on ideas developed by Henri Bergson and William James, where the specious present has a longer duration than the instantaneous present of the physicist and mathematician. He therefore asked the question: does the specious present include the past alone, the future alone, or both past and future? Saltmarsh's complex theory can be summarised in proposing a one-to-one correlation between what he calls the 'supraliminal mind' and the 'subliminal mind'. There is an overlap between the specious presents whereby information which exists in the specious present of the subliminal mind may still be in the future of the supraliminal mind, fixed in a temporal order, but because the information is within the mind system, it can be perceived by the supraliminal mind.

Finally, Saltmarsh (1934) considered the very nature of existence. If all events are determined, then there is no such thing as free will and therefore he believed that time is an illusion. This would mean that all events exist at the same time and that we move around picking up events as they move into our specious present. He found this particular hypothesis lacking and opted instead for a belief in free will and the conclusion that the world is not entirely

determined. Saltmarsh believed that free will has duration, otherwise it would not exist, and that it is able to influence the course of events. To Saltmarsh, time is plastic, with some chains of events that are completely determined and others less so. With this particular theory, Saltmarsh concluded that premonitions only occur for determined events and he regarded premonitions of non-determined events as impossible. How this relates to premonitions where the future event is averted, for example, those premonitions where the percipient changes their intended action and thus survives (where such an action would have likely resulted in death), is hard to say. Saltmarsh's paper represented a ground-breaking attempt to assess the evidence and garner some conclusions regarding the evidence. Was Saltmarsh justified in his conclusions? Certainly he was far more positive than Sidgwick (1888) in his assessment of much of the same evidence.

Another major contribution relevant to the topic of precognition is Dunne's (1927) book *An Experiment with Time*, which aroused the public's interest in the topic of precognition like no other book before it. Unlike the other case collections discussed in this section, that are composed of different accounts suggestive of precognitive experiences, Dunne's contribution was a review and analysis of a collection of his *own* dreams, collected over an extensive period. Dunne not only believed that he commonly had precognitive dreams, but that most people probably experience premonitions during sleep, but forget them, or do not recognise them for what they are. Dunne's experiences ran from the mundane, dreaming of an unexpected trip he was to take in the future, to the more interesting. One of the aspects that raised Dunne's story in the mind of the public at large was his claim that he often dreamed the winner of particular horse races (a kind of get-rich-quick scheme). As Sidgwick (1888) noted, such dreams were not uncommon, and have continued to be reported in the literature (for example, Godley, 1950).

The public were not only fascinated with the examples provided by Dunne (1927), but with his discussion of the problem and his proposed explanation. Like Saltmarsh (1934), for Dunne the nature of time was important to the question of precognition. He argued that time can be conceived of as a length travelled by a moving 'now'. Movement involves both time and space, but since he begins from a position where time is considered 'space', he argued that this requires another kind of time (calling it 'Time-2'). Ordinary time is conceived as a fourth dimension of space, along which moves a three-dimensional field of presentation, which is the 'now' point. Dunne next argued that this Time-2 must also have some duration or length and so argues that it must be travelling along another type of time, *ad infinitum*. There is thus an infinite series of Times, and an infinite series of now points, each of which has one dimension more than the last. Obviously Dunne's starting position seems extremely tenuous. There seems little reason in proposing an infinite number of time dimensions, merely as a need to explain premonitions.

Dunne (1927) proposed that each infinite now point is associated with an infinite number of self-conscious observers, each of whom is conscious of himself in observing. Observer 1 (who is the self) is a three-dimensional field moving along Time-1. Observer 2 is the four-dimensional field that moves along Time-2, and so on, *ad infinitum*. Dunne argued that the future is determined and that each observer or self is merely an automaton moving along their already determined paths. Premonitions are therefore explained because any event in the future of Observer 1 is already present to an Observer 2. Dunne argued that only an ultimate observer, whatever that might be, can intervene and has the capacity for consciousness and free will. This higher-order observer can only intervene at Time-1 and nowhere else. As such, this ultimate observer usually attends to Time-1 and Dunne believed it is difficult to break

this association, although he proposed that sleep might reduce the connection. When this occurs, Observer 2's perspective is taken, and a wider view becomes available. The difficulty in describing the four-dimensional perspective of Observer-2 in the three-dimensional perspective of Observer-1 is what Dunne suggested as the cause of the fantastic or unrealistic nature of some premonitions.

Dunne's (1927) theory was taken with some seriousness at the time and received wide public attention. Yet there seems little justification for the proposal of an infinite regress of time and consciousness and so elaborate a theory, when so little is actually known. Broad (1935) criticised Dunne's use of the term 'observer at infinity' which appeared in early editions of his book because, as Broad pointed out, a hypothesis which proposes an infinite number of observers cannot have a final or last observer. Ultimately, as remarkably influential as Dunne's book was, it has been little supported, although it is still often hailed as the best evidence for precognition. We return to the topic of Dunne's premonitions in Chapter 5.

Rhine (1954) published a more conventional case collection review of all of the cases sent into the Parapsychology Laboratory at Duke University, and noted that precognitive cases seemed to make up a much larger percentage than had been the case in the earlier published collections.⁴ Rhine argued that the increase in reports of precognition seemed to be a trend which it was suggested reflected an increased perception of respect regarding dream reports, where Sidgwick, for example, considered the dream evidence as evidentially the weakest. Rhine asked the philosophical question as to whether precognition is merely a form of ESP that is independent of time, or whether premonitions are a distinct phenomenon which reflects a fundamental difference between the processes?

Rhine (1954) reviewed some 1,324 precognitive ESP cases and compared them to 1,966 contemporaneous ESP cases. Obviously, this then is a large-scale study, dwarfing all of those published before. The proportion of dreams to waking experiences in the precognitive group was much greater than it was in the contemporaneous one; 68 percent of the precognitive cases were dreams and waking experiences made up only 32 percent, while in the contemporaneous group the percentages were reversed, being 35 percent and 65 percent respectively. Rhine noted that most of the psi experiences reported seem to take one of four psychological forms:

intuitive an experience perceived without logical or sensory antecedent.

hallucinatory a waking experience where the premonition is experienced by an effect on the sensory system.

unrealistic dreams symbolic or at the least not directly representing future events.

realistic dreams accurate depictions of future events.

These categories were similar to those that both Sidgwick (1888) and Saltmarsh (1934) developed. Realistic dreams made up the majority of the cases in the precognitive group (60 percent), followed by unrealistic dreaming and intuitive accounts (19 percent) and hallucinatory experiences were reported the least (6 percent). Overall only 36 percent of the cases carried conviction, that is the percipient believed it portrayed a future state. However it is worth noting that Rhine (1954) reasoned that premonitions seemed to be believed less often than contemporaneous reports, because dreams carry less conviction (19 percent compared to 37 percent).

⁴This is Louisa, wife to J. B. Rhine.

Rhine (1954) concluded that precognition is real and that it appears to be the same as ESP. She proposed that recall of the future may not be substantially different from recall of the past and this is borne out to some degree by the predominance of realistic experiences, reflecting an accurate premonition of a future event, much like a memory is a record of a previous event. Rhine also believed that just as some people have a photographic memory, some individuals may have photographic premonitions. In fact, many individuals in reporting their precognitive dreams have marvelled at the fact that the experience was just like “remembering” the future.

Rhine (1955) followed up her review of the precognitive accounts with a further analysis of another selection of similar reports with the purpose of discussing implications for intervention. As has already been touched on briefly, authors dealing with precognition have often raised issues regarding the implications for free will and volitional freedom. Rhine specifically argued that, at least on the face of it, the fact that there are accounts suggesting accurate foreknowledge of a future event suggests that to some extent the future must somehow exist. She likens the situation to a roll of photographic film, where the future is fixed, merely waiting to be unwound. However, Rhine recognised that if such were the case, then volitional freedom would be an illusion. However, she also believed there to be a case that occurrences of premonitions do not necessarily imply an absence of free will, depending on the nature of time, the nature of freedom and the nature of precognition.

To help shed some light on the issue, Rhine (1954) examined a large number of cases and identified some 433 cases which seemed to show conviction. Out of the 433, some 191 cases reported that the percipient had tried to prevent the event that had been foreseen. Rhine reported that of the 191 cases, for every unsuccessful attempt to stop fulfilment of the precognition, there were two successful accounts. Rhine also argued that cases with highly dramatic character should be regarded with suspicion. The very fact that accidents may be remembered more often than simple banal events raises significant doubt regarding certain accounts. Rhine herself concluded that if there was evidence for an ability to prevent events in the future, that it was seemingly modest and quite unpredictable. Rhine notes that there does appear to be a class of reports where a threat is avoided, just as information from the present might initiate a normal threat response and that future information sometimes causes people to act in a beneficial way.

Out of all of the 433 cases Rhine (1954) considered, only 9 met the high standard of evidence that Rhine had demanded if she was to make any worthwhile assessment of the phenomenon. However, even of those nine cases which Rhine believed were the best evidence for premonitions, there was still the possibility that errors in memory could embellish any of the accounts to render them more interesting and dramatic than they really were. Finally though, Rhine identified three, which were the best precognitive accounts she could find, one of which is reproduced:

A mother had a waking picture of her eldest son, Herbert, dead in the bath tub. It haunted her so that she made a special point of listening that nothing went wrong, but she did not tell him her impression although she told her younger son, Peter. After a couple of years Herbert went away and when he came home for a holiday she still remembered it.

One evening on this visit she heard him whistling and singing in the bath tub. She was dressed to go out but could not leave. After a while she heard the water running out but did not hear him singing so she opened the door, and there he lay exactly as she had seen him two years before. There was a gas heater and the window was closed and he had apparently been overcome by fumes. She immediately opened the door and windows and called the doctor and he was

revived. If she had not been there, he doubtless would have died. (Rhine, 1955, p. 27)

Rhine (1954) believed this was an interesting premonition which highlighted the importance of free will. Had the future been determined, the son would have died. Had the mother not attended to her son based on the feelings of concern, the son would have died. However even with this, one of the best cases of precognition in the spontaneous cases, there are a number of normal explanations which might explain the account more parsimoniously than the precognition hypothesis. Significantly, the mother reports that she dreamt about her son dying in a bath tub, but does not report how often she had anxiety dreams about the health of her children. The fact that the mother had been so concerned as to investigate her son's condition, implies not only that she was a concerned mother, and therefore concentrated her attention more readily to attend to her son, but also there is the possibility that the lack of sounds or smells from the bathroom might have initiated real concern based on her general anxiety over his safety.

The cases presented in Rhine (1954) are interesting in providing an overall assessment of the spontaneous cases. Rhine's positive inclinations towards paranormal hypotheses aside, she identified only three cases as providing impressive evidence on which to begin to consider the problem of volitional freedom. This seems a very small selection from which to draw any great conclusions except maybe that the question is left begging. This is Rhine's conclusion also and it is left for future researchers to continue their investigation of the problem.

There have been more general case collections published over the years (e.g. Nicol, 1961; Prasad & Stevenson, 1968) along with a great many single premonitions which feature regularly in the annals of psychical research (e.g. Bender, 1966). However, it is a fair conclusion that many premonitions are linked to death and thus we turn to examination of a sub-set of cases, specifically concerning precognition of disasters.

1.2.1.1 Premonitions of Disasters

The word 'disaster' is derived from the Latin for *unfavourable aspects of the stars* and related to the portents found in nature that ancient societies saw in all manner of natural phenomena, pertaining to future dangers.⁵ Many authors have noted the prevalence of premonitions associated with death and crisis (e.g. Saltmarsh, 1934; Sidgwick, 1888; Murphy, 1960; Mundle, 1964) and perhaps it is not surprising that some of the most impressive premonitions are associated with disasters. Three sets of cases in particular will be examined: those associated with the sinking of the *Titanic*, those associated with the Aberfan coal-tip disaster, and the case of the exploding church.

The Sinking of the *Titanic*

Parapsychologist Ian Stevenson published a review and examination of paranormal experiences associated with the sinking of the *Titanic* (Stevenson, 1960). Stevenson reported a great fascination with the spontaneous cases that he had collected and that the rationale for collecting such material was very much influenced by the work of Cox (1956a, 1956b) which will be discussed in greater detail in Chapter 6. Cox reported that from his review of train accidents,

⁵The etymology of some related words: 'premonition' comes from the Latin, *prae* meaning 'before', and *monere* meaning 'warn'. The word 'hazard' is derived from the Arabic term *al zar*, meaning 'the dice'; 'chance' is derived from the Latin *cadentia*, referring to the fall of a die. Finally, 'random' comes from the French *randir*, which means to run fast or gallop, though Skeat (1882) suggests that the origin is actually derived from the irregular and violent motion of a river, full-to-the-brim.

he had discovered that fewer passengers travelled on the day of an accident than were travelling at other times. For Cox, this was evidence of some kind of death-avoidance behaviour. Whilst Cox regarded conducting a similar review of ship sinkings as being much more difficult (due to increased demand, etc.), Stevenson argued in spite of this particular concern that the *Titanic* sinking was an event which captured public interest that still exists today, and that he had frequently come across associated paranormal accounts which demanded reporting. The cases also highlighted the possible importance of unconscious precognitions, those hunches or feelings which cause a person to do something for no conscious reason.

When the *Titanic* was constructed in 1911 it was the largest ship in the world (882.5 feet long and weighing some 46,328 gross tons). It employed an unusual design, where fifteen watertight bulkheads could be closed almost instantly to seal off compartments of the ship. Unfortunately, these bulkheads extended only partway up the hull of the ship, a design feature that would come to have serious repercussions in the actual disaster. The ship's furnishings were luxurious and top-of-the-range, with the first class passengers travelling in extreme comfort. When finished, the *Titanic* sailed from Belfast to Southampton and, on April 10th 1912, set off for New York. She stopped briefly in Cherbourg and Queenstown. At 11.40 pm on the night of April 14th, under full speed, she struck an iceberg which ripped a gash in five of her forward watertight compartments. Water rapidly rushed into the damaged front compartments and tipped the bow down so that seawater ran over the top of the fifth bulkhead and flowed into the next compartment, and so on. The ship did not capsize but tilted down so that the stern gradually rose while the bow sank, and it finally descended into the ocean at a steep angle. The ship sank in less than three hours at 2:20 am on April 15th. The *Titanic* carried far too few lifeboats and many of these were not filled when they were lowered and rowed away. She carried about 2,207 persons, including crew; 705 persons survived the sinking to be rescued by the liner *Carpathia* which arrived at the scene too late to save more. About 1,502 persons lost their lives, although an accurate total will never be known.

Stevenson (1960) argued that the general conviction of unsinkability, extending almost up to the actual moment of sinking, has relevance to the reported parapsychological experiences, as he believed it unlikely, although not impossible, that any normal inferences could have led to a premonition. This is in contrast to the sinking of another vessel, the *Lusitania* about which only a few apparently parapsychological experiences have also been recorded. The *Lusitania* was sunk in 1915 by torpedoes from a German submarine. The Germans not only issued general announcements of their submarine warfare against merchant ships but strongly hinted in the newspapers of New York their intention of sinking the *Lusitania*. Obvious expectations of danger hung around the *Lusitania*'s last voyage: complete confidence accompanied that of the *Titanic*'s. Stevenson argued, reasonably, that there was no doubt an expectation of doom that had accompanied the passengers of the *Lusitania*, but that the passengers of the *Titanic* should have been blissfully unaware of the unlikely fate that was to befall their particular journey.

One of the strongest examples in Stevenson's collection is reproduced below,

Mr. J. Cannon Middleton, an English businessman, booked passage on the *Titanic* on the 23rd of March. About a week later i.e. ten days before the sailing date, he dreamt "that I saw her (the *Titanic*) floating on the sea, keel upwards and her passengers and crew swimming around her." The following night he experienced the same dream. These dreams made him "uncomfortable," and he was subsequently "most depressed and even despondent." Mr. Middleton did not cancel his passage until about four days after the first dream. He did so then because he had received

a cable from the United States telling him that for business reasons he should postpone his sailing for a few days. After he cancelled his ticket, Mr. Middleton told members of his family and friends about his dream prior to the actual sailing and sinking of the ship. They subsequently testified that he had done this. Two of the persons to whom Mr. Middleton had told his dreams mentioned in their reports that Mr. Middleton had told them (in addition to the details given above) that in his dream he himself "seemed to be floating in the air just above the wreck." Mrs. Middleton stated that her husband "never dreams" and had certainly never had a dream of this kind before. (Stevenson, 1960, p. 157)

There are a number of conclusions that can be drawn from Stevenson (1960). Certainly some of the cases that Stevenson presents are dramatic and seemingly inexplicable, if they occurred in reality in the ways described. However, out of a large number of people involved in some way with the disaster, only two of the percipients had any personal ties to the *Titanic* and this is hard to reconcile with the idea of psi being a kind of precognitive death-avoidance behaviour. For those cases which are not connected, there are a range of explanations as to why they may have an experience which is labelled at some time afterwards as paranormal and being related to the sinking of the *Titanic*.

Errors in recall *might* lead to people identifying a particular dream as precognitive. For example, one might frequently dream of ship sinkings (or horse races) and forget the dreams, only to be reminded that they had experienced a seemingly precognitive dream when prompted by the learning of a real-life incident. It is also possible that the event is learned of, but not consciously understood (hearing people talking about the event but not recognising to what they are directly referring) and that this prompts a subsequent dream which then appears to be precognitive, after the percipient learns of the event proper. There is a great deal of psychological research, such as the very well known work by Loftus (1979), which highlights the unreliable nature of eyewitness testimony and for in-depth review of the research with specific emphasis on anomalous experiences see French (2003) and French and Wilson (2006).

Does the fact that only two people were connected with the sinking undermine the basic theoretical position that psi might be needs-based? Cox (1956a) argued that ship sinking data is hard to investigate for a variety of reasons, the main one being that ships are usually over-subscribed. In the case of the *Titanic*, this could well be a distinct possibility. As the largest, most luxurious, and fastest ship of her type in the world, the maiden voyage of the *Titanic* was a once in a life-time opportunity. It is plausible that once people had chosen to board the ship, and had purchased a ticket, that it would take a considerable feeling of dread to dissuade the person from boarding and thereby prevent their future death. Thus although precognition may be needs-based, other psychological needs may overwhelm or otherwise conflict, meaning that precognitive information may only play a small part in critical decision making, be it conscious or unconscious. Stevenson (1960) believed there is another possibility as to why there is a lack of premonitions by those connected to the events (i.e. by loved ones): that during the emergency, thoughts are not focused on family and friends (as sentiment might suggest) but that all concern is directed to immediate affairs.

Stevenson (1960) concluded that the anecdotal accounts do provide evidence for a kind of precognitive death-avoidance which sometimes is acted upon and sometimes is ignored. He did not attempt a study of the cancellation of tickets along the lines of Cox's investigations of railway accidents for two reasons. Firstly, the *Titanic* made only one voyage and this provided no baseline for comparisons of bookings before and for the fatal voyage. Secondly, a clerk of the White Star Line acknowledged that many persons had refused to sail on the *Titanic* because

of a superstitious fear of maiden voyages. Fear of danger associated with a maiden voyage seems to be remarkably long-standing, and which may stem from the possibility that a ship that has not been sailed may not fare well once a journey has begun. Obviously in the case of open sea travel this can represent a very real danger. Although there may have been little rationale for fearing a maiden voyage on the *Titanic*, with the claims that it was unsinkable well known, there may still have been a large degree of apprehension associated with the ship's first sailing.

In a subsequent extension to his original paper, Stevenson (1965) added seven more paranormal experiences which seemed to be associated with the sinking of the *Titanic*. Stevenson found the anecdotes that he presented as being certainly suggestive of a precognitive effect. Is he justified in doing so? That some people would have found the maiden voyage of the *Titanic* something to fear means that for any report, however striking, there must be some doubt as to whether the apprehension was a manifestation of a widely held naval superstition.⁶ Therefore, despite Stevenson's conclusions regarding the *Titanic* reports, they do not provide strong evidence of precognitive death-avoidance behaviour, dramatic and compelling though they are. However, as part of a growing 'faggot' of evidence, there is enough suggestion of strength to pursue further investigation.

The Aberfan Disaster

Consultant psychiatrist J. C. Barker reported his review of premonitions associated with the Aberfan disaster (Barker, 1967b). The disaster occurred in October 1966, when a massive coal-tip slid down a mountain-side onto the Welsh village of Aberfan, killing some 144 people. Over 120 of the victims were school children killed inside the local primary school, which was destroyed by the avalanche. The disaster was unprecedented in Welsh history and was particularly tragic, given the large number of children who died. On visiting the area immediately after the disaster, Barker was inspired to imagine whether there might be people in the country who had received a forewarning of the disaster and so made an appeal in a national newspaper for any premonitions or precognitions that people might have had about the disaster. Barker claimed that there were some two-hundred reports of premonitions associated with the disaster, although he himself had access to 76 cases. Some 36 of the reports were from dreamers, the rest were claims of waking visions, or feelings of impending doom. Barker reported brief versions of 32 of the cases, for example (p. 171) 'Dreaming of screaming children buried by avalanche of coal in mining village. Woke up screaming', and gives three in detail, one of which is presented here:

She was an attractive dependable child, not given to imagination. A fortnight before the disaster she said to her mother. . . 'Mummy, I'm not afraid to die.' Her mother replied, 'Why do you talk of dying, and you so young; do you want a lollipop?' 'No', she said, 'but I shall be with Peter and June' (schoolmates). The day before the disaster she said to her mother, 'Mummy, let me tell you about my dream last night.' Her mother answered gently, 'Darling, I've no time now. Tell me again later.' The child replied, 'No Mummy, you must listen. I dreamt I went to school and there was no school there. Something black had come down all over it!' The next day off to school went her daughter as ever. In the communal grave she was buried with Peter on one side and June on the other.' . . .

⁶A pertinent experience is reproduced by N. N. Taleb in his book on randomness, *The Black Swan*, who quotes from the diary of E. J. Smith, the captain of the *Titanic*, "But in all my experience, I have never been in any accident. . . of any sort worth speaking about. I have seen but one vessel in distress in all my years at sea. I never saw a wreck and never have been wrecked nor was I ever in any predicament that threatened to end in disaster of any sort." (cited in Taleb, 2011, p. 44)

The day [she] went to school the clock stopped at 9.00 am. Had it not stopped, her mother would not have been late going to school and this is what saved [the mother's] life. Curiously [the father] had apparently always felt that his daughter would not live out her normal life span. (Barker, 1967b, p. 173)

This case, like so many others from the annals of psychical research, seemingly supports the proposition that some people are responsive to future events particularly those that involve death and disaster. However, Barker (1967b) reported a number of facts which make this particular case evidentially a lot weaker. Firstly and most importantly, the danger regarding the coal tip was well known in the village, with a number of villagers raising concern over the ever-present danger. This alone represents a reasonable normal explanation for any premonition regarding the disaster. *If* this concern was widely known (perhaps as a rumour), the fact that the coal tip was directly over the primary school would no doubt have directed concern to the school children. It is not surprising that these concerns made their way to the pupils and that anxiety over an ever-present danger might be reflected in anxiety dreams by people in the village, or those aware of the danger.

Moreover, as the disaster was caused by a coal-slippage, that is, the shifting of many tons of semi-loose material down a hill-side, it is possible that there were normal sensory precursors to the disaster, perhaps in the form of low-frequency sounds or vibrations (and thus perhaps a form of hyperaesthesia had informed percipients). Such noises would not have to be continuous or dramatic to potentially inform of some impending danger. Rainfall and high winds might also have impacted the tip and caused concern (consciously or unconsciously) to those who observed those circumstances and deduced a potentially dangerous outcome. Thus conscious dismissal that there were fears of a disaster before the event does not immediately rule out unconscious priming. However, there is also the added fact that the experience reproduced above was compiled by the local minister and merely signed by the parents. It is easy to see how the minister's prior beliefs about the supernatural could have influenced his retelling of the account, if not his desire to comfort the parents, and thus the extent to which the account is accurate is difficult to assess.

In assessing the accounts connected with the Aberfan disaster as a whole, Barker (1967b) noted that very few of the correspondents had any connection with Aberfan, nor did any of them have sufficient information to avert the disaster. For Barker, the principle difficulty is that the majority of premonitions are very vague and with so many humans dreaming every night, some dreams might, simply by chance, match a future event. The events of Aberfan encouraged Barker to set up a bureau to collect premonitions in the hope that cataloguing experiences might lead to potential disasters being averted (this discussion is continued in Chapter 5).

The Case of the Exploding Church

The number of premonitions reported by individuals directly associated with either the *Titanic* or Aberfan disaster were much lower than the actual number who lost their lives in those disasters. However, we conclude this section with an examination of a well documented case where, fortuitously, nobody died.

Edeal (1950), reporting in *Life* magazine, gives an account of an extraordinary sequence of events. On the evening of Wednesday March 1st 1950, a small local church in Beatrice, Nebraska, blew up. The walls exploded outwards and the roof caved in. On the evening of the explosion the town was in the grip of a particularly cold spell of weather and the local pastor had decided, fatefully, to light the church furnace early so as to provide warmth for

choir members who were to have their scheduled practice later in the evening.

When the church exploded the 15 people who made up the choir were supposed to be well into their rehearsal time and yet, in the aftermath of the explosion, it became apparent that *not a single person had been killed*. There was no single explanation for why the different choirists had not attended the meeting on time and probability expert Weaver (1963) in his book *Lady Luck: The Theory of Probability* used the Beatrice event as an example of a highly improbable happening. He calculated that the chance of all 15 choir members being late on that particular night was $\frac{1}{4}^{10}$, this number includes the fact that some of the choir members were dependent upon others, and puts the chance of the event at one million to one. Weaver notes that had all 15 had independent reasons for being late, the chance would have been nearer to a billion to one.

Although none of the choir members reported anything like a premonition or precognition associated with the potential disaster and Weaver uses the case as an example of an unlikely coincidence, parapsychologists such as Dossey (2008) cannot help but see this case as an example of unconscious non-intentional psi, manifested as hunches or feelings that prevent future fatality and as such this case fits in well as an example of spontaneous death-avoidance.

In regards to the evidence for premonitions of disasters, like those associated with the *Titanic* and the Aberfan disaster, the best cases seem interesting. However, the whole area of sporadic or spontaneous premonitions as evidence for a paranormal ability seems inconclusive. To agree with Sidgwick (1888) and to some extent with many others who have collected and examined claims of premonitions, people seem to believe that predicting the future is possible, but that the evidence is at best suggestive. If we refer back to the 'faggot' hypothesis, although each individual case does seem to be weaker than would be hoped for, together they do represent, if not strength in numbers, then a curious collection which may point to a theoretical underpinning. Importantly though, the spontaneous case reports can be seen as being the very impetus for conducting empirical research into the existence of precognition, without which there would be no reason to conduct experiments into something which goes against current understanding of the world. *If* premonitions do really occur, it would seem though that the main consensus would be that premonitions are associated with death or crisis.⁷

1.2.2 Empirical Tests of Precognition

Empirical testing for precognition began, like the experimental search for extrasensory perception, in the lab of J. B. Rhine at Duke University. Rhine had ignored the problem of precognition until he was satisfied that the case for contemporaneous paranormal phenomena was established (mainly because if evidence for telepathy was met with scepticism by conventional scientists, evidence that participants could successfully know the future would be even harder for them to accept). The first experimental test of precognition was conducted in 1933 but not reported by Rhine until 1938 (Rhine, 1938). In these early tests, the participant (Hubert Pearce, who had already been identified as a successful participant in earlier ESP research) had to guess the order of a pack of Zener cards (the ubiquitous cards bearing five symbols that seem strangely esoteric to the general public, but were chosen to be of equal visual weight) *before* they were to be shuffled. After the guesses were recorded, the pack was shuffled and the card order compared against the participant's guesses and scored for matches (or hits). As with

⁷We return again to premonitions of disasters and the work of the premonitions bureaux to avert disasters in Chapter 5.

the tests of contemporaneous ESP, it was reported that the number of hits was significantly in excess of what was expected by chance alone. Rhine (1977) reports that a member of staff at the Duke University Parapsychology Laboratory (Charles E. Stuart) thought it possible that some kind of clairvoyant effect might be influencing the order of the cards during the shuffle (known as the “psychic shuffle”) and thus the above-chance scoring might not have been due to precognition on the part of the participant, but the result of clairvoyance on the part of the shuffler (and at this stage there was substantial evidence for clairvoyance, as compared to the difficult problems posed by precognition).

This criticism led to the use of a mechanical shuffler that, it was argued, could not be the *source* of the effect (Rhine, 1941). With the advent of PK testing and results which supported the PK hypothesis, it became apparent that though the mechanical shuffler had been used to avoid the “psychic shuffle” it might be possible that participants’ PK was influencing the order of cards by interfering with the mechanical shuffler. (The PK research had already demonstrated that dice in a mechanical dice-throwing machine were seemingly influenced by the thoughts of the participant, e.g. Cox, 1951; Forwald, 1952). The PK research made it difficult to determine whether the findings in the precognition tests were the result of information transfer from the future. Again, given the contemporaneous nature of the PK influence, the precognition hypothesis appeared to be the less likely explanation. Given this problem in determining whether actual precognition has occurred, researchers developed a novel method that hoped to tip the explanatory balance back in favour of precognition. They hoped to do this by using a method that was argued to be outside of the scope of the PK ability of a single subject. This technique, called the “weather cut”, made use of the temperature readings in prespecified newspapers printed *after* the guesses were collected (e.g. Rhine, 1942; Rhine & Humphrey, 1942).

An example procedure using this weather cut is as follows: First participants filled in their guesses on a record sheet (putting an appropriate symbol in a space provided on the record sheet, identifying the symbol that will be chosen). Once all the trials had been completed, the targets were generated by somebody blind to the subject’s guesses. This was done using a newspaper and recording the temperature extremes as reported. The smallest number denoted the number of revolutions each deck (of 5) were to be given in the mechanical shuffler. The total of the temperature extremes was used to cut a deck of 25 numbered cards. The number resulting from this cut was the number used in cutting the decks after they have been removed from the shuffling machine. It was thus argued that PK would not be the best explanation for the results of a study involving this method (because the numbers obtained were dependent on the weather and this is a system that PK could not have much influence over), and thus precognition is the more plausible explanation for any above-chance scores obtained during testing. However, some have argued that the PK influence does not have to be on the weather system itself, but on the temperature recording device, the person making the recording or even the mechanical processes involved in printing the temperature in the newspaper (Roll, 1961). Later developments at ensuring experiments are tests of *actual* precognition have involved calculations which are outside of human abilities (Mangan, 1955) and, later, radioactive isotopes to again help ensure that the ability to predict through any normal means is severely reduced (Schmidt, 1969). It is clear then, that ensuring any test of precognition *is* testing what it sets

out to test is a somewhat contentious issue.⁸

The literature regarding experimental support for precognition has progressed rapidly since Rhine published his first work in the area and a discussion of every study is beyond the scope of this thesis. However, given the historical importance of the work, two particular empirical cases will be discussed, followed with a more general discussion of a number of meta-analyses. More specific discussion of empirical tests of precognition will also be covered throughout the thesis in relation to the regime of testing undertaken.

As Rhine had almost single-handedly developed the science of parapsychology, it was for other researchers to investigate and replicate his findings. This task was taken up in England by the well respected and sceptical researcher S. G. Soal, a mathematician who worked, in his early years, at the University of London.⁹ Soal was one of Rhine's strongest critics and spent a great deal of effort attempting to reproduce the findings that Rhine had spectacularly achieved. (Soal favoured cards bearing the picture of one of five different animals, rather than the symbol cards used by Rhine.) His own efforts were essentially negative, and Soal concluded that Rhine was prone to a number of 'errors', including not providing enough detail in his papers to reproduce the effect, inadequate randomisation of cards and a positive bias towards a paranormal explanation.

Disappointed at having spent a great deal of effort and not achieving a significant result, Soal became despondent. However, on the suggestion of Whately Carrington, Soal re-investigated his data, looking for a precognition effect, something Soal had not looked for at the time. On doing so, Soal identified two participants who seemed to show above-chance scoring, and his subsequent successful experiments with the two participants were held up by nearly all commentators as being of tremendous importance when assessing the case for precognition. In fact, Soal's initial scepticism, followed by his fortuitous reanalysis and subsequent championing of his results was part of the reason that his 1943 work (Soal & Goldney, 1943) was considered until the late 1970s to be some of the best experimental evidence in support of a paranormal hypothesis of precognition (e.g. Roll, 1961).

At the time there was some minor suggestion that something unusual had been happening in the procedure. For example, one of the participants alleged they had been drugged by a cigarette during testing. Scott and Haskell (1973), publishing in the important scientific journal *Nature*, showed that the pattern of scoring in Soal's work hinted at some kind of manipulation of the data, presumably as part of a fraudulent attempt to influence the results. Although the paper raised serious doubts about Soal's work, it was not until some time after his death in 1975 that Markwick (1978) provided unequivocal evidence that Soal had indeed manipulated his data to demonstrate a precognitive effect. Reportedly, Markwick had herself been influenced by a dream about Soal, and hoped to vindicate him, but found that Soal had often re-used the same random sequences when selecting targets and that wherever there were deviations from the often-used sequences, that the difference was in the favour of hits. By manipulating his data in such a way, enough spurious hits were accumulated to give the impression of a precognitive effect. Interestingly, Markwick reported that it was the lack of access to a computer that had fortuitously led to her discovery of the repeating patterns. Had she not 'eye-balled' some of the data, she stated, she would have been unlikely to discover the fraud. Today Soal is

⁸It should be noted that the precognition experiments of this thesis do not generally attempt to ensure actual precognition from a methodological standpoint, but instead argue from a philosophical perspective that precognition is an *internal* process that involves recognising a future state in the present, rather than the result of contemporaneous ESP or PK.

⁹It is an interesting fact that Soal was awarded a D.Sc from the University of London in 1944, primarily for his work on precognition.

totally discredited and is often ignored in reviews of the area. However, despite the ease with which Soal has been forgotten from parapsychology, his work was certainly some of the most methodologically rigorous of the early precognition studies, even if the results are questionable. Thus, it is important to bear in mind that much of the empirical impetus to continue research into precognition was inspired by Soal's work and that no research since that time has achieved quite the same effect.

Much more extensive discussion of the empirical support for precognition will occur with regard to the specific areas to be investigated. However, there are a number of meta-analyses and reviews which allow the entire body of empirical support for precognition to be assessed, if only from a general perspective. Morris (1982) reviewed what he called the evidence for true-precognition, that is, where the only explanation for the phenomenon can be a real precognition of the future rather than a clairvoyance or inference theory, which Morris argues is an alternative explanation for much of the original card-guessing research (as discussed above). Mangan (1955) used dice to generate numbers which were used to obtain entry into a random number table. Given the complex nature of the task, that is with 'fair' die the complex physical interactions between the dice should provide a random entry into the random number table, alternative theories regarding the phenomena, specifically that they might be a consequence of psychokinesis, were dismissed. Between 1955 and 1970 there were at least twenty-seven published studies using Mangan's technique. Of the 27, 7 reported significant overall deviations from chance (two-tailed) and fifteen reported significant internal findings based on earlier research. Two reported results which were against the hypothesis. Overall, Morris concluded that there is evidence for true precognition (though again Morris highlights the difficulty in establishing actual or true precognition over other psi hypotheses).

Honorton and Ferrari (1989) conducted a meta-analysis of 309 studies of forced-choice precognition, a methodology that requires participants to choose between members of a fixed set, for example the Zener card symbols, published between 1935 and 1987. The data set comprised of nearly two million individual trials, with more than 50,000 participants. Certainly, such a large data set allows for some interesting conclusions. Overall Honorton and Ferrari claimed that there is a significant effect in favour of precognition ($p = .00025$). However only some 30 percent of the experiments showed significant hitting at the 5 percent level. This effect is claimed across all experiments and it is argued is evidence against the results being the preserve of a few successful researchers (e.g. Akers, 1987). They also argued that the fact that 70 percent of the reports were non-significant was evidence against any file-drawer criticism. Honorton and Ferrari trimmed their data of certain outliers and were left with some 248 experiments from which to draw overall conclusions. They argued there is no effect that the quality of the study influenced its results, nor that the quality of the publication source was important. The year of publication also had no effect on the significance and Honorton and Ferrari claimed that this is evidence that despite supposed improvements in methodology, the basic effect has not disappeared. Finally, Honorton and Ferrari reported that it would require over 14,000 unreported null results to reduce the significance of the meta-analysis to that expected by chance (46-to-1) and that this number is much higher than Rosenthal's recommended safe ratio of just 5-to-1. Looking at the types of participants recruited across the experiments, Honorton and Ferrari found no difference between students, volunteers or animals, although they did report that 60 percent of experiments employing selected participants reached acceptable levels of significance, compared to only 21 percent of those using other participants. Overall Honorton and Ferrari concluded positively in favour of

interpreting the literature as showing a significant precognitive effect. They argued that the overall effect sizes are small but robust, and are comparable to effect sizes in medicine.

Steinkamp, Milton, and Morris (1998) conducted a similar meta-analysis, again of forced-choice experiments, but comparing clairvoyance with precognition. The report examined cases published between 1935 and 1997 in a number of parapsychology journals. Specifically, they compared studies employing similar methodologies and found that the overall size of the effect for clairvoyance was .009 and for precognition was not very different at .010. The authors concluded that there was no support for the idea that contemporaneous ESP works any better than precognition (when conditions were similar). Whether this means that precognition is a type of ESP is as yet, an open question.

Finally, it appears worth briefly discussing the precognition experiments of Daryl Bem (a well-respected mainstream psychologist, with an interest in parapsychology) whose recent papers have gained wide attention both inside academic parapsychology and outside in the wider scientific community (Bem, 2003, 2011). In these papers, Bem outlines a range of novel methodologies that each follow a simple structure: the experiments generally look for a conventional psychological effect reversed in the temporal direction. For example, one experiment asked participants to choose between one of two curtains on a computer screen, and after making their choice the computer would randomly present erotic imagery to them. The precognitive hypothesis predicted that participants would choose the curtain that was to contain the erotic imagery more often than chance would expect.

Similar results are reported for reversed habituation-type experiments and even precognitive influences on recall. Bem's large body of research appears to have significantly shown the existence of precognitive influences in a wide range of psychological tests, and in a way which should make independent replication and verification a simple task.¹⁰ Unfortunately, as should become clear as more and more of the experimental evidence for precognition is examined, Bem's is not a particularly unique claim (and it is arguably the academic reputation of the author that has driven the wide interest into this work, rather than the novelty of the claim), a sentiment echoed by Alcock (2011a).¹¹ Alcock is exceedingly critical of all the research described in Bem (2011), primarily concentrating on two specific practices. Firstly is Bem's habit of not correcting for multiple analyses (and thus declaring a result significant, when such a correction would reduce the result to non-significance), despite showing evidence in the paper that he understands that such corrections should be used. Secondly, Alcock declares Bem's methodology is suspect, as he appears to change procedures mid-experimentation, thus making it difficult to understand what his various constructs and measures are actually testing. Despite Alcock's criticisms, the ongoing debate continues (e.g. Bem, 2011; Alcock, 2011b).

Alcock is not alone in making such damning criticisms. For example Wagenmakers, Wetzels, Borsboom, and van der Maas (2011) argue that there are significant problems with Bem's statistical analyses and advocate for the use of a Bayesian t-test when analysing the results (which, when they do just that, they argue are non-significant). Interestingly Wagenmakers et al's criticisms are not just levelled at parapsychology, but experimental psychology in general (thus the error is not Bem's alone). Bem, Utts, and Johnson (2011) respond by arguing that although there is apparent advantage to such Bayesian statistics, that poor selection of priors can lead to incorrect conclusions, and this is their explanation for Wagenmakers et al's findings.

¹⁰The simple reason why this thesis does not contain replications of any of Bem's research is because the entire experimental phase was completed quite some time before Bem's research was published.

¹¹Alcock closes this particular paper with the phrase, "plus ça change, plus c'est la même chose" and we wholeheartedly agree that Bem's claims are not substantially different from those that have come before.

Given the contentious nature of this debate, it is not surprising that it is also ongoing (e.g. Bem, 2011; Wagenmakers, Wetzels, Borsboom, Kievit, & van der Maas, 2011). Unfortunately despite attempts to independently replicate some of Bem's claims, the findings have been generally mixed (e.g. Savva, Child, & Smith, 2004; Savva, Roe, & Smith, 2005; Ritchie, Wiseman, & French, 2012), although in their very recent review article Bem, Tressoldi, Rabeyron, and Duggan (2014) present a meta-analysis of 90 replications of the original Bem research and conclude that there is very considerable evidence and support for precognition in the dataset.

Further discussion regarding the empirical evidence for precognition will continue throughout the chapters of this thesis. It is clear though that there is *some* evidence that experimental tests of precognition appear to be related to the same kind of phenomena seen in the sporadic cases. However, it is also clear that separating out a precognition effect from contemporaneous ESP or PK is inherently difficult and a contentious issue.

1.2.3 Questions Raised by the Evidence for Precognition

Having reviewed some of the evidence for precognition, we can end this section of the introduction by asking and attempting to answer some questions raised by that evidence.

1. Does precognition imply that the future is determined?
2. Is intervention possible and, if so, what initiates the experience?
3. Are precognitions more often about planned rather than unexpected events?

In some ways these are the deepest questions of precognition and extensive discussion of determinism is beyond the remit of this thesis. At its heart is the mystery of how something that does not exist (the future) can have an influence on the present. Therefore it appears reasonable to ask whether precognitions involve direct acquaintance with a future state or whether they merely have the appearance of such? The philosopher C. D. Broad (who was very much interested in parapsychology) speaks specifically on this topic (Broad, 1968). Broad identifies two theoretical difficulties with the very notion of precognition. The first he calls *the alleged epistemological difficulty*. This essentially equates to the fact that any precognition appears to require perception (or as Broad refers to it "pre-perception") of the event in the future and thus must be (p. 189) "directly acquainted" with the event. Broad himself argues that this particular difficulty is false and that our perceptions are not as much tied to the present as we might imagine. He gives the example of how the sound of a gunshot, fired some distance away, will take time to reach our ears (by which time the event which caused the noise is over) and thus he argues that we are not always in direct contact with the present, although we are often uncritical regarding the simultaneous nature of some perceptions. As the event which caused the noise (the gun being fired) has ended before we perceive the noise of the gunshot, Broad believes there is no inherent reason why precognition cannot be perceived before the cause comes into being.

Broad (1968) believed that *the causal difficulty* is much more troublesome. The causal difficulty relates to the fact that as the future has no existence it cannot cause anything. There is thus no chain of events which could have caused the precognition. Broad (1968, pp. 194–195) concludes that "no case of ostensible precognition can possibly be one of genuine precognition", although he argues that the best cases of ostensible precognition will be of some importance to researchers interested in "*paranormal powers* which are already admitted or strongly suspected

to exist in some humans beings alive in the flesh.” Thus, Broad believed that precognition, whilst evidence of paranormal functioning, was not evidence for information processing from the future. Stevenson (1970a) agrees with Broad that a non-existent future cannot be the cause of anything and argues that any precognition is the product of an inferential process based on a present contingency.

Are future events entirely determined? Premonitions that appear to exactly describe future events suggest that *some* events are entirely determined and thus there can be no intervention. However most argue that whilst there appears to be some level of determinism, there is plenty of evidence that precognitive experiences are often not heeded and thus there appears to be a behavioural component regarding whether we act on premonitions or not (e.g. Stevenson, 1970a). A good example:

A particularly impressive case... occurred in the death of Robert Morris, Sr., the father of the financier of the American Revolution. Having dreamed that he would be killed by the firing of a cannon from a ship he was to visit, he tried to avoid going on board the ship. His grounds for such reluctance seemed so absurd that he was eventually persuaded by the captain that it was entirely safe to board the ship as no guns would be fired until he left. At the end of the visit, the captain ordered that a saluting gun should be fired only after he raised his hand when the party, including Robert Morris, had safely reached the shore. But as the captain's boat was still within range of the gun a fly lighted on his nose, and he raised his hand to brush it off. This was taken by the ship's gunner as the signal to fire and a fragment of the discharge hit Robert Morris and wounded him fatally. (Stevenson, 1970a, p. 193)

It appears that the percipient had every opportunity to change his behaviour and not die. But psychological pressures bore down on him to override his perceived sense of danger. Cases such as this imply intervention is certainly possible (and it is necessary that for precognition to be useful, positive intervention in the face of imminent danger must be possible).

Rhine (1955) noted that the question of intervention is a difficult question to investigate empirically, because presumably successful intervention eliminates the “target” and thus removes the ability to check whether there has been a “hit” or not. As mentioned above, instead, Rhine hoped that by investigating the spontaneous case reports with regard to whether intervention was attempted or not, some understanding would be attained. In nearly two-thirds of the cases examined by Rhine, no intervention was attempted. This is in spite of the fact that most of the precognised events were serious and it would have been desirable to prevent the event occurring. Of the remaining cases, some intervention was attempted and Rhine concludes were successful in the ratio of two-to-one. The reasons for failing included: the information precognised was too vague, a second person was involved who was not cooperative, the event was unpreventable even though intervention was attempted. Of those precognitions where intervention was apparently successful, there were occasions when the precognition allowed the danger to be avoided by an intended action (for example, leaving a hotel early where there was subsequently a fire) or the danger was avoided by not completing an intended action (for example not getting on a train that was booked in advance).

Stevenson (1970a) argues that in 27 percent of the 125 precognitive dreams analysed percipients attempted to avert the outcome. Hearne (1984) reports that in his case collection almost 60 percent had attempted to warn somebody in connection with the premonition. Whether or not the future can influence the present and whether intervention is possible are still open questions.

4. Is precognition more common closer to the event?

As already mentioned some have argued that as the evidence for ESP is allegedly well established, the further one gets away from the present, the more difficult it is to create accurate inferential predictions. From this kind of perspective (seemingly favoured by e.g. Broad, 1968; Stevenson, 1970a), the ability to predict the future is based on psi information gained in the present and then inferred into the future. A consequence of this perspective is that the greater the temporal distance between premonition and event, the less likely it is that the premonition is based on inferential ESP and the more parsimonious the explanation of true-precognition becomes.

Rhine (1942) compared precognition scores over a period of one to three days as compared to 10 days. The results suggested that the longer interval contributed more to the overall significance than the smaller. On the basis of the result, Rhine and Humphrey (1942) set out to directly compare these time intervals and found that the result for the 10-day period was more significant than that of the shorter interval. Osiris (1955) tested a single subject using a simple symbol guessing experiment conducted remotely through the postal service. He obtained significant results in both the short time frame (1–7 days) and the longer (30–33 days).

Two independent studies by the same authors involved prediction of targets one day in advance (McMahan & Rhine, 1947) and one year in advance (McMahan & Bates Jr, 1954). Although not separately significant, combined the two experiments are significant (although as part of the early precognition research there is some debate as to whether the results are true precognition).

Anderson (1959) also directly tested the success of precognition across two time frames, either a few days or a year. This was a remote testing experiment (the participant was in France and the experimenter in the US) and responses were recorded and posted to the experimenter. The procedure involved randomly designating half the trials to be checked immediately and the other half to be checked in a year's time hence. Randomness was achieved with a 10-sided die and complicated mechanical computation (in an attempt to ensure true precognition). The immediate trial targets were generated in the same way. The year trials were stored securely and the targets were generated a year after they were received. The combined results were significant, as were the one-year results (the immediate results were at chance). Anderson concluded that the results suggest precognition is not limited to events close in time.

Hearne (1984) reports that in his experience with spontaneous cases, the latency period between premonition and event was hours (18 percent), days (29 percent), weeks (27 percent), months (21 percent), years (4 percent), and over 10 years (.6 percent). This suggests, like other similar collections, that premonitions are more frequently experienced closer to the event onset.

Orme (1974) reviewed some of the case collections already considered and concluded that precognition generally appears to be related to nearer events than those more distant in time. He reports that the Pearson product moment of the correlation between the logarithmic temporal progression (essentially testing the number of premonitions across an ever increasing time period) was $-.964$, $p < .0001$, which he argues shows a clear linear relationship. Orme (1979) also analysed the prophecies of Nostradamus and found a trend that there is a quantifiable decrease in occurrence between the prediction and event (although there are breaks in that trend which Orme suggests might be due to the ease with which modern researchers are able to find connections between recent events and the written quatrains of Nostradamus).

Again, it appears to be an open question whether precognition is more common as the future event gets closer in time. It is of course possible that if precognition is something

different from ESP that there is a confounding influence of inferential ESP, which makes it hard to distinguish between apparent and true precognition.

5. Are precognitions more often specific or vague?

Rhine (1955) found a variety of specific and vague precognitive experiences in her case collection and argued that even the vague premonitions could be enough to encourage a concern of danger in the percipient, which appeared to have helped them escape later disaster. On the other hand, Barker (1967b) noted that none of the premonitions regarding Aberfan allowed the percipients to help in any way and most were vague and indistinct.

6. Are some people more prone to precognitive experiences?

7. Is it an acquired skill or something you are born with?

8. Are there personality constructs that predispose a person to having veridical precognitive experiences?

9. Is there a familial link to premonitions?

Rhine and Pratt (1957) argue that there appears to be no racial distribution of psi. Tests have been conducted around the world and there is no good evidence that rates of scoring differ depending on geographic or demographic considerations. Precognitive experiences appear to be relatively common, for example, Ryback and Sweitzer (1988) reported that one in two college students polled had experienced a precognitive dream (and one in three adults had experienced an allegedly psychic dream). Rattet and Bursik (2001) also confirmed a high incidence of precognitive experiences in their sample consisting mostly of college students and Blagrove, French, and Jones (2006) found that just under half of their sample reported an experience of a seemingly precognitive dream (their sample was composed of students and the general public). (Some of this research, for example Blagrove et al., 2006 has attempted to argue that the cause of individuals believing that they have experienced a precognitive dream is due to a misunderstanding of probability theory, or some cognitive deficit that makes probabilistic estimates flawed, see also Blackmore & Troscianko, 1985. Whilst this is a legitimate and reasonable line of argument, we are here concerned with assuming that precognition *does* occur and not with explaining away experiences as the result of cognitive deficits.) Barker (1967b) reported that percipients in his collection ranged in age from 10 to 73, though women were more often percipients in a ratio of 5:1, and Hearne (1989) confirms that most percipients of precognition are seemingly female.

Furthermore, Hearne (1984) reports his analysis into the types of people who are likely to experience such premonitions. Hearne published an appeal for premonitions in a national newspaper and received some 450 responses, 127 of which were of multiple precognitions. Hearne forwarded copies of the Eysenck Personality Inventory to each of the 127 multiple precognition respondents. Hearne argued that the study was nomothetic and intended to help build up hypotheses, thus there was no effort to investigate the veracity of the claims. Given the extensive statistics quoted in Hearne's study it appears worth reiterating his main findings. Approximately 69 percent of those sent questionnaires returned them. Of those, approximately 90 percent were female. The ages ranged from 23 to 80, with a mean age of 46 years and people in their thirties being the largest "group". Half were in employment, the other half were unemployed, retired or "housewives". Just over 70 percent of the sample were married, 6 percent single, and the remaining percentage divorced, separated or widowed.

44 percent reported between 2 and 10 precognitions, 35 percent between 10 and 50, and 22 percent over 50 precognitions. Age of first premonition was between 3 and 64 years, but the top modal range was between 10 and 15. The majority (43 percent) estimated the frequency of premonitions as once a year or longer, 39 percent reported frequencies of months and 9 percent one or more a week. Approximately half reported that the premonitions occurred in batches and 35 reported no such pattern. Dreams represented the largest form (44 percent) followed by waking-thoughts (34 percent), waking imagery (12 percent) and hypnagogic imagery (9.7 percent). In regards to some more personal aspects, Hearne reports that over 50 percent were neutral in their attitude to premonitions, 35 percent were unhappy and 12 percent were pleased. The majority of the respondents believed in life after death (63 percent), 34 percent were uncertain and 3 percent did not believe. One respondent was a member of a Spiritualist church.

In terms of other family members who also have reported premonitions (as reported by the original respondent), 48 percent reported that their mother also had premonitions, 9 percent the maternal grandmother and 4 percent the maternal grandfather. The maternal grandmother, but not the mother had premonitions in 17 percent. The father had premonitions in 19 percent and in 4 percent the paternal grandmother. In 6 percent the paternal grandmother but not the father. In 10 percent both mother and father. In 39 percent the daughter had premonitions and 25 percent the son reported them.

In regards to medical history, Hearne (1984) reports that 43 percent reported some form of serious illness (not necessarily mental) and 36 percent were currently taking medication, and in regards to personality factors, as measured by the Eysenck measure, the respondents were more neurotic than the normal population and scored higher on the lie-scale, but were not significantly different on the extroversion/introversion measure. This contrasts with Rattet and Bursik (2001) who found that extroversion was associated with precognitive experiences. Orme (1974) suggests that precognition is a common belief in schizophrenia, but he notes that they lack the ability to give specific details.

One interesting source is the work of Cohn-Simmen (née Cohn), who has investigated the Scottish tradition of second sight, generally taken to be the ability to experience prophetic visions. Cohn (1999) constructed pedigrees from families that reported a history of second sight, which were analysed with the help of a geneticist, looking for established patterns of inheritance. Cohn concluded that there was some evidence that second sight had an inherited genetic component, though it is difficult to exclude cultural and social factors entirely.

It appears that there is some evidence that some people may be predisposed to experiencing premonitions (for example the gender bias as believed by Hearne, 1984). But again it remains an open question as to whether these identified predispositions are directly related to veridical precognition, or are a consequence of some kind of sampling bias.

10. Are precognitions primarily about the self?

11. Is a personal connection important to precognition?

Stevenson (1970a) reported that in his analysis of precognition reports, 49 percent (of 125) were about events happening to the percipient. Furthermore, close family members were dreamed about in 31 percent of cases, distant relatives, friends and acquaintances in 11 percent and strangers in 9 percent. Stevenson sees this finding as suggesting that there is often seemingly a personal connection between the percipient and those involved in the event. However, Stevenson also notes that there appears to be one aspect that bucks this trend, that is the

Table 1.2: Percentage of events of precognitions as adapted from Stevenson (1970a) (p. 201)

	SPR series (1934)	Miscellaneous European and American Series
Death	35	47
Violent	9.4	33.5
Natural	25.6	13.5
Accidents	20	17
Illnesses	6	4
All other events	39	32

precognition of disasters, where, for example, percipients of the *Titanic* disaster were on the whole *not* personally connected with the event. Similarly, Barker (1967b) reported that very few of the percipients who contributed to his Aberfan disaster collection had any connection with the disaster. However, it might be argued that *all* the participants learnt of the disaster and like many “witnesses” (direct and indirect—through the national media, say) that they were all to some degree emotionally influenced by the sudden death of a large number of children in a preventable accident. If the Aberfan disaster was a national tragedy, would it be something that influences only those in the locality? Perhaps the same is true for the sinking of the *Titanic*?

Hearne (1984) reported that premonitions in his sample were categorised as follows: people known to the percipient (44 percent), persons unknown (30 percent), persons not known personally (18 percent), the self (.5 percent) and objects (2 percent). Thus, unlike other collections, most of Hearne’s premonitions are not about the percipient explicitly.

12. Is precognition more common in certain states?

Most of the case-collection studies have noted the propensity for precognitive experiences to be dream reports over other types of experience. To reiterate then, Sidgwick (1888) found 66 percent of premonitions occurred in dreams. Saltmarsh (1934) that 68.1 and Green (1960), 68.8. Rhine (1954) reported 68 percent also. Some have argued (e.g. Stevenson, 1970a) that this is evidence that precognition is achieved more easily during the sleep process, though there is no indication why that might be so. Stevenson suggests that it might be because the barrier between consciousness and the unconscious is weaker. However, it might just as easily be due to the lack of external stimuli. Experimental tests of precognition that have utilised the dream state have not been demonstrably more successful than the forced-choice tests (Schmeidler, 1994).

13. Is precognition more likely with danger?

14. Are some targets more successfully precognised than others?

As Stevenson (1970a) clearly states (p. 200), “Death is the commonest single event related to precognitive experiences” (see Table 1.2). Stevenson also notes that there is some evidence that percipients may have a specialisation: one person may seem to favour premonitions of aircrashes, another with ship sinkings. He argues that this should not be surprising to us, as we expect those with an interest in topics to more frequently dream of them (conventionally speaking) and that it should not surprise us that this is true of precognitive dreams. Another interesting question that Stevenson asks is if danger is important for precognitive experiences,

why are wars or battles not more associated with precognition? Stevenson tentatively answers this question by suggesting that the degree of shock or surprise contributes to the likelihood of a precognitive experience. As already noted, there were seemingly many more precognitive experiences associated with the sinking of the *Titanic* and very few associated with the sinking (in wartime) of the *Lusitania*. The former event was entirely unexpected, the latter was a wartime tragedy that was not totally unexpected.

Hearne (1984) also confirms that death is the most common topic of spontaneous premonitions making up 53 percent of premonitions in his sample. Hearne also notes that some of his respondents did specialise in one type of premonition, for example plane crashes.

15. Is precognition future-ESP or a different thing altogether?

16. In simultaneous testing, is scoring higher for contemporaneous or precognitive targets?

Nash (1982) reports that of the twelve published experiments which have compared scoring levels for contemporaneous targets against precognitive targets, eight have shown no significant difference, two have favoured precognition, one favoured clairvoyance (contemporaneous) and the other was mixed. In only the last four were participants blind to the condition they were being tested in, which Nash argues could clearly have influenced the results through expectation bias. To test whether there was better scoring on contemporaneous versus precognitive trials, Nash essentially had participants call one guess for two different targets, in the hope that the difference in scoring would indicate a preference for one type or the other of ESP. The difference between the two sets of scores was significantly different and Nash argues that the results show that ESP scores on present targets are higher than for precognitive scores. This is somewhat confirmed by Tart (1983) who found a significantly higher rate on present targets as compared to precognitive targets in a review of some 85 experiments.

Finally, as mentioned above, Steinkamp et al. (1998) conducted a meta-analysis of forced-choice experiments comparing precognition and clairvoyance and concluded that contemporaneous ESP is no better than precognition. Whether these findings suggest precognition is the same thing as ESP is, as yet, an open question.

17. Can precognition occur diametrically?

This question appears to have been first asked by Nash (1960). Diametric tests of psi generally involve tasks that require two separate acts to score a hit, for example an act of telepathy and an act of PK, e.g. Osis (1953). Nash was interested in whether she could obtain evidence for precognition when a hit required *two* separate acts of precognition. Nash's 28 participants were tested over a 6-month period in 40 separate sessions (one trial per session). Ten blank cards were placed into numbered envelopes (0–9) at the start of the experiment. Participants were informed after making all of their guesses (at the end of testing) random numbers would be transcribed onto the sheet (each participant had a different order) and that these numbers would identify the envelope to be used as the target for that trial. The target envelopes were assigned numbers using the stock exchange (and this is again an attempt to ensure this is a test of true precognition). The cards in the envelopes were numbered approximately two-years after the original guesses were made. Hits were scored if the number written by the participant matched the number on the card in the envelope appropriate for the random number transcribed on the sheet. Thus for participants to get a hit they have to seemingly precognise both the number on the card and the random number transcribed on the sheet. A novel procedure indeed. Unfortunately there was no above-chance scoring in this study.

It is clear that there are a range of questions open to the researcher interested in precognition. Very little has actually been settled and in fact the most basic question, “does precognition occur?” still remains an open and controversial one. What is not in dispute is that there is a range of evidence, both experimental and anecdotal that has been taken by some as evidence for precognition. Assuming as we will that there is some evidence that parapsychology, and precognition in particular, is worth further examination, there is another aspect that we wish to consider: that is, whether evolution can assist us in the investigation of precognition.

1.3 Psi as a Darwinian Adaptation

Before examining some of the theories regarding evolution and parapsychology, two strands of evidence will be presented which paint a picture in support of the idea that precognition is a kind of adapted perceptual ability. First, there will be a brief examination of the animal psi research. If animals show psi, perhaps this rules out both the idea that psi is a manifestation of a soul power (assuming animals do not have a “soul”) or that it is a uniquely human ability (a higher-order function perhaps). Secondly, there is a brief examination of research into altered states of consciousness. We have already noted that dreams are the primary source of precognitive experiences. Do altered states of consciousness suggest a biological basis for psi? Although some of the research and theories to be covered are not directly precognitive, given that we cannot yet be sure that precognition is something entirely different from contemporaneous ESP, such research and theories may be of some relevance when examining precognition.

1.3.1 Wonder-Animals and Anpsi

It is interesting to note that accompanying the increased pet ownership that began in the mid-nineteenth century, there was an increasing interest in the abilities of animals (both normal and paranormal) and subsequently a greater frequency of paranormal cases associated with pets (Lachapelle & Healey, 2010). Although animals had been associated with trickery and performances from much earlier times as part of travelling circuses and the acts of street performers, it was the intimate acquaintance and close personal bonds that were formed between owner and pet that led to significant interest in the possibilities of animal psi.

The infamous tale of Clever Hans is one of the best known cases of a *wonder-animal*. Briefly, Wilhelm von Osten, a German, taught his horse Hans to tap his hoof and count, perform arithmetic, tell the date and even read. At the time Clever Hans caused great excitement and there was much speculation in the media about how the feat was achieved. Psychologist Oskar Pfungst conducted an intensive investigation into Hans and produced a book-length report which revealed his conclusions (Pfungst, 1911). Given Hans’ abilities, there was speculation that the phenomena were due to some kind of telepathy or psychic connection between owner and horse, and yet Pfungst’s conclusions had implications for not only the study of animal intelligence, but also for those conducting psychical research. Specifically, Pfungst concluded that Hans did not have any special powers at all, not for arithmetic nor mind-reading. However, he identified that Hans was able to muscle read his owner’s reactions, even speculating that the increased number of rods and cones in the eye of the horse heightened its perceptual experience for movement. Given the almost, but obviously not quite, imperceptible nature

of the stimulus that Hans was reacting to, Pfungst's findings had important implications for testing other wonder-animals and human beings. We can be sure that Hans was not attempting to deceive anybody and yet some observers were convinced that a paranormal mechanism was the only explanation for the obtained results. (Pfungst's findings also somewhat supported Faraday's conclusions decades earlier, that the Spiritualist phenomena of table-tipping was caused by unconscious muscle action and not the action of spirits).

It is quite interesting that J. B. Rhine's first scholarly paper (written with his wife Louisa) was an investigation of Lady Wonder, the mind-reading horse. Despite Pfungst's investigation, Rhine was reportedly completely convinced of Lady's psychic abilities, which included guessing numbers off a writing pad, and declared Lady, "The greatest thing since radio!" (cited in Pigliucci, 2010, p. 84).¹² To be fair, with further investigation they discovered that the horse was using obvious bodily cues, much like Clever Hans (Rhine & Rhine, 1929b).

Along with the feats of the wonder-animals, there is another class of cases which are best described as extraordinary, usually involving tales of strong connections between master and pet. Victor Hugo reportedly gave away a beloved pet dog to a Russian aristocrat, that apparently travelled all the way from the count's estate (outside Moscow) to Paris. Such extraordinary homing behaviour, otherwise known as psi-trailing, is reported with relative frequency (Sheldrake, 2011). Other pet related phenomena were reported by Camille Flammarion (astronomer and psychical researcher) who received numerous reports, often from pet owners, that described telepathic communication as their animals lay dying. One owner wrote of a sad childhood event in which his parents had decided to have his adored dog put down following a complaint of which he had not been told. On the night of his pet's death, the boy experienced a vision of his dog in his bedroom looking sad, lying on the floor and then disappearing. He later found out that his mother had put the dog to sleep that night, 2 km away. Premonitions of dying pets in dreams were certainly one of the most reported occurrences, along with pets in danger sending out telepathic distress signals or pets trying to warn their owners were also common (Lachapelle & Healey, 2010).

Zohar (1983) gives what she calls one of the best accounts of animal death-avoidance in her book on precognition, entitled *Through the Time Barrier*. At the start of the Second World War, the Baines family of Wimbledon had built an outside bomb-shelter in the garden, but abandoned it because of damp, for a metal shelter in the house. They continued to use the indoor shelter between 1941–44. On the 30th June 1944, some 12 hours after a bombing raid that had blown out the windows of the family house, it was noticed that the pet dog, a spaniel called Merry, had gone missing. After a search, the dog was found in the abandoned garden shelter. Taking the dog back into the house, the family later found Merry missing, and again found her in the garden shelter. This is reported to have happened at least three times and because of her reluctance to leave, the family decided to sleep in the garden shelter, having cleaned it. At 2.50 am a bomb fell outside the house, destroying several properties, including the Baines's. It has been suggested that Merry's behaviour saved the family, although Zohar notes that there is the possibility that the humans themselves were the source of the precognitive influence and that the dog was merely the apparent focus.

A further interesting case of an animal seemingly being aware of imminent death (though this time there is no attempt to avert the outcome) is that of Oscar the cat, who lives in a dementia unit in the U.S. According to one of the doctors who worked in the unit, primarily

¹²Rhine and Rhine (1929a, p. 463) concluded, "There is left, then, only the telepathic explanation, the transference of mental influence by an unknown process. Nothing was discovered that failed to accord with it, and no other hypothesis proposed seems tenable in view of the results."

prompted by the nurses' observations, Oscar (who is apparently not a friendly cat) appears to 'know' that patients are soon to die and was observed around over 25 patients, who died soon after (Dosa, 2007, 2010). The doctor, Dosa, does pose a conventional explanation for Oscar's seeming ability (that cell death produces compounds that Oscar is capable of detecting), however he favours the paranormal explanation. Nickell (2013) appreciates the sentiment of Dosa's work (that dementia is a disease that requires public sympathy). However, he argues there is little good evidence for Oscar's abilities being anything more than wishful thinking and an unsupported anecdote. Clever Hans, Lady the Wonder Horse and Oscar the cat suggest that without careful experimentation it is very difficult to eliminate the more prosaic explanations for behaviour which may appear to some as paranormal.

With the rise of experimental parapsychology, some researchers attempted to take the controlled, laboratory methodology employed by Rhine and to test and find paranormal effects using non-human animals. Specific tests of animal precognition will be covered in more detail in Chapter 6. Morris (1977) in a chapter written for the *Handbook of Parapsychology* (Wolman, 1977) entitled *Parapsychology, Biology and Anpsi* describes a range of empirical studies using a variety of methodologies and species. One interesting case, that of Chris the "Wonder Dog", highlights the difficulty of conducting this kind of animal research (Wood & Cadoret, 1958). Although there was evidence that the dog was responding to body language cues, much more difficult to explain was the significant scores obtained when Chris was tested calling through a pack of Zener cards. Morris (1977) suggests that Chris's owner may have been the true psi responder in the tests with Chris, and that Chris was picking up unconscious bodily cues (rather than it being a case of true animal psi). It is often difficult to tease apart these kinds of factors.

Schmidt (1970) found that a heat lamp connected to a random number generator (RNG) would stay illuminated significantly more often when his own pet cat had access to the area under the lamp. Schmidt also reports cockroaches were shocked via electrocution on a grid (again connected to an RNG) more often than chance and, as Morris (1977) reports, this is a kind of anpsi-missing (the phenomena of psi-missing and the decline effects are proposed paranormal explanations for significant below-chance scoring). As Morris notes, there is a philosophical difficulty in interpreting this research, because the human investigator invariably knows the expected outcome and presumably wishes for a certain type of outcome. Thus whatever result is achieved can be explained just as easily by the experimenter's psi powers 'doing' what the experimenter wants. Similar research includes that by Peoc'h (1994) who found that in an experiment with young chicks tested in a room absent of light, a lighted candle (carried by a randomly controlled robot) spent more time in the vicinity of the chicks than compared to a no-chick condition ($p < .01$ (this was a test of psychokinesis not precognition).

The most interesting anpsi research highlighted by Morris is that which investigates the behaviour of a cohort of animals before a random selection are destroyed. (An experiment of this type is described in Chapter 6, along with more extensive discussion of similar experiments.) Morris (1967) recorded the behaviour of rats in an open field and found that those that were randomly chosen to die 10 minutes into the future showed a tendency to freeze compared to those who were not killed. Craig and Treurniet (1973) found their rats were more active 24 hours before being destroyed than the rats who remained alive. Related to his work with rats, Morris (1967) showed that goldfish which were randomly chosen to be netted and lifted out of the water exhibited significantly more activity beforehand, than fish which were not removed. A follow-up study did not find an effect (Morris, 1977).

Although anpsi research where the individual is subjected to harm or death is perhaps the most interesting (in regards to this particular thesis) there are a few experiments which have attempted to selectively breed animals who are successful in a psi-test, in the hope that the offspring of the successful animals will perform better in the future (a selective breeding element is part of the methodology of the experiment reported in Chapter 6). Robinson (1980) used a T-maze to test mice for ESP. The most successful animals were bred together, as were the least successful (like with like) for six generations. After the third generation, Robinson notes that the high ESP males were socially dominant over the low ESP males, but no evidence supporting a heredity ESP effect was found.

Lépes (1992); Lépes and Argibay (1994) reported a large scale ESP test of the common *Drosophila melanogaster* fly. Over many generations the flies were subjected to a simple T-maze test. The maze was in darkness with one arm of the T-maze ending in a light source, and the other arm in darkness, but with an enticing food source. Although the flies would initially navigate to the light source, after some time flies would move towards the food source (in darkness). A separate group of flies were tested in a similar condition afterwards, and it was hypothesised that some of these flies would gain knowledge of the food source via ESP from the earlier run flies. The next generation were bred from these “successful flies”. Results of these experiments were interesting but mixed.

Given that the spontaneous cases appear to concentrate on pet experiences, it should not be surprising that there are a core of experiments testing the abilities of animals to respond to psi, based on the observations of their owners. One famous set of experiments by Sheldrake and Smart (2000a) report work with the dog Jaytee, who the owner believed would behave in a way which indicated he was telepathically aware of when she was on her way home. Although Sheldrake and Smart are certain that the results of experimental testing support this conjecture, others who have conducted their own testing with Jaytee are not convinced that psi best explains the results (Wiseman, Smith, & Milton, 2000). Radin (2002) has suggested that one of the reasons for Jaytee’s lack of consistent psi responses may be due to environmental influences (e.g. fluctuations of the Earth’s geomagnetic field).

Sheldrake has not only been involved in the testing of other dogs that appear to know when their owner will return (e.g. Sheldrake & Smart, 2000b), but also involved in other psi tests of other species, including the study of an allegedly telepathic parrot (Sheldrake & Morgana, 2003). Although none of these tests have been strictly precognitive, they do suggest that if the results are due to psi responses originating from the animal, that psi is not unique to just the human species, but in fact can be found in a wide variety of species.

Relevant to the thesis that psi may be a form of death avoidance, Alvarez (2010) investigated short-term precognition (sometimes referred to as presentiments, or anomalous anticipation) in finches. Alvarez was interested in whether the finches would make appropriate predator detection calls *before* presentation of short video clips that showed a snake, as compared to during a control condition (where a video was shown, not containing images of the snake) and reports that there were significantly more calls during the experimental period than during the control period, which he takes as evidence for precognition.

In their article on anpsi research, Dutton and Williams (2009) review both the empirical tests of animal psi and the spontaneous case reports (some of which have been covered here), and argue that what appears to be important is the close bond between animal and human being. They suggest that it is this relationship which is important and that the psi experience is mediated by it. Although an interesting proposal, this is somewhat diametrically opposed

to the position being developed within this thesis (that psi is an evolved death avoidance behaviour, that does not require any other individual except the self).

Finally, as with the discussion of Soal's work and the influence that he had on the positive assessment of the evidence, until after his death when his work was discredited, it is worth noting that one particular researcher, a Walter Levy, did appear to make some considerable progress into animal psi research in the 1970s. A medical doctor turned parapsychologist, Levy worked with Rhine at his research laboratory (the Rhine Research Lab) and was touted to become the next leader of the lab when Rhine retired. Levy's interest was in animal psi and in a short time he had published significant research findings that seemingly proved that psi could be shown experimentally using animals. A fuller account can be found in Rogo (1985), but it is enough to say that Levy was discovered falsifying his results and left parapsychology. Though his work is not referred to often (as the results were fraudulently obtained), Levy no doubt had some impact at the time, dramatic as his results were (e.g. Levy, Mayo, Andre, & McRae, 1971; Levy, Terry, & Davis, 1973).

Overall, wonder-animals and anpsi appear to be like all the other areas of psychical research and parapsychology. It is hoped that there is a nugget of truth within the accounts. But it is important to realise that given the results of Pfungst and Levy, caution must be expressed at any great conclusions about where the psi is coming from—if there is any at all. However, it appears that psi may not be the single preserve of our species, a fact which must support the assertion that psi is an evolved ability or function.

1.3.2 Altered States: Both natural and induced

If psi has evolved, it must be dependent on biology. Another source of evidence that suggests psi has a biological basis is that of the altered states of consciousness (ASCs). The early history of psychical research had a close connection with the investigation of such states, including mesmerism, hypnosis, dreaming, waking hallucinations, etc. (Inglis, 1977).

It should not be surprising that, as the dream state was seen as important in the reporting of spontaneous paranormal experiences, dreaming has also interested experimental parapsychologists (though research into precognitive dreaming has been limited). There is some suggestion that dream recall is important in psi-scoring. Honorton (1972) ran a small precognition test in a group setting and found that participants who stated that they recalled at least one dream a week scored significantly better than those with less frequent recall. Johnson (1968) also reported a significant positive relationship between ratings of dream recall and scores on an ESP test. Haraldsson (1975) conducted a large-scale group-testing of 223 students in 100 precognition trials, who also filled in a questionnaire. Although dream recall did not significantly influence scores, there was a positive relationship between dream recall and reported frequency of precognitive dreaming. Other sleep related states that may be psi conducive include those associated with falling asleep (hypnagogic) and waking up (hypnopompoc), although the chaotic nature of these states may lead to experiences being more frequently identified as anomalous (Sherwood, 2002, 2012).

Direct testing of ESP in the dream state represented a major development for the field (e.g. Ullman & Krippner, 1970, 1973). This was part of the move away from forced-choiced testing using Zener cards to the use of free-response measures. In 1962, a dream laboratory was set up at the Maimonides Medical Center, which was used by Ullman and a number of other researchers to establish a programme of dream ESP studies, conducted in a controlled

setting. Participants were generally monitored during their sleep (by EEG) in a remote location. After the participant was secured in their sleeping location, a target (usually an art print) was randomly selected and given to an agent. The agent was sent to a secure remote location. When the EEG monitoring suggested that the participant was experiencing a period of REM sleep, the monitor signalled the agent to telepathically send the target image to the participant. After 10–20 minutes of REM activity, the participant was awakened and asked to describe the dream in detail. After describing the dream, the participant could return to sleep and the process was repeated during subsequent periods of REM activity. The same target was often used for the whole night. In the morning the participant was generally shown eight to twelve prints, one of which had been the target for the previous night. The pictures were ranked by the participant (along with a confidence rating). The typed dream descriptions and sets of art prints were sent to independent judges. For example, if 8 prints were included in a set, a hit was scored if the target picture was ranked 1–4 and a miss 5–8. Although the procedure was slow at producing large numbers of trials, it did appear to be a successful methodology.

One of the major adaptations of this procedure involved the use of a multi-sensory kit, so that the sender could immerse themselves in the process of sending. (This kind of methodology is one of the reasons why the free-response methods were extolled over the forced-choice methods. They appear to be much more naturalistic and more engaging than the artificial forced-choice methods, where the participant can quickly become habituated.) Rather than use a static art print, that the sender might find boring after a time attempting to send it, a multi-sensory kit was developed for the possible targets, such that the print was accompanied by other relevant objects. For example, for the target picture depicting the crucifixion of Jesus Christ, the kit contained associated objects such as a crucifix, a small cross, a picture of Jesus, some thumb tacks and a red felt-tipped pen. During the participant's REM activity, the sender was supposed to pin the Christ image to the small cross and colour simulated blood with the pen. An example of the dream reported for that particular target:

We were going to be sacrificed, or something, and there were political overtones. . . what we would do is pretend we were the gods. . . in looking at the so-called king, chief or whatever the native was. . . it would be almost like. . . looking at one of these totem-pole gods. (Ullman & Krippner, 1973, pp. 122–123)

This particular study was rated as significant by the external judges (with a hit rated for each of the eight nights).

Although not common, precognition was sometimes tested using a similar procedure. For example, Krippner, Honorton, and Ullman (1971, 1972) tested a so-called 'special' subject (that is someone who claims to be able to consistently score above-chance). As a precognition design, the targets were not chosen until the morning after the dream mentation was recorded. Both studies reported significant numbers of hits, suggesting that precognition results were also possible using this type of methodology.

The Maimonides lab produced a lot of other studies (van de Castle, 1977) but closed in 1978 due to lack of funding. A later meta-analysis of the dream studies conducted at Maimonides reported a highly significant p-value, $p < .000002$ (Child, 1985). Of course, given the success of the Maimonides results, other researchers have sought to replicate the findings. However, given the resources involved there have been no dedicated labs set up to do the same kind of dream research, though dream ESP research has continued. For example, Roe, Sherwood, Savva, and Baker (2005) used a dream diary procedure rather than a dedicated dream lab. Of course this type of methodology means that participants will usually not wake

up during REM sleep and that dream content must be judged from the diary contents (which may not contain all the relevant information). Most of the post-Maimonides research has eschewed the expensive testing laboratory and utilised less controlled settings (for example participants dreaming at home e.g. Roe, Sherwood, Farrell, Savva, & Baker, 2007). In their review of much of the dream research Sherwood and Roe (2003, 2013) conclude that whilst post-Maimonides research has generally been successful, on the whole dream research has yet to achieve the kinds of results seen at the Maimonides labs.

Finally, much more recently, and specifically regarding the induction of precognitive experiences, Pablos (2004) reported a pilot study using Rivastigmine (a chemical Pablos argued potentiates REM sleep), which reported an increase in the incidence of dreaming and with it a significant precognition result.

Much research into such ASCs argues that by enhancing internal attention (or reducing attention to outside stimuli), attention to the psi influence can be enhanced. Some have argued that the high incidence of psi reports amongst friends and relatives (and the low incidence amongst strangers) is evidence of a learning influence over extended interaction (Stevenson, 1970b).

As was noted in an earlier section, many spontaneous cases are reported in altered or dream states (e.g. Rhine, 1962). Broad (1953) suggests that psi is a process that occurs continuously on an unconscious level (he called these psi *interactions*) and distinguished them from psi *experiences* which entered consciousness. Thus it might be seen that one of the problems of psi is *not* whether psi occurs or not, but what conditions facilitate its detection. Honorton (1977) suggested that the ancient texts of India claimed that certain paranormal powers were associated with meditation and that the eight stages of Raja yoga can be seen as a system of progressive psychophysical noise reduction. A number of studies have seemingly reported a significant association between psi-scoring and meditation (e.g. Schmeidler, 1970; Dukhan & Rao, 1973). Braud (1990) concluded that there is a trend for positive results in tests of PK that utilise meditation (though again, this is not found in all such studies). Schmeidler (1994) in her large-scale review of ESP experiments, concludes that although meditation appears to be a psi-conducive state, that it appears to be so only if the meditators believe that the outcome of the study will be successful.

Related is the work on ESP in the hypnotic state, which itself has quite some history. For example, Mesmer demonstrated what was described at the time as mental suggestion at a distance, which today would be interpreted as ESP (Inglis, 1977). Despite the reported early connection between hypnosis and psi-like experiences, the experimental work has not conclusively proved the hypnotic state as specifically psi-conducive. For example, Fahler (1957) compared the performance of four participants in a clairvoyance and precognitive task, in both a waking and hypnotic state, and found that scores in both the clairvoyance and precognitive tasks were significant only in the hypnotic state (scores were at chance in the waking condition). Casler (1962) compared the influence of suggestion on performance in a clairvoyance task and found that participants given explicit suggestions to score for high ESP produced significant ESP scores as compared to those in a no-suggestion hypnosis condition and those in a waking group. Honorton and Stump (1969) investigated hypnotically induced dreams and found that participants were able to significantly match their dream imagery with the target image (as compared to three random images). Honorton (1977) concludes the research into psi-performance and the hypnotic state suggests that hypnosis can enhance psi-scoring. Interestingly from their meta-analysis of some 29 studies investigating hypnosis

and psi, Stanford and Stein (1994) concluded that results appeared very much dependent on the experimenter and on the selection of participants. Parker (2005) argues that the hypnotic state is a potentially complex one, and thus parapsychologists have generally avoided utilising it, in favour of some of the other states, which are considered (if not less complex than hypnosis) more capable of producing positive psi results. Luke (2011) reports that hypnosis continues to be a state that is suggested as being psi-conducive, but results are mixed.

Parker (2005) notes, with the hypnotic state appearing complex, and the dream studies being difficult and expensive, an altered state was sought that might avoid these problems. Much of this research argues that by reducing attention to physical sensations, we can direct attention internally. This leads us to one of the most controversial paradigms, the ganzfeld technique. Because the ganzfeld is not often used in a precognition setting, it will only be covered in brief here. Because dream studies and hypnosis studies often require a lot of effort for a relatively small number of trials, a methodology was developed that attempted to remain akin to these more demanding methodologies but was easier to collect data in. Participants in the ganzfeld are usually reclined in a comfortable chair in a quiet environment. Two halves of a ping-pong ball are placed over the eyes and a uniform source of light is shone in front of the face. This results in a diffuse "sea of light" which disorients the participant. Often white noise is also played to the participant through headphones. It is generally argued that the ganzfeld technique induces in the participant a state akin to sleep onset (hypnagogia). However, Wackermann, Pütz, Buchi, Strauch, and Lehmann (2000); Pütz, Braeunig, and Wackermann (2006); Wackermann, Pütz, and Allefeld (2008) continue to argue that the ganzfeld may not induce a hypnagogic state, based on investigations of electrical brain activity.

The status of the ganzfeld as a psi-conducive state has been extensively debated (e.g. Bem & Honorton, 1994; Bem, Palmer, & Broughton, 2001; Milton & Wiseman, 1999, 2001; Storm & Ertel, 2001), but it is widely seen as one of the best methodologies for producing replicable evidence for psi. Honorton (1977) reviewed more than 80 experimental tests covering the range of phenomena discussed and concluded that together they argued for a significant effect:

Psi functioning is enhanced (i.e., is more easily detected and recognized) when the received is in a state of sensory relaxation and is minimally influenced by ordinary perception and proprioception. (Honorton, 1977, p. 466)

This suggests to Honorton that something like filter theory is relevant. Filter theory, as proposed by Henri Bergson, suggests that the brain and nervous system function as filters, protecting us from being overwhelmed by huge amounts of mostly useless information from outside ourselves, leaving behind just the important and useful. The states discussed above serve to reduce the bombarding of information on the filter and instead reduce the information flow to a trickle, in the hope that psi information will be more readily discernible. Storm, Tressoldi, and Di Risio (2010) analysed a large number of free-response studies, comparing the ganzfeld with other types of procedure (e.g. dream studies) and concluded that the mean effect size for the ganzfeld was significantly larger than for the other types of free-response studies. The debate as to whether the ganzfeld induces a psi-conducive altered state is still ongoing.

The final aspect of altered states that will be covered here is the use of drugs in an attempt to induce psi-conducive states. It should not be surprising that given that both the dream state and relaxation/sensory deprivation have been suggested as being conducive to psi-experiences and psi-scoring, that researchers have sought to induce such states through the use of chemicals, and perhaps make the evidence already achieved stronger and more reliable. It is interesting that some of the earliest ESP experiments reported by J. B. Rhine attempted to investigate

the influence of drugs on psi performance, something that if demonstrated would suggest a biological mechanism for psi (Rhine, 1934). These early experiments seemed to show that heavy doses of narcotic drugs reduced the usual significant scores to a non-significant level. The stimulant caffeine has been claimed to raise scoring levels (Averill & Rhine, 1945). Other drugs have also been tested (such as dexidrine and sodium amytal), the results of which are mixed (Cadoret, 1953). Interestingly Luke and Kittenis (2005) report that anomalous experiences are almost entirely absent from non-psychedelic drugs (such as heroin), which is interesting giving that relaxation itself has been reported to be psi-conducive (Braud & Braud, 1973; Braud & Braud, 1974; Braud, 2002)..

Tart (1977) highlights a major difficulty in conducting drug-based research (focusing on hallucinogenic drugs). Tart argues that despite standardising doses, etc., individual differences with drug experimentation (recreationally) means it is difficult to ensure a standardised drug-based experience in the lab. One person's experience may be very different from another, though the amount of the substance and conditions that it is administered in, may be very similar. For example, a famous study by Orne and Scheibe (1964) showed that participants in a seemingly benign experimental setting can have very different reactions. In their study, Orne and Scheibe had participants sit in a small room in a hospital. The researcher who briefed them was dressed in a casual manner and they were told that they were to sit in the room for two hours, after which he would return and interview them regarding their behaviour. If the participant could not stand waiting, they were to press a "panic button" on the wall and this would bring the researcher to end the experiment. No participants pressed the button, and when interviewed participants were generally bored, but not overly distressed by the experience. In a second condition, participants were greeted by a well dressed and seemingly important "psychiatrist", who took a long medical history. Participants were asked to sign a detailed release form, and in plain sight was a tray of hypodermic needles, marked "emergency tray". In this condition some participants pressed the panic button ending the experiment early and many reported feelings of anxiety. This study illustrates that the way an experiment is set up and presented to the participants can result in very different outcomes, thus as Tart (1977) argues, differences in drug experiments may not be due to the effect of the drug, but rather due to individual differences to the experimental set-up.

In his comprehensive review of research into paranormal phenomena and psychoactive drugs, Luke (2012) argues that given the obvious role that altered states play in spontaneous case reports of anomalous experiences, psychoactive (or mind-altering) drugs are an obvious candidate for better understanding psi. Luke agrees with other authors in the area that it is probably important that such substances seem to have a direct impact on temporal and spatial awareness that makes them useful in facilitating or inducing psi experiences (presumably by altering brain chemistry). As Luke states, it is entirely possible that psi is mediated in the brain through the action of endogenous chemicals (although he argues against concluding that they are the cause of psi, but rather merely part of the process). Given the controversial status most psychoactive substances have enjoyed over the last hundred years, it is not surprising that most of the experimental research investigating psi and drug states have been exploratory, and that results are mixed. Luke argues that those experiments that have used the conventional card guessing tests are not interesting enough for participants partaking in a psychedelic experience and that more engaging experimental techniques must be used. There are also a range of methodological problems associated with such tests, for example individual responses to substances is difficult to control for. Ultimately, though it is without doubt that this area

of massive interdisciplinary crossover, is only now beginning to receive the attention that it deserves, and thus it is for future researchers to determine if such states are psi-conducive or merely appear to be.

From this brief overview of ASCs it should be clear that there are many states (both natural and induced) that may be conducive to the manifestation of psi and one might interpret much of that evidence as supporting the contention that psi has a biological basis. Although the research into altered states is still very much ongoing, and results are on the whole mixed, it is encouraging that much of the research suggests a biological basis for psi. If psi is biological in nature, and shared between different animal species (as suggested in an earlier section) then it is likely that psi (and precognition specifically) is an adapted behaviour that has evolved because it is advantageous to the individual (this line of argument is continued in the next section).

1.3.3 Evolution and Psi

Finally, if we accept that there appears to be some good evidence that psi has a biological basis (and we will cover some of the physiological evidence in the next chapter), are there any evolutionary theories of psi? This sub-section will give a brief account of a number of different theoretical perspectives that have been put forward for the existence of psi as an evolved ability. None has been universally accepted within parapsychology and at best represent the field recognising that evolutionary theory may be an important aspect to consider. It is also worth recognising that the theories generally divide between those who see evidence for psi as evidence for some kind of deity or soul power, and those who see psi as something that has evolved and is therefore biologically useful.

The famous British biologist Sir Alistair Hardy postulated in his book, *The Living Stream* (Hardy, 1965), that the research findings of parapsychologists might profitably fill in some of the gaps of neo-Darwinism. Based on his Gifford Lectures, the book bore the subtitle *A restatement of Evolution Theory and its Relation to the Spirit of Man* and was unashamedly a work of what is known as natural theology (a theological discipline that argues that understanding nature can help with theological problems).

The book itself is presented in two halves. The first presents the accepted theory of evolution and the second seeks to demonstrate where the inadequacies of the theory lie and how the theological and psychical considerations might fill in those missing gaps. Hardy (1965) was a staunch believer in natural selection as the primary process of biological change and yet was troubled by certain biological evidence that seemed to defy conventional explanation (for example, the physical similarity between the Tasmanian wolf and the placental wolf, both genetically very different creatures, but both physically very similar). Hardy suggested that there is unconscious telepathic communication between individuals of a species, which helps stabilise common behaviour patterns until such a time that genetic adaptations convert them into fixed instincts. Hardy does not suggest that the psychic phenomena acts directly upon the genetic structure of an organism, but only that changes in habit are facilitated by this common connection which increases the genetic fitness of the species. This would allow advantageous behaviours to spread amongst individuals through a common psychic transmission. Behaviour is a major selective force. Hardy argues that the organism chooses its behaviour and the right type of environment, rather than the environment selecting for organisms. Hardy gives a specific example: the beaks of different birds demonstrate that they are used for different

purposes and on the whole to eat different things. Hardy argues that it is the change in behaviour that first drives the beak changing, not the beak changing and then behaviour following.

This is not a theory of precognition, more a synthesis by Hardy (1965) of the evidence for telepathy (which he is convinced is genuine) to help explain what he refers to as problems in evolution, and it is not clear how precognition would fit into his theory (if it could at all). Rupert Sheldrake's theory of morphic resonance (which appears to suggest that there is something akin to a shared memory between members of a species) appears to be quite similar to Hardy's proposal (e.g. Sheldrake, 1981).

Biologist John Randall, in his book *Parapsychology and the Nature of Life* (Randall, 1976) gives an exhaustive account of the state of parapsychology and along with it proposes an evolutionary theory of psi. Randall's theory is dualistic in nature, arguing that mind and matter are two very different things. He seriously suggests that life arose on this planet through the action of mind: it is this mind process that first initiated the coming together of the simple chemical precursors of life and that throughout the course of evolution this same mind process has intervened, more or less, to produce creatures with an ever increasing capacity of mind.

As Randall (1976) notes, precognition is a much harder phenomena to place in such a theory. He suggests that precognition may be evidence that the mind, or some part of it, sits outside space and time as we currently conceive of it. It is extremely controversial to presuppose that mental processes are not only independent of biology but have, during the course of evolution, directed its evolutionary progress. Randall's theory, like Hardy (1965) before it, does not really offer an evolutionary explanation for psi, but attempts to shoehorn evolution and the paranormal together. It is clear that these types of theories appear anathema to evolution (as will be discussed below) and do not appear to offer much in the way of testable ideas.

Kennedy (2004) is another parapsychologist who is also interested in what the purpose of psi is. Like Hardy (1965), Kennedy argues that there are certain mysteries in biology (such as the origin of life) that hint at a need for a para-biological explanation. Kennedy argues that despite the general parapsychological consensus that psi is needs-based or motivationally driven, that both psi in the real world, and psi in the laboratory suggest quite the opposite: that psi does not appear to be useful in any evolutionary sense, and is quite sporadic and capricious. Kennedy concludes (partly on the basis of a personal experience) that the purpose of psi experiences may be 'transformative' and that dramatic personal experiences of psi may have a positive existential influence upon the experient. He therefore sees psi as a manifestation of this para-biological process whose purpose is to induce an expansion of consciousness (whatever such a thing might actually mean).

Specifically then, Kennedy (2004) believes that psi experiences direct attention to what he calls the 'spiritual realm' and away from materialistic self-interest. On the other hand, the capricious and sporadic nature of psi is explained as a way of preventing the use of psi to achieve 'material gain'. Kennedy argues that the spiritual mode of life need not be extra to biological life on Earth, and could have evolved concurrently with it.

The perspective outlined in Kennedy (2004) seems more akin to those of Hardy (1965) and Randall (1976), and again, is not a theory directly dealing with precognition. Ultimately, it will take the testing of the different perspectives (the materialistic such as the research outlined in this thesis) and the more spiritualistic (as Kennedy espouses) before we can begin to determine whether one perspective is more or less correct than any other.

A different perspective is proposed in a paper by Jules Eisenbud (Eisenbud, 1976), who was a psychiatrist and influential member of the parapsychological community.¹³ Entitled 'Evolution and Psi', the paper outlines Eisenbud's musings on the two topics and postulates a whimsical theory. First, Eisenbud remarks on the lack of investigation into the genetic aspects underpinning psi, and laments that there are no palaeontological observations that appear to require psi as an explanation. Thus, Eisenbud argues that, despite the extensive historical records, it is difficult to predict which way psi appears to be going. Is it a vestigial remnant and thus part of our evolutionary heritage? Or is it something like our brain matter, that is likely to increase in size and ability for as long as we are around? Without data we cannot make conclusions.

Eisenbud (1976) remarks that his theory of evolution and psi first came to him during some chance observations related to the predating behaviour of a nursing cat (as compared to a cheese-baited mousetrap). Eisenbud pondered the fact that, despite the inviting nature of the cheese (and the presumably repulsive nature of the cat), it seemed as though the opposite were true: that weeks might go by without the mousetrap being sprung, compared to the often hourly successful predation abilities of the nursing cat. Eisenbud proposed that psi mediates an ecological relationship between the predator and the prey. That populations have a kind of pecking order, where the vulnerable individuals represent a kind of 'roving reservoir', not just the sick and old, but also the aberrant and deviant, and that as animals go about their business they are somewhat at the whim of this ecological relationship. If a needy predator comes within range, then the weaker individuals seemingly offer themselves up in place of the stronger or more important members of their species (and to reiterate, this relationship is presumed to be mediated by psi). According to Eisenbud, this process enables populations to not get too large or too small. The vulnerable individuals are offered up to the predator, who in turn is satiated with just those individuals that it benefits the population to lose (for the good of the species). The cat and the mouse are part of a biological system that work together for their mutual benefit.

Finally, Eisenbud (1976) deals specifically with precognition and its relation to evolution. Eisenbud asks the question that if there is information transfer from the future, does this imply that evolution has an overarching plan, toward which it is groping? He wonders if the results obtained in the animal precognition research are merely deviations from the game (in the sense that removing the animals from their natural ecological niche has released them from some of the pressures that usually keep their psi powers in check). He concludes that ultimately it is unlikely that the issues will ever be settled.

The theory outlined in Eisenbud (1976) is certainly interesting, though whether there is any merit in comparing a static mouse catching machine a couple of hundred years old with a predator millions of years in the making seems unlikely. It is certainly novel to suggest that different species might use psi information to enter into a mutually beneficial (it might even be described as a *psi-symbiotic*) relationship. However, it does appear that at the heart of Eisenbud's theory is the idea that prey enter into a psi-mediated death sacrifice for the good of the species. Modern neo-Darwinism appears to argue that individual genes are the most important level of selection (Dawkins, 1976) and that individuals and species are related inasmuch as they share percentages of genes. Thus, as Humphrey argues below, if the individual gives up his life for the greater good of the species, the genetic complement that led to that sacrifice will be lost

¹³Eisenbud was best known outside of parapsychology for his widely criticised investigation of 'thoughtographer' Ted Serios, (e.g. Eisenbud, 1977).

and with it a reduction in the genes that led to the sacrifice. Cooperation amongst species of the kind suggested by Eisenbud (psi-mediated or not) appears to run counter to the modern understanding of evolution (though good evidence for psi could swing that interpretation back). Whether organisms form symbiotic relationships mediated by psi has not seemingly been considered before. Though such a thing may account for occasions that may look like apparent precognition (seeming foreknowledge of a predator attacking the group, say), it is clear that Eisenbud's theory cannot explain true or actual precognition.

Thouless and Wiesner (1946) argued that psi had been sufficiently established through spontaneous case reports and empirical testing and that it appeared to be a single paranormal capacity to react beyond the normal limits imposed by space and time. They suggested that psi-determined behaviour is like normal behaviour and is correlated to some degree with events in other parts of the world. They argued that if there was no direction or specificity to psi information, then it would provide no useful information; that is, that if every part of the universe was just as likely as any other part of the universe to stimulate a particular psi system, then such a system would essentially be random and useless for any particular task. Importantly, Thouless and Wiesner suggested that early life may have consisted of organisms that were bombarded with psi information randomly from the rest of the universe and that the evidence that exists today suggests that over time, a selection pressure has been to filter down the different sources, so that only those that are relevant and useful are acted on. They also argued that the conventional senses that we identify today evolved after the psi-sense: psi is seen as a pre-sense. Thouless and Wiesner had little to support their speculation. However, they did suggest that if sensory perception is a later acquired process normally suppressing the effects of psi, that any factor which favours sensory perception will be detrimental to psi. They argued that if the psi-signal is ignored by the conscious mind, it may be demonstrated through involuntary, or unconscious actions. As discussed above in the subsection on altered states of consciousness, there is a range of evidence that suggests that reducing sensory input can seemingly facilitate awareness of psi. Is this because the vestigial psi-sense is usually swamped by the later acquired conventional sensorimotor system and associated processing? Physiological evidence for psi will be covered in the next chapter, but there is also a range of evidence which suggests concentrating on physiological responses in psi experiments may provide some of the most compelling evidence for psi functioning.

Furthermore, if we take psi to mean precognition, is it possible that a selection pressure has been to respond to precognition that is relevant to the organism and that natural selection has driven this process? Being able to identify specific information regarding the future would have enormous benefits to the survival of the organism. Of course this hints at an answer to one of the major questions: if precognition is an evolved ability, why is it not better? This conception of precognition as a pre-sense hints at a range of answers to this question. Perhaps it is difficult to identify future relevant information specific to the individual from a sea of information. Presumably this pool of information does not come with specific identification information attached and perhaps what we identify as psi is a vestige of an aborted evolutionary pathway. Perhaps the conventional senses provided reasonably good information that was on the whole useful (and allowed some future inferences) as compared to a very unreliable, but occasionally spectacularly correct psi-sense. It may be that psi is an atavism (akin to Haidinger's brush, as mentioned at the onset of this chapter). It is clear that the psi-as-a-pre-sense perspective is not without its difficulties, but certainly helps explain some of the idiosyncrasies of psi research, without recourse to a theory that runs counter to current understanding of evolutionary theory.

Certainly related to Thouless and Wiesner's position (and perhaps one of the most well recognised theories of psi) is Rex Stanford's *psi-mediated instrumental response* (PMIR) model (Stanford, 1974a, 1974b, 1977, 1982, 1990). PMIR was first described in Stanford (1974a), where it was proposed that organisms actively scan the environment via psi, based on their current needs, and discover appropriate information accordingly. The active scanning component of PMIR was dropped altogether in Stanford (1978, 1982), mainly on the basis of parsimony. Stanford argued that because the organism cannot know where important and useful information will derive from, random scanning of the environment in the hope of hitting such information appears to be a formidable challenge. In the light of the evidence for precognition this challenge is made considerably more difficult, as the temporal future increases the potential location of such information exponentially. However, Stanford confirms that the goal-oriented or needs-based nature of psi remains important. In the absence of psi-scanning, Stanford suggests a psi-filter mechanism instead determines what information is needs-relevant.

Generally speaking for Stanford (1990), responses to psi information are predicted to be the same as if the information had been acquired through the conventional sensory system. Moreover, any response initiated by psi information should be accompanied by a conventional adaptive preparation to respond (for example the orienting response). Thus Stanford argues that some of the parapsychological research that examines biological responses (such as EEG) seem to support this contention. Directly relevant to the work described in this thesis is the fact that Stanford argues that the strength of the PMIR is directly related to the importance of the need. Thus, we might expect that as death represents the annihilation of the individual (something that most organisms will presumably fight against), that PMIR will be at its most strong when the organism needs to avoid future death.¹⁴

Stanford (1990) argues that PMIR is basically achieved through the interaction of psi information and the organism's current dispositions and plans of action. Essentially PMIR can weight the current needs based activity system such that advantageous behaviour is initiated (or planned behaviour that will bring about disadvantageous circumstances avoided). Stanford (1982) gives a personal example of an occasion when he was driving home and felt an urge not to take the planned route. Later he went back to examine the route he would have taken and found that indeed there was a situation that could have been personally dangerous. As Stanford suggests, PMIR may influence the timing of actions.

Finally, Stanford (1990) suggests a number of reasons why, if psi is adaptive, does it not help organisms in an obvious manner in everyday life. In other words, why is psi so very elusive? The main reasons that Stanford cites for the lack of effectiveness of PMIR are generally psychological:

... they include behavioral rigidity, inhibition, stereotypy, response chaining, strong preoccupations, and the unavailability of relevant ready responses. (Stanford, 1990, p. 138)

Such factors might well explain why only a small number of people avoided the disasters described previously (such as the sinking of the *Titanic*). It might also explain why animals do not appear to be more psychic than humans (because stereotyped behaviour and behavioural rigidity are limiting factors in lower organisms) and that although humans appear to have

¹⁴Some organisms, for example the male praying mantis, are killed by the female during the process of mating, thus death avoidance would actually reduce reproductive success (at least death from the female). But even in this example one might argue that the male mantis' need to reproduce is *greater* than even the need to survive, suggesting that on some occasions, personal survival might not be the greatest need an organism can have.

greater freedom than other animals, their strong preoccupations and unwillingness to change behaviour might well act against PMIR also.

The model of PMIR as developed by Stanford (1990) was proposed to explain both the spontaneous case reports and experimental evidence for psi and to show how psi might work without the percipient ever being aware that they were using psi advantageously. Stanford himself notes that this conception of psi working independently of will is in stark contrast to that supposed by many parapsychologists, where participants are very often expected to will a specific outcome (for example in the Rhinian card guessing tasks). Stanford even suggests that laboratory settings may have a direct influence on results (the closer the setting is to real-life, the more likely that natural psi is to be invoked) and that there are many spontaneous case reports that do not appear to support the notion that psi is purely a conscious faculty, where despite ignorance about the importance of directing psi, the percipient has acted in a way which suggests psi has helped them (presumably unconsciously).

Recent years have seen a resurgence of interest in PMIR. For example, a series of experiments (Luke, Delaney, & Sherwood, 2008; Luke, Roe, & Davison, 2008; Luke & Morin, 2009) have attempted to test for precognition in a concealed test (i.e., the participants are not aware that they are taking part in a psi task) where participants are “rewarded” by success with a positive follow-up task (e.g. rating pornographic images relevant to their sexual orientation) and failure results in a “punishment” (e.g. a monotonous number task). It is thus argued that if PMIR is true, participants will use their psi ability to succeed in the task, despite not being aware that success or failure is contingent on their responses on the concealed test. Most of the experiments produced results in line with the hypotheses. Interestingly, Hitchman, Roe, and Sherwood (2012) attempted a replication of this work but found that participants who were not required to take part in the contingent aspect (some of the participants were randomly allocated out of that contrast) scored better on the psi test than those whose performance would have a direct impact on which reward or punishment they would be subjected to. Ultimately this runs counter to the expectation from a PMIR perspective. The Hitchman et al. (2012) replication failed to reach statistical significance.

Although PMIR is an interesting model, it is clear that there is a divergence between the death-avoidance hypothesis being developed in this thesis and the model as proposed by Stanford (1974a). Namely, the active scanning component of PMIR (which was dropped in later revisions of the model, e.g. Stanford, 1990) is not how psi is conceived of here. Instead, on a very basic level, precognition is seen as something internal to the organism: the state of the future is learnt and this future state is somehow conveyed backwards in time internally to the organism *in the past*. Stanford also neglects to discuss the evolutionary aspect of his model, though he is clear that psi responses should be adaptive. Yet again further research is required to establish if PMIR adequately explains psi in all its forms.

In his 1987 presidential address to the Parapsychology Association’s 30th annual convention, Richard Broughton (Broughton, 1988) made the point that though parapsychologists frequently treat psi as though it is an ability, they rarely ask what it is for, and makes a very clear statement that (p. 195) ‘if psi is a product of human evolution then its function is to help ensure the individual’s survival’. This is very much the underlying position taken in this thesis. Broughton argues that psi will not only be used in a crisis, but that it will be used by the individual in all events (akin to PMIR). He argues that it is likely that psi is unevenly distributed in the population and that some are going to be better or more successful than others. Psi may not just have evolved to avoid plane crashes but also in intuitive decision making processes to

provide a competitive advantage over other individuals. Broughton provides a novel reason why psi appears to be so very elusive. He argues that ostentatious displays of psi may get the person into trouble and thus psi has been driven to be kept subtle and hidden. (He proposes that this explains one of the great disparities within the conception of psi: on the one hand, it appears to be boundless and unrestricted by conventional standards—i.e. time and space—and on the other hand the evidence is very weak and not impressive.)

Broughton (1988) argues that John Maynard Smith's concept of an evolutionary stable strategy (ESS) may be crucial to the understanding of psi. The theory of an ESS essentially argues if the majority in a given population adopt a certain behaviour or strategy, it cannot be bettered by an alternative (rarer) strategy, and that ultimately there may be a range of costs and benefits to displays of psi, ostentatious or otherwise, which may explain the unreliable and elusive nature of psi. It appears premature to conclude that the lack of success in establishing psi in an experimental setting is because ostentatious displays of psi are not advantageous for the individual, but it is certainly an interesting argument.

Broughton (2010) continues to argue that psi is evolutionarily advantageous, but has most recently suggested that psi has evolved concurrently with what Suddendorf and Corballis (2007) call *mental time travel* (MTT), the ability to mentally model future contingencies (i.e. what is going to happen). Broughton suggests that what he calls *anomalous intuition* biases imagery and memory selection during MTT, thus producing anomalous experiences. As such, Broughton is one of a small number of parapsychologists who envision of psi as a potentially normal part of our evolutionary heritage, aiding in the struggle for survival, and certainly not supernatural in origin. Again it will remain to be seen if this type of perspective (which is very close to the perspective running throughout this thesis) will be any more useful than other theories about what psi may be for.

Taylor (2003) made a comprehensive review of needs-serving models of psi, in direct response to Broughton (1988) and his suggestion that parapsychologists must answer the 'why' of psi, rather than the 'what', 'when' or 'where', and in essence Taylor's conception of psi as needs-serving is not remarkably different from the precognition-as-death-avoidance presented in this introduction. Taylor argues that any neo-Darwinian theory of psi must make three basic assumptions: a) that has psi an underlying genetic basis, b) that there is variation of that genetic basis within the population and that c) possession of any amount of the genetic compliment that confers a psi ability improves that individuals biological fitness.

Taylor (2003) argues that psi is natural (as opposed to supernatural) but weak or rarely occurring. Like others, Taylor asks the obvious question that if psi is a natural phenomenon, why is not better or more frequently seen? One possibility that Taylor suggests is that there are 'high evolutionary hurdles' that effectively prevent psi from improving over what we currently witness. Taylor suggests that just as no biological organism has evolved wheels instead of legs (despite the fact that wheels are more efficient than legs for some types of ambulation), presumably because the hurdle to get from working legs to working wheels is too great for the incremental process of evolution to achieve. Thus Taylor argues that the evolutionary steps from no psi to a fully formed psi ability are too great a hurdle for biological evolution. This certainly appears a reasonable argument, given that evolution does not appear to have any overarching plan towards which it is groping.

Another possibility suggested by Taylor (2003) is that psi itself may co-opt existing biological systems that evolved for an independent (and entirely unrelated) reason. Thus if the psi ability co-opts the existing perceptual system, it may frequently be misidentified as a

normal percept, making identification of information as extrasensory difficult. Of course, the question remains, why has the ability to make this distinction not evolved to be 'better'? A final possibility is that the cognitive processing power may be too high and that as attention is not unlimited, processing is directed to the normal senses first, leaving little in reserve for psi. Again, if this is true, why has the ability to prioritise highly important psi information not evolved?

Taylor (2003) raises the pertinent question as to why we do not see *more* examples of individuals utilising psi in a needs based manner to avoid disaster or expedite personal reward. Although Taylor agrees with Stanford (1990) that psychological pressures may bare down on individuals such that they do not take the advantage that psi gives them, Taylor ultimately argues for a cocktail party like aspect to psi, where scanning is directed along a needs based line, not on a universal system. Taylor thus proposes the term *evolution's need-serving psi* (ENSP) for a conception of psi that is similar to PMIR. It is entirely unconscious and to avoid a cognitive overload only periodic scanning of the environment occurs. The more relevant information is to increasing an organism's fitness, the more sensitive ENSP is. In many ways Taylor (2003) and the work outlined in this thesis are very close in their opinions regarding psi (both were developed independently) and it is clear that there are a number of parapsychologists who envisage of psi as a biological and evolved process that is used by humans in their day to day existence, and is not supernatural or fantastic.

Carpenter (2004, 2005) proposes a model of psi and the mind in what he calls his 'first sight' model. Essentially the model posits that psi is a leading edge phenomena: ESP is the leading edge of the mind as it moves from one experience to another and PK is the leading edge of the mind as it moves from one intentional effect to another. This model is an attempt by Carpenter to take psi processes from something that are unusual and at times capricious, to making them normal parts of psychological processes, and in essence argues that psi functioning suggests that there is a borderline between the self and the outside world that is transitory and is neither wholly self nor wholly outside, but contains elements of the two. Psi is seen as part of the normal sensory processes, and in a way, begins the whole process. Carpenter argues that it is never conscious, but that it sets off a chain of association that inevitably sets up a relationship between what is going to be perceived consciously, and what has been detected through this leading edge process. Carpenter admits a strong parallel between this model and Stanford's PMIR model (Stanford, 1990).

Because psychological motivations are important in Carpenter (2004, 2005), and because avoidance of death can be assumed to be an important motivational factor (in humans at least), it is clear that certain death-avoidance research may be seen as supporting Carpenter's conception of psi. On the other hand, some of the anpsi research which uses very basic organisms (such as bacteria) suggest that psi can occur without any psychological processes to speak of, a fact which appears to undermine Carpenter's fundamental conception of what psi actually is. Again, only with further testing will we begin to determine which, if any, perspective might elucidate the phenomena under investigation.

Unlike most of the arguments presented in this section that do not deal directly with precognition, Hearne (1984) and his 'Group Replenishment Theory' was developed directly from Hearne's work on precognition (e.g. Hearne, 1982a, 1982c). Hearne concludes that the preponderance of females in his sample (which he argues exceeds a reporting bias that might account for the disparity) suggests that there is a biological connection between the female gender and premonitory experiences. Hearne appears to favour an evolutionary argument

based on the the mother-child dyad, but notes that the early age of premonition onset suggests a disconnect between sexual maturity and premonitions. Along with the evidence he presents in terms of a propensity for female family members to also have premonitions, Hearne argues that this suggests premonitions are innately female. Hearne also suggests that the high neuroticism of his sample may indicate that emotionality is important and that perhaps premonitions are mediated in “primitive emotion areas of the brain” (p. 765). Thus he concludes that premonitions have a biological function to confer a survival advantage, primarily through the female. Interestingly he also notes that there is a significant correlation between age at first premonition and the number of offspring which he interprets as evidence for those experiencing premonitions to be naturally more fecund than others.

The essence of this argument is that during times of great disaster, the psychic females are capable of preventing the total annihilation of the group (by avoiding and surviving the disaster) and then because of their high fecundity, quickly replenish the group. He does suggest that there is an inherent check in the process that prevents an increase in the numbers of these precognitive females and that for Hearne (1984) this explains the fact that evolution has not resulted in all humans seemingly possessing these abilities. Hearne (1989) extends this argument: that if psi exists it should be the product of an evolutionary process and aid in survival. However, in an evolutionary arms race, counter-measures may also develop. Thus he suggests that more advanced species may be *less* psychic because their overall psi abilities have been “neutralised”. With this hypothesis in mind, he suggests that simple organisms may be the best subjects for psi experiments. Hearne raises some interesting questions regarding the evolution of a precognitive ability. Without further experimentation it remains an open question whether Hearne’s speculation is a useful one.

The final argument of this section is different to the others covered, as it essentially argues that for evolution to progress as conventionally understood, psi cannot exist. Nicholas Humphrey (who held the Perrott-Warrick post) presents a powerful evolutionary argument *against* the reality of psi, in his book *Leaps of Faith* (Humphrey, 1996). He argues simply that the views espoused by many parapsychologists of an interconnectedness of humans and other life forms, of personal immortality and of a lack of individual boundaries, are anathema to natural selection: without a struggle for existence evolution cannot occur. Evolution pits every biological entity against every other and by doing so there is no room for the spiritual assumptions made by many parapsychologists.

Whilst the argument presented in Humphrey (1996) is without doubt correct for those theories that seek to explain parapsychological phenomena in terms of a connectedness or spirituality, it cannot be true that it is impossible to develop a theory of the paranormal (or a part thereof) that remains credible in the light of evolutionary theory. Those theories which treat psi as though it is an evolved perceptual ability (vestigial or otherwise) do not fundamentally undermine evolutionary theory. It must be recognised though, that most of those theories temper the limits of psi and thus it is the limitless version of psi that Humphrey is arguing against.

It is interesting that Levin (1996) recognises the import of the lack of evidence for the evolution of psi. He asks the question, simply put, that if psi is useful why is it not better than it appears to be? The fact that the evidence for psi does not suggest a strongly evolved ability, appears to be a damning criticism (others have agreed with Levin that this is an obvious question to ask e.g. Taylor, 2003). Levin considers possible responses to this argument. Firstly, he argues that we could dismiss neo-Darwinism, though he does not see this as a reasonable

response given the evidence for evolution as compared to that for psi. Another possible response to the argument is that psi *can* be observed frequently in the natural world and he provides a number of suggestions of phenomena that might have a parapsychological basis (e.g. the coordination of ants in a colony). Levin argues that although most conventional biologists imagine such difficult-to-explain behaviour has a “normal” explanation at its root, that there is the possibility that psi may be that explanation (however unpalatable it may be to those same conventional biologists). Again, Levin considers this possibility unlikely, given the undeveloped nature of psi in humans (that is, if animals *do* use psi in a useful way, why is psi functioning so very fickle in humans?).

Levin (1996) proposes ‘a more promising line’ of argument based about the physical characteristics of psi as information transfer. He argues that the evidence suggests that psi is very different from the other sensorimotor systems we are used to (for example, with regard to precognition it appears that time is not the barrier it is in other systems). Levin suggests that we may then be bathed in psi influences from around the planet, and that the “infinitesimal signal-to-noise ratio may well make it impossible to isolate immediately useful information” (p. 225). Promising though this argument appears to Levin, he states that there is evidence for useful psi information in the literature and that if such events are possible, why has evolution not improved it? One possibility, though Levin deems this implausible, is that the processing required to identify the relevant signal is far above the processing of any evolved biological system. The animal research (already covered) suggests that this isn’t so.

Another possibility suggested by Levin (1996), is that perhaps psi requires too much energy and thus it is “evolutionarily inefficient” as compared to other biological systems. Levin thinks this is a reasonable argument which possibly explains the lack of evolution. Another perspective that Levin discusses is Broughton’s views on evolution and psi (Broughton, 1988 already discussed above), though he argues that the kind of ESS suggested by Broughton should lead to an evolutionary arms race amongst individuals of the same species (and he also doubts that there is an advantage to hiding psi ability from a competing species).

Finally, Levin (1996) poses the question whether there is a moral dimension to psi, that puts it outside of the remit of ‘bad’ outcomes, and so it can only be used for ‘good’. Levin notes that this will be an unpalatable explanation to many. Levin himself argues for a dualist explanation of psi, similar to that proposed by Hardy (1965) (discussed at the beginning of this section).

It is clear, yet again, that there remains a choice between a limitless, theistic, dualistic conception of psi, that as Humphrey (1996) and others recognise is anathema to Darwinian evolution, and a more biological, more tempered version of psi, which recognises the possibility of the subtle and possibly vestigial nature of psi, perhaps a remnant of a failed biological experiment, perhaps hidden by strategy or evolutionary arms race. It is this second conception of psi (and precognition) that we believe nearing the end of this introductory examination of the problem of precognition, offers the most promise. Only further experimentation and hypothesis testing will allow researchers to begin to decide whether any of the arguments put forward have a veridical basis, or whether wholly new theories are required. This is part of the motivation for the research outlined in this thesis.

1.4 Working Hypothesis

Precognition may or may not exist. Yet, it is without doubt that there are a variety of claims that state precognition not only exists, but occurs and is demonstrable within the confines of a scientific laboratory. Premonitions have been reported since time immemorial and it appears clear that many cases of spontaneous precognition are associated with death. The idea of psi as a pre-sense or that it is needs-based is less 'far-out' than the alternative position, that is to argue that precognition is somehow connected with supernatural or theistic elements (such as a soul). Like Orme (1974), our primary assumption is that "precognition is communication between a future part of the organism and its present" (p. 358).

We have already noted that Humphrey's outright criticism that the paranormal and evolution are mutually exclusive does not necessarily apply to those theories that argue for a tempered and limited version of psi (as opposed to a limitless conception of psi). If precognition is an evolved ability, it is clear that there are numerous potential reasons why psi is not as useful as it appears it could be. Perhaps a) it is limited by biology (costly, etc.); b) it is limited by physics (many-worlds theory, a consequence of the physical property of the psi-signal, or information leaking into other dimensions); c) historical (atavistic or even an aborted fossilised pre-sense); d) psychological limitations (cognitively intensive, low in motivational hierarchy, social pressures etc.). Perhaps, as others have suggested, psi usefulness is not without cost and ostentatious displays of psi can prove almost as costly as not using it. At this stage we do not have enough understanding of the phenomenon to possibly decide.

All of the studies in this thesis are concerned with the overarching question regarding the reality of the paranormal and specifically of precognition. The primary aim of this investigation is to first capture the elusive quarry and find evidence that requires recourse to the paranormal hypothesis of precognition. There are a range of strong claims for precognition.

Given the conclusions made as part of an overall review of the evidence for precognition, the secondary aims are based on the evolutionary considerations already discussed. Does conceiving of precognition as a kind of death-avoidance behaviour improve our chances of obtaining the extraordinary evidence needed to accept paranormal claims into conventional scientific thinking (by assisting in the development of more successful methodological procedures)? The basis of the secondary emphasis is simple. That some spontaneous cases look as though they are a kind of death-avoidance behaviour and that some evidence suggests a genetic basis for this proposed ability. Ultimately, if no evidence for precognition is found, then these secondary aims (evaluation of precognition as death avoidance) will be of much less interest, as they are contingent on the assumption that precognition is an ostensible phenomenon.

Retrocausal Influences in Electrodermal Activity

2.1 Presentiment I

2.1.1 Introduction

The decision about what kind of empirical testing to begin with came from a general review of the current trends in precognition research.¹ Perhaps some of the strongest experimental findings within recent decades have come from investigation of the presentiment effect (e.g. Radin, 1997b, 2003). The presentiment effect is a precognitive phenomenon which itself falls into a class of phenomena best described as *time-reversed effects*. Time-reversed effects are essentially conventionally understood psychological phenomena that occur in the temporally *opposite* manner to their conventional counterparts. For example, in this chapter we are going to be investigating a type of physiological arousal to stimuli. Rather than investigate arousal *after* stimulus presentation (as conventional psychologists might), we are interested in how the participant responds *before* stimulus presentation. Can participants show a differential response to stimuli before onset? This is what is meant by time-reversed. We are *not* interested in any *normal* mechanism by which a significant differential response to stimuli before onset might be misconstrued as paranormal (such as expectation) except to identify confounding variables and eliminate or account for them as best we can. Instead we aspire to use the future as an impenetrable barrier which should effectively screen stimuli from participants. If then they show this differential response we can begin to question what that result means for our understanding of cause and effect.

A number of time-reversed effects have been proposed in the literature along these lines including the presentiment effect (Radin, 1997b), the precognitive habituation effect (Bem, 2003) and time-reversed cognitive interference (Klintman, 1983), to name the most widely investigated. The precognitive habituation effect is in essence a reversed mere-exposure experiment (Zajonc, 1968). The mere-exposure effect is a robust finding that participants show preference for stimuli they have previously been presented with (even if that exposure is very brief). The precognitive habituation experiments have claimed to show that participants demonstrate preference for stimuli that they *will* be shown in the future. It is referred to as precognitive habituation because aversive stimuli are *less* arousing because they will again

¹Some of the results of Presentiment I were collected as part of the requirements for an undergraduate psychology degree subsequently awarded from Goldsmiths College.

be shown to participants in the future. Time-reversed interference will be covered in detail in Chapter 3 but is claimed to be a Stroop-based effect reversed in the temporal dimension (e.g. Klinton, 1983). From a conventional perspective participants are faster at responding to the second stimulus (by reading a written colour-word) on trials which are congruent (a red block is followed by the word "red") than trials which are incongruent (a blue block is followed by the word "green"). The time-reversed interference claim is that the congruence of the trial influences responses to the first stimulus, before congruence is established. Finally, as already mentioned the presentiment effect is another proposed time-reversed effect. Where conventional scientific understanding predicts that people will show physiological arousal after presentation of emotional stimuli, Radin (1997b) claimed to have found physiological arousal *before* presentation of emotional stimuli, compared to neutral stimuli. These time-reversed interference claims have generated quite some excitement since they were first proposed and are regularly hailed as providing some of the best evidence in parapsychology (e.g. Dossey, 2008).

The presentiment effect also appears to be a good candidate as a starting paradigm as it eschews participants' conscious guesses (the standard response measure in parapsychology) and instead attempts to show through unconscious physiological responses that the participant is indeed responding to precognitive influences. Using a participant's physiology rather than relying on conscious reporting has long been seen as a holy grail of parapsychology (e.g. Broughton, 2004) partly because it side-steps many of the issues raised in the debate between forced-choice and free-response tests.

Some aspects of the biological evidence for psi were covered in the introductory chapter and it was noted that there is evidence that suggests that if psi is real that it is an unconscious ability that is relayed to the person through unusual bodily emotions and sensations. The presentiment work outlined in Radin (1997b) an attempt to build on the historical opinion that a successful physiological measure of a paranormal effect would be particularly impressive and he employed a simple methodology to investigate this very possibility: by measuring his participants' physiological responses before, during and after presentation of emotionally arousing stimuli and investigating whether the responses before presentation of the stimuli resemble those after presentation and compared responses across randomly selected calm or emotional stimuli (the calm pictures consisted of pastoral scenes and neutral objects and the emotional stimuli were violent or pornographic images).

Radin (1997b) hypothesised that precognition somehow involves the feeling of a future emotional state and that this may somehow initiate the precognitive experience. For example, a future state of feeling fear or disgust might somehow be conveyed backwards in time to influence behaviour in the present. Radin argued that the arousing stimuli initiated participants' orienting response. The orienting response is argued to be an evolutionary response to change in the environment such that it disposes the organism to respond quickly to danger or threat (flight or fight). It is characterised by a range of simultaneous physiological changes, including a rise in EDA, pupil dilation, etc. Radin argued that it is very easy to provoke an orienting response in the laboratory by presenting a provocative stimulus: not only pictures, but sounds, words, electric shocks and so on, could also be used to investigate the effect. It is important to note that this effect is not specific then to the stimuli, only that the stimulus be arousing. Radin also argued that an organism's general level of arousal is cumulatively influenced by arousing stimuli, so that the strength of the response diminishes quickly as the novelty wears off and habituation sets in. Due to this particular aspect, Radin made use of a small number of

targets, which were randomly interspersed with a larger number of control stimuli.

Radin (1997b) was unusual in the annals of physiological investigation of psi, in that it was specifically interested in using a precognitive methodology based on the argument that it might be easier for participants to detect their own future thoughts or feelings, than another person's contemporaneously. Radin identifies a further reason for beginning this particular investigation, in that a methodology which employs a precognition design can remove the need for expensively shielded testing rooms, as there is no way (other than paranormally) to receive information about a stimulus before it has been presented or generated. Radin goes so far as to state that sensory leakage is completely avoided using a precognition design. However, of course, it might be argued that because in tests of precognition the "guess" exists before the target is generated, there certainly exists the possibility that the existing "guess" could be used in an unscrupulous manner to help generate matching targets. Given the type of data used in this study, this particular concern does not seem credible.

Radin (1997b) outlines the first of four experiments into the presentiment effect. Participants sat in a comfortable chair in front of a computer display. Electrodes were attached to their hands to record EDA and a further monitor attached to measure heart rate and blood volume pulse. When participants were ready they were presented with a random target stimulus (displayed for three seconds) followed by rest periods (of between less than one second and 30 seconds). Each testing epoch lasted 13 seconds, which included 5 seconds before, 3 seconds during and 5 seconds after stimulus presentation and all physiological measures were recorded over 40 trials. Radin made use of violent and pornographic images and, because of this, restricted his choice of participants to adult male volunteers. Radin's method of analysis involved transforming the raw EDA, heart rate and pulse data, and correcting for so-called 'drift', whereby physiological measurements drift over time. Radin tested 8 participants (3 women, 5 men, despite noting that only men were selected) in a total of 260 trials, 104 were randomly selected by the computer as emotional, 156 as calm (the stimuli were drawn from a pool of 120 images, 79 of which were calm and 31 emotional). The first three people ran 20 trials in a single session, the remaining five ran 40 trials in a single session. The results of the first experiment were claimed to significantly show a presentiment effect in line with his prediction: that is, that participants in the study showed clear differentiation between the neutral stimuli and arousing stimuli in their physiological responses *before* stimuli presentation.

Interestingly, Radin (1997b) argues that the experimental design has a built-in control, and that the physiological responses after presentation of the stimulus must show the normal responses shown extensively in the literature for there to also be a presentiment response. That the normal response is found seemingly supports Radin's assertion that what is seen is tantamount to time-reversed arousal. Radin reports a significant drop in blood volume pulse for emotional targets both before and after presentation (although no effect is found for the heart rate data). In his second study, Radin had three participants contribute only 40 trials. A one second period was used instead of three seconds. Only EDA was recorded, but Radin reported similar results to his first experiment. Experiment 3 had 16 participants contribute 640 trials, using a procedure almost identical to the first experiment, except that the procedure was fully automated. The results again were similar to the previous studies. Finally, Experiment 4 tested four people who contributed 40 trials each. This method was the same as the previous experiments, except the target was chosen immediately before it was presented, rather than chosen at the start of the trial. Radin argued this meant it became a test of true precognition because during the epoch before the target is chosen there is no way of inferring what the

choice will be. The results were similar to those found in the earlier experiments. Radin concluded that he had demonstrated a paranormal presentiment effect using this methodology.

Discussing alternative hypotheses, Radin (1997b) dismisses the idea that the results are somehow due to chance and argues that the software ensures that there are no systematic artefacts providing cuing (in terms of hard drive noises, etc.) that might influence participants' responses. The possibility that the effect is due to an artefact of analysis is equally dismissed given the demonstration of normal effects after presentation, alongside the paranormal effects before. Examining the order of his targets, Radin employs a number of statistical tests to demonstrate that the order of presentation for all stimuli was random and that the participants were not performing better than chance because of any bias in the order of presentation. Radin also discounts the idea that some anticipatory or expectation response might be the explanation for the results including unconscious strategies that might be employed by participants after a given type of stimulus. This 'gambler's fallacy' hypothesis involves participants reacting in specific ways after the presentation of different stimuli. For example, after presentation of an unexpected violent image, participants *may* begin to expect another presentation of a similar stimulus and therefore their physiology may reflect this anticipatory preparedness, rather than be a consequence of a presentiment of a future state. (Only if there was a systematic bias in the source of randomness could this type of artefact produce the same kind of differential effect hypothesised.) Radin claims that such normal explanations cannot fully explain the effect he finds. Without replication and further testing, it is difficult to determine if Radin's claims are overstated and unwarranted. Radin again reports statistical analyses that he claims show that no such bias can account for the findings. Radin's conclusions are not tentative. He reports evidence for a robust and demonstrable paranormal effect, which not only fits neatly into a certain interpretation of the previous evidence, but which Radin takes as evidence for some kind of quantum non-locality of mind and space.

The claims made in Radin (1997b) are certainly impressive and have the potential to influence much of psychology. It is crucial to establish if conventional forward-in-time effects have an accompanying time-reversed influence. Much of the best evidence for the paranormal points to the existence of this very type of effect. It seems to be a direct experimental test investigating the experiences that were described in Chapter 1; for example, those surrounding the Aberfan mining disaster. Where in the spontaneous reports there are cases of people seemingly saved from death by a hunch or intuition of a negative event in the future, Radin's presentiment effect seems a direct analogue: that people can be affected by a future state, albeit unconsciously. This makes the presentiment effect a good target for a first experiment in this thesis.

Bierman (1997) attempted to replicate Radin (1997b) and found that the response before being shown a highly emotional visual stimulus was significantly greater than the response before presentation of calm pictures ($p < .007$).² While certain normal explanations are evaluated in the discussion, Bierman concludes that the paranormal explanation of anomalous mental cognition is the only satisfactory answer. Bierman and Radin (1997) presented their findings outside of a conventional parapsychology journal, in *Perceptual and Motor Skills*. In this brief paper they summarise their joint findings and yet again claim that the effect is strong and has implications for other researchers working in similar paradigms. The findings outlined in their 1997 paper are expanded upon in a paper self-published on the internet in 1998 (Bierman &

²Of course the 'response' is the participants' physiological activity in the present, and the use of the word response does not imply that participants are actually responding to future stimuli.

Radin, 1998).³ The 1998 paper summarised three different studies of the presentiment effect. Study 1 was a straightforward replication using different materials and equipment, which is claimed to show a significant difference in response before the target stimuli compared with neutral stimuli. Study 2 involved a slightly different methodology, whereby uninformed participants were shown 10 images, one of which would be a target stimulus. This second study was an attempt to reduce expectancy effects, that might reasonably be expected to occur by over-exposing participants repeatedly to the emotional stimulus, and maximise the surprise element of the target stimulus. Although a trend was identified, no significant effect was reported.

Finally, Study 3 used the same participants as Study 2, who were, after participating in Study 2, no longer uninformed about the potential for stimulating images to be shown, and were tested using a very similar methodology to that used in Study 1, with some attempt to improve performance using the most stimulating images from the earlier studies and so on. Again, another trend was reported in the data, but the difference between the target images and the non-emotional images was, as in Study 2, non-significant. Bierman and Radin conclude that although not consistently shown across the different methodologies, there is certainly evidence for what they call “macroscopic time symmetry”, which they believe is a better explanation than the idea that different anticipatory strategies could account for any pre-stimulus response. This time-symmetry is predicted to be related to the beginning of conscious experience rather than stimulus onset.

Radin and Bierman have both argued that evidence for the presentiment effect should be identifiable in normal psychological research that employs similar methodologies. An experiment by Bechara, Damasio, Tranel, and Damasio (1997) is suggested as an example of an occasion where the presentiment effect may contribute towards full understanding of the situation. Bechara et al. present evidence to support their claim that normal people employ two systems when they make a decision, a conscious decision making process, and an unconscious process, which biases the conscious system according to previous experience. The implications of this study are clear: that normal participants have unconscious decision making mental systems which are capable of physiologically influencing the body, and thus nudging along the conscious faculties into a more efficient and effective decision making process. Bierman and Radin (1998) believe that psi processing or precognition exists to supplement those elements proposed by Bechara et al. and that they influence the unconscious processing in a similar way, which is recognised by consciousness as a hunch or feeling.

The main objective of this first study, Presentiment I, was to test the claim that there is time-reversed arousal before presentation of an emotionally stimulating image and it would employ a similar methodology to that used in the previous presentiment effect research. Participants’ skin response would be measured and any anomalous difference or variation investigated.

The secondary focus of this thesis is an attempt to investigate whether evolution might suggest alternative methodologies which might prove more successful than those already employed in previous research. As Radin (1997b) suggests, there is a body of precognitive experiences and empirical reports that seem to suggest that precognition might be related to future fear or death feelings, somehow impinging themselves on a present state. The

³Given that this research was in part stimulated by papers pre-published on the internet and subsequently edited, it is difficult to identify an exact time-line of research and draft publications, and as far as possible the historical account is given, that is, papers are discussed in the order that they have been produced, in an effort to understand both the development of testing and the influence of researchers upon one another.

emotional stimuli used in the earlier presentiment research fall into two categories: violent or pornographic. That violent or pornographic images cause conventional physiological responses has been widely studied in normal psychology. However, concerns were raised early on with respect to the ethical considerations involved in presenting participants with extremely graphic images of horrific violence or strong pornography. Although previous researchers had not reported any concern over the stimulus set used, it was hoped that by a slight adaptation of the procedure, that the claim could be tested in a comparable but less ethically contentious manner. Spider stimuli were quickly identified as being directly amenable to this particular purpose. Spiders have been shown to elicit the same kind of arousal in physiological responses to violent or pornographic images in conventional experimentation. They are also arguably less ethically challenging than the other stimuli types because all participants will be randomly exposed to the presence of spiders as part of their daily lives, whilst the same cannot be said for hardcore pornography and extreme violent images. Thus, though spider images may well induce feelings of anxiety, it was anticipated that such feelings would be short-lived and similar to that experienced during a random encounter with a spider.

Geer (1966) reported a simple experiment whereby students were classified into high-fear and low-fear groups (in relation to their fear of spiders) and their galvanic skin response (GSR, a measure comparable to EDA) measured in relation to being shown images of spiders and snakes. Geer found that those students who rated themselves as being highly afraid of spiders showed a greater increase in GSR in response to the spider picture and for a longer duration than the low-fear group. Wilson (1967) followed up Geer's research comparing ten subjects who reported an intense fear of spiders and ten who did not. He found that using an index from their GSR response to spider and neutral stimuli, he could perfectly differentiate between those who reported fear and those who did not. That such a differentiation can be found by simply measuring skin conductance is relevant to the question as to whether replacing previous types of stimuli with spider images will continue to elicit the same response in participants and it is worth noting that Radin (1997b) argued that the presentiment effect was not dependent on the stimuli used, but on the response that it caused.

Katkin and Hoffman (1976) were interested in investigating a predicted gender difference between men and women in their response to spider stimuli. Specifically, based on Geer's work, Katkin and Hoffman predicted that women would show a more pronounced GSR reaction than men, although found only fearful participants, irrespective of gender, showed prolonged autonomic arousal when first presented with an image of a spider. This is encouraging as it is evidence that gender differences need not be predicted when investigating a forward-in-time or backwards-in-time arousal effect. Prigatano and Johnson (1974) reported a similar finding to Geer, but concluded that it was probably a reflection of the intensity of the stimuli. Rather than it being specifically related to spider-fear, they argued that the effect represented general autonomic arousal.

Directly relevant to the secondary working hypothesis of this thesis, that precognition, if real, might represent death-avoidance behaviour, is the work of Öhman and Soares (1993) who investigated EDA responses to a set of stimuli, including spider images. Öhman and Soares argue that there are evolutionary reasons to believe that organisms have evolved to associate anxiety and aversion with potentially deadly situations through the simple argument that any behaviour which promoted getting into a deadly situation would presumably lead to those genes becoming extinct leaving more "cautious genes" to propagate. Seligman's (1971) preparedness theory of phobias suggested that human phobias reflect an evolutionary pressure

which has resulted in avoidance as the main adaptive outcome and that these are developed through a kind of selective association akin to Pavlovian conditioning.

Mineka (1992) reported conditioning work with rhesus monkeys which ultimately demonstrated that the animals were capable of learning a wide range of phobias against a wide range of objects, but that evolutionary relevant fears (such as a fear of spiders and snakes) could not be subsequently extinguished. Öhman and Soares (1993) found significant differences between EDA responses to fear-relevant stimuli (i.e. spider images presented to spider phobics) and fear-irrelevant stimuli and that furthermore this is a robust finding with implications for arguing an evolutionary basis for a fear of spiders. Öhman and Soares make the argument that a fear of spiders has evolved because of the deadly consequences spiders could have had when our ancestors encountered them. As such, a fear of spiders can be argued as a specific type of fear of dying. Therefore spider stimuli represent a potentially excellent stimulus-type within the context of investigating the presentiment effect.

Returning to parapsychological research, there is some research that links high anxiety with above-chance scoring in a psi-task. For example, Freeman and Nielsen (1964) were inspired by the findings of normal psychology that there is a significant correlation between EDA and the emotionality of words (McCurdy, 1950). Freeman and Nielsen used a novel precognition task that involved participants rating a target word (in terms of liking or disliking) in one of five spaces on a score sheet and a hit was scored if ratings were written in the to-be-chosen space. They predicted that participants would score higher on words that were rated as liked than disliked, but found that only the high anxious group produced above-chance scoring ($p = .02$).⁴

In our regime of testing, it was hypothesised that for those people who reported a fear of spiders, they would be expected to show a strong pattern of arousal after presentation of the stimuli and, by that reasoning, one might predict that presentation of the stimuli might be accompanied by a paranormal, time-reversed effect. As well as providing an ethically less problematic methodology, the use of spider stimuli would also provide a kind of control group. It might be expected that even those who rate themselves as being not afraid of spiders would show some arousal to spider images, but that their reaction will be less pronounced than those who rate themselves as being afraid of spiders. By that reasoning, there should be a similar time-reversed effect in both groups but it should be more pronounced for those fearful of spiders, if the presentiment effect is real.

Aside from using spider images, Presentiment I would use a one-trial methodology as employed by Bierman and Radin (1998), whose Study 2, made use of just such a one-trial methodology. In contrast, Radin (1997b) had used twenty-five participants over twenty-nine trials and therefore each participant viewed a different order of neutral and target pictures over the twenty-nine trials. As such, it is certainly possible that any bias in the random presentation of stimuli, or certain kinds of strategies on the part of participants, may have influenced the results. Specifically, if the presentation of the images is non-random (for example, there may be too many alternations in the sequences of emotional and neutral images), this bias may be learnt by the participants and their reactions may be based on normal information rather than

⁴We also have reason to suspect such a relation because of a result obtained in a study omitted from this thesis due to word-length requirements. Over 400 respondents filled in an anxiety questionnaire (the Taylor Manifest Anxiety Scale, Taylor, 1953) and this score was correlated with the paranormal experiences scale (administered as part of the Anomalous Experiences Inventory, Kumar, Pekala, & Gallagher, 1994). We obtained a small but highly significant Spearman's rank correlation ($r_s(414) = .225, p < .0001$). This possibly suggests that high anxiety may result in either more active monitoring of internal signals (including the psi-signal) or a lower threshold for psi, though other more prosaic explanations may also explain the result (e.g. intolerance of ambiguity).

paranormal phenomena. In addition to the possibility of learning, there is the possibility that strategies may be employed by participants whereby they might expect one arousing image to be followed by another (that is, in a sequence which comes, neutral-neutral-emotional it is certainly possible that there might be a general leaning there for participants to expect this sequence to be followed by another emotional image). That this should have no systematic effect if the source of randomness is true, must be weighed against the potential for relaxation effects, general arousal of participants, and expectancy or gamblers' fallacy to have a small effect on the data, which might be interpreted as being a paranormal effect.

Because multiple trials introduce a great many potential confounding factors into this already difficult research area, a one-trial 'surprise' methodology may be a better test of the claim. Having only one target event randomly presented to each participant, also seems more in line with spontaneous case reports, where very few involve multiple correct precognitions within a short time frame. Therefore, it was hoped that employing a one-trial procedure would eliminate any implicit learning effect and would be, for those people who reported a fear of spiders, an event which should produce some significant feelings of fear in them. Although having a small number of trials might be criticised at having reduced the potential to find the paranormal effect, the number of participants to be tested in the experiment was increased over the previous empirical tests, and the use of the spider-fear element also provided a further novel adaptation to this type of experiment.

The hypothesis of Presentiment I was therefore that, in line with previous research, those participants who rate themselves as being afraid of spiders will show a larger physiological response *before* the onset of the spider image than those participants who are not afraid of spiders.

2.1.2 Method

2.1.2.1 Design

Presentiment I employed a mixed design where participants' electrodermal activity (EDA) was the dependent variable (specifically a 20-sample period 2.5 seconds to .5 seconds before stimulus onset). The one between-subjects factor was spider-fear (fear versus no-fear).

2.1.2.2 Participants

Data was collected until 60 participants successfully completed the task.⁵ The computer software provided an automated validity check of EDA values and participants were dropped from the study if their data did not pass this check (which involved ensuring none of the values indicated a failure in data collection, usually the result of an electrode disconnecting). Most participants were undergraduates at Goldsmiths College, University of London. There were 14 males and 46 females with an age range from 16 to 35 years with a mean of 21.9 years and a standard deviation of 3.6 years. Participants' spider-fear was recorded thus: 32 fear and 28 no-fear. Participants volunteered and were thereby self-selecting.

⁵It should be noted that a redundant "control" condition was run (with a further approximately 60 participants) but is not discussed further in the final version of this chapter. Briefly, the control condition was identical in all ways to the experiment described, but replaced the spider image with another neutral image. As we would not expect to find any paranormal result in the control condition it is redundant, although would have been of some interest *if* the results had provided evidence suggesting a presentiment effect. Given the results obtained, further discussion of this condition has been dropped, mainly on the basis of clarity.

2.1.2.3 Materials

An EDA measurement device was constructed especially for the study by a departmental research technician, RD.⁶ The equipment consisted of two inputs for the connecting leads to be inserted into, which were in turn connected via circle clips to the electrodes (Biotrace centre-snap). The EDA device had a parallel port output which conveyed participants' EDA, with a sampling interval of 50 ms, and a sample rate of 20 Hz, to a desktop PC. Data was measured as resistance in kilohms and then converted to microsiemens for analysis.

A short computer program was also written by RD in Visual Basic to present the stimulus material in a randomised order. The program distinguished between two blocks: in the first block five neutral pictures were randomly presented. After the fifth picture the second block began, where there was random presentation of four neutral pictures and a spider picture. Each image was presented for 1 second, with a 7 second blank screen before picture presentation. Images were selected by RD from the IAPS database Lang, Bradley, Cuthbert, et al. (1999). Randomness was achieved using standard pseudo-random algorithms and the internal clock as seed generator. The software was also able to log the EDA from the device. The software was programmed to load all of the pictures into the memory to prevent the noise of the hard-disk spinning as the pictures were loaded, thus preventing sensory leakage. It was considered possible that had images been read directly from the hard disk, that differences in location, size, fragmentation, etc., may have resulted in a physical signature before presentation, although given that the spider image was shown only once, this was considered only a minor possibility, though by loading from RAM it was hoped to avoid it.

The Fear of Spiders Questionnaire (FSQ) (Szymanski & O' Donohue, 1995) was used as the means of determining fear of spiders. The scoring of this measure involves assigning spider-fear to participants who answer yes to one or more questions. Szymanski and O'Donohue reported that the FSQ is able to differentiate spider phobics from non-phobics ($p < .01$) and so it was deemed useful in the current context.

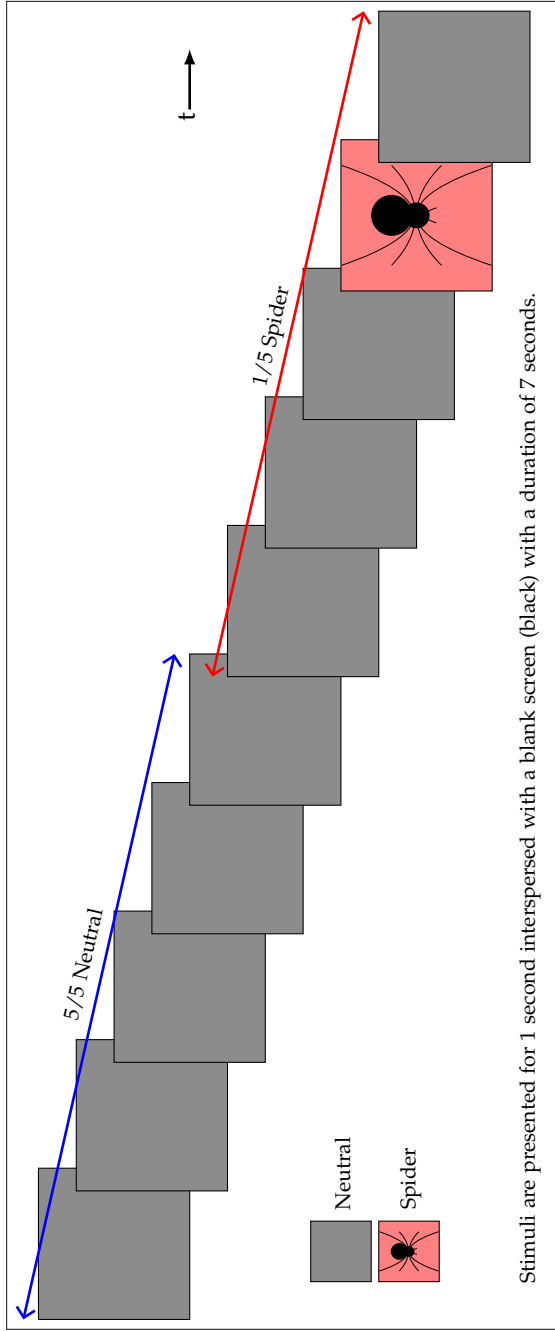
An additional small program was written, again in Visual Basic, to extract the relevant data from the raw data. This program was run after all data had been collected and was the means by which the large pool of physiological data was reduced to the specific period of interest. The data at 2.5 to .5 seconds before onset of the spider stimulus was extracted and a mean value produced for each participant.⁷ This mean was logged to a data file for analysis. It is important to understand that the extracted period of EDA covers a period of time where the computer display is blank for all participants and that the data is pre-spider presentation (that is, none of the physiological data analysed contains the conventional physiological response to the emotional spider stimuli). To further clarify how the specific epoch of analysis relates to the spider presentation see Figure 2.1. Because the methods of constructing the experimental data were different to that employed by Radin (1997b), no attempt was made to expressly repeat his analysis.

2.1.2.4 Procedure

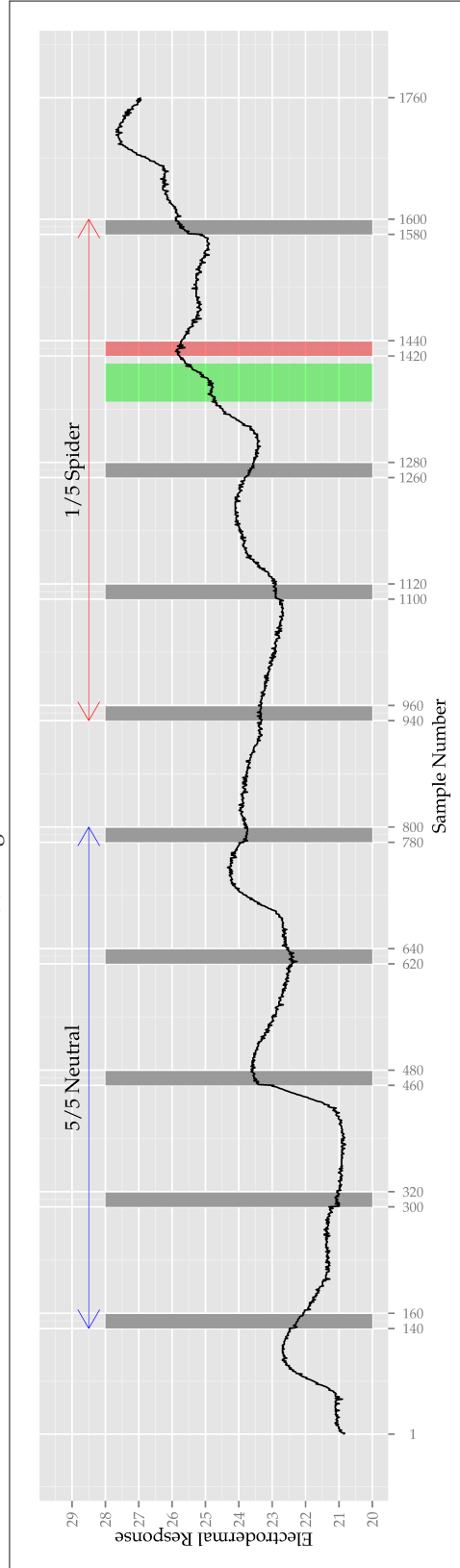
Participants were recruited from the local area on the basis of showing interest in taking part in a short psychological experiment. Once a brief cover-story had been told to the participants (that their skin conductance would be measured whilst watching some neutral images, thereby

⁶As the EDA device was custom made equipment, it was necessary to rely on RD to ensure that the equipment functioned as required.

⁷It is not clear from the records why this epoch was used.



(a) Diagram of Procedure



(b) Example physiological response

Figure 2.1: Diagram of procedure and example physiological response. (a) shows a simple diagram of the procedure. Ten images are presented to participants. The first 5 are all neutral (5/5) and one of the second 5 is a spider image (randomly presented as one of the images 6–10). (b) shows an example physiological response. This plot presents EDA for each sample taken during the experiment. The red bar denotes the spider image and the green bar is the period used in the statistical analysis (Presentiment I). This plot clearly shows that the stimuli were presented for 1 second (covering 20 samples), interspersed with a blank screen of 7 seconds (covering 140 samples).

establishing a baseline) the electrodes were first attached to the participants, one at the base of the index finger and one at the base of the middle finger of the left hand, palm facing upwards. The leads connecting the EDA measurement device with the electrode had centre clips, which were attached to the centre snaps of the electrodes. The participants were then given the instructions for the study which involved being asked to sit back, relax and watch ten neutral pictures to allow a baseline reading before the experiment. This set the cover story, as being uninformed to the nature of the study was important. When the participants had finished reading the instructions and had confirmed understanding, the experimenter began running the software, after which the experimenter left the testing cubicle leaving the participant on their own. The participant was presented with a seven-second blank screen followed by a stimulus picture presented only for one second. The first five pictures were randomly selected using standard randomisation procedures (5/5 neutral). After the fifth picture a spider image was presented at a random position from six to ten (1/5 spider). The software ended after presentation of a final blank screen after the last picture.

Once the software had finished, the experimenter re-entered the cubicle removed the electrodes and presented the participant with the short spider-fear questionnaire.⁸ After the questionnaire had been completed the participant was debriefed as to the aims of the study. Although the experimenter was blind to the individual participant spider-fear status, he was not blind to the fact that the participants will be exposed to a spider stimulus. Figure 2.1 illustrates how the experiment was set up, and provides an example physiological response (and with it the timing and spread of the stimuli and the period extracted for analysis).

2.1.3 Results

Summary statistics for Presentiment I are presented in Table 2.1 (this is the means and standard deviations for the two fear groups, for the extracted period before spider image onset). Before data collection the expected presentiment result was that those who rated themselves as spider-fearful would have a higher EDA score before presentation of the spider stimulus, than those who were in the no-fear group.⁹

Table 2.1: Summary statistics (means in microsiemens and respective standard deviations) for the two spider-fear groups for Presentiment I

Spider-Fear	\bar{x}	σ
Fear	17.7	5.7
No-Fear	21.7	9.3

It is important to stress that for Presentiment I no attempt was made to establish whether there was a significant conventional response to the spider stimuli (as would be needed for

⁸Administering the FSQ after participants have been presented with spider stimuli could mean that responses are biased by the preceding exposure. Because the questionnaire has to be administered at one point (either before or after) and because it was deemed important for participants to be uninformed as to the spider aspect of this experiment, it was pragmatically decided that the measure would be given after the experiment. It was also decided that given the very general nature of the measure (that one or more affirmative score would put the participant in the spider-fear group) that a participant who has undergone the experiment and still does not answer any of the questions in the affirmative is very likely to be the very sort of no-fear participant being sought.

⁹Because of the delay in publishing the results of these two presentiment experiments, it is worth noting that although the contrast of analysis was most certainly planned before data collection, it is true that the data could be analysed via a range of techniques and previous discussion of this data (as part of conference presentations, etc.) have used different statistical techniques, for example parametric versus non-parametric, not in an attempt to trawl the data for significant differences, but to use the most appropriate statistical technique in this application.

there to be a significant presentiment response). As the raw data for Presentiment I are now irretrievably lost, it is not possible to establish the presence of a conventional differential response based on fear to the spider stimuli. The decision *not* to investigate the presence of a conventional effect was based upon a desire to concentrate primarily on the paranormal aspect of the experiment, though with hindsight it is regrettable that such an investigation was not undertaken.¹⁰ The difference we see between the fear and no-fear group before presentation is in the wrong direction to that expected and is non-significant: $t(44) = -1.96, p = ns$ (one-tailed).

Subsequent deliberation on the procedure of Presentiment I has highlighted a major problem with the analysis of this experiment. Namely that by pooling the participant scores as means, we lose the ability to detect any small influences in the data (as a psi influence would be expected to be). Thus, although it was hypothesised that our two fear groups would have a mean difference that differed statistically from one another, it would have been much better to standardise individual participant scores, to account for individual differences. This problem was addressed in Presentiment II.

Finally, some brief points about power should be raised. This study was one of the earliest attempted independent replications to test a paranormal effect that was claimed to be robust and replicable. It essentially repeats the procedure of Bierman and Radin (1998) Study 2, where they tested 32 participants in a single-trial methodology. Although no significant result was reported in this study, the authors do argue that there is a presentiment trend in the data. Given the problems identified in the introduction, this methodology was expected to be the least problematic in terms of confounding factors which made it the most appealing despite the obvious low power of the design. Given Bierman and Radin's results, we developed a methodology that would hopefully *improve* the experimental results by comparing two fear groups who *should* respond in a differential manner to the stimulus (this doubles the number of participants used in Bierman and Radin's study). Thus we estimate our study to have more power than Bierman and Radin's Study 2.

Because we have not retained post-stimulus onset responses, it is impossible to determine whether our participants showed the appropriate conventional responses (i.e., that the fear group have a larger physiological response to the spider image than the no-fear group). Using the appropriate values from the next study (Presentiment II) which is identical to this study in many ways, we can assume that the participant response will be similar which gives us a power rating of approximately .43 for the conventional effect based on the differential response between the fear and no-fear group. This is obviously low power. For us to have power of .80 we would need a total of 172 participants for this contrast (compared to the 60 we currently have). As the presentiment effect is expected to be smaller than the conventional effect our power to detect an even smaller effect is reduced further.

Given that our results are in the opposite direction to that predicted (i.e., that the fear group have scored *less* than the no-fear group, it does not make sense to consider the difference as being due to presentiment. The *pwr* package of R (Champely, 2012) provides a power test based on the t-test (*pwr.t.test*) given the parameters of this research. Bierman's reanalysis of the Öhman data using spiders and snakes (Bierman, 2000) gives an effect size of approximately .156 which if we were using that value to determine the number of participants required to establish a presentiment effect today, would require approximately 500 participants in the two fear groups. Given the loss of the participants from technical issues found in both Presentiment I and II, recruiting and testing an even larger number of participants would represent a very

¹⁰This oversight was rectified in Presentiment II.

difficult and time-consuming process outside of the scope of a regime of testing such as this. Increasing the numbers of trials would allow us to increase the power of the study without the need to recruit such a large number of participants, but a multi-trial methodology brings with it other problems and confounds.

2.1.4 Discussion

No parapsychologist expects to be able to produce paranormal results on demand. If it were so easy, the field of parapsychology would have already demonstrated the existence of psi, and there would be no controversy surrounding parapsychological research. Despite Radin's claims of a robust and replicable effect, there was always the possibility that finding *no* statistically significant result was likely given the elusive nature of psi, and that therefore, despite much effort, it may be difficult to come to any conclusions. Although it may be argued that the study lacked sufficient power to detect the small hypothesised effect, it is true comparing the two spider-fear groups is a novel way of testing for this type of influence and one which has the potential to elucidate the nature of the phenomena, once established. It is also true that our methodology compares with previous research, and attempted to specifically enhance the presentiment effect. Although a null result from a low-powered study is generally considered "scientifically useless", the fair conclusion in regards to this null result must be that we have failed to find evidence for presentiment.

On the basis of the initial claims evaluated in the introduction, there at first seemed, if not strong evidence, certainly interesting evidence that humans may be able to react in the present to emotional events in the future. Although much of the previous research has used pornographic and violent images to elicit an emotional response in their participants, these were considered ethically dubious to use within this research framework and it was hoped that based on previous normal findings (that spider stimuli elicit physiological responses in those who report a fear of spiders), that replacing the violent or pornographic images with spider images, might provide, not only an ethically more palatable procedure, but test the secondary hypothesis of interest within this thesis (of whether evidence for precognition is evidence for an evolved death-evasion faculty). In hindsight it may have been useful to run a small evaluative test as to whether the ethical claims being made are substantiated by participant preference. We could easily have gathered ratings of all types of stimuli (spider, violent and pornographic) and determined whether the spider images are rated as less psychologically distressing than the other types of images.

Another noteworthy point is that despite the strong claims of the earlier presentiment work, and the way that the research fits well within currently established precognitive empirical studies, this current experiment highlights a range of potential difficulties, which immediately demote the presentiment effect as a candidate for the best evidence for a paranormal claim. Specifically, the use of a physiological response measure requires a technical proficiency to maintain consistent standards during testing. Some 40 participants had their data lost, mostly through mechanical error, with electrodes detaching, but some due to the individual's physiological constraints (e.g. skin was sweatier than normal, etc.). Methodological problems such as these have not been reported in the presentiment literature and yet, as problems that dog most psychophysiological research (conventional or time-reversed), it is important for researchers to report whether they had such methodological problems and how they were dealt with. It is also a question of whether the physiological response provides a certain amount

of 'wobble-room' in regards to the analysis of results. In one of his studies, Radin reports correcting participants' scores for physiological drift, and yet there are any number of methods of tidying up data or statistical methods that could be employed in this type of research. In this particular study it would have been of interest to examine both the physiological responses to the spider stimuli (conventional) and to examine the before-stimulus response for a wider time-frame than was employed. This was not done primarily to prevent fishing expeditions and the data to be analysed was planned pre-experimentation. However, the wide scope to analyse and reanalyse a seemingly varied physiological measure leaves any researcher open to the claim that they have cherry-picked significant trends from their data, from a wider, non-significant pool.

Given our conclusions regarding the low power of this study, it is not possible to claim that the presentiment effect is not veridical. However, it was hoped that by replicating the experiment as closely as possible, more might be learnt about the so-called presentiment effect.

2.2 Presentiment II

2.2.1 Introduction

Although Presentiment II was to be a direct replication of Presentiment I, it was thought pertinent to review any relevant literature that may have been published in the intervening period. Bierman (2000) reported three directly related presentiment studies, each using data that had not been collected as part of a parapsychological investigation, but were conventional psychophysiological studies. Study 1 involved re-analysing a study where participants' skin response had been measured during presentation of emotional images, with some being images of snakes and spiders. Although a suggestive effect was found, participants were not split into fear versus no-fear groups and the effect reported overall did not reach statistical significance. The second study involved making use of a gambling-based methodology, where participants' skin conductance response was measured whilst playing a gambling simulator. Bierman claims that overall significance for the gambling study is suggestive (not quite reaching the .05 α level). The third data set re-evaluated by Bierman was an emotional priming study, where participants had their skin conductance response measured whilst viewing various different types of stimuli. This last study was reported to be significant ($p = .043$). Bierman's evaluation of the research findings is somewhat optimistic given the overall lack of paranormal effect found in the data and Bierman goes so far as to suggest that precognitive influences may well be common in normal data sets. That Bierman reports selectively dropping data from his analysis (based on various criteria) suggests that there was a certain amount of post-hoc fishing within the data sets to search for any potential effect. If this is indeed so, the fact that the presented findings may represent the best of any number of potential significant or non-significant statistical results is certainly not encouraging.

Schmidt and Walach (2000) reported their review and meta-analysis of EDA studies as used specifically in tests of remote staring and studies involving intentional real-time changing of another volunteer's physiology (so-called direct mental interaction with living systems or DMILS). Although they did not specifically investigate the presentiment literature, they note that many of their conclusions can be applied to studies of presentiment. They argue that the DMILS literature itself is not as homogeneous as they had expected and that there

is a large degree of variation between researchers in both practical matters (such as climate control in the testing location) and definitions of terms and standards (such as what exactly is being measured and how long after electrode attachment is data collection started). Schmidt and Walach find the entire body of research essentially flawed: none mention the standards of psychophysiological research, most do not employ standard methodologies and so on. All parapsychological research that utilises a physiological measure is considered potentially flawed to the point where their conclusions can be dismissed by the wider scientific community and they urge parapsychologists to make use of the standards set in psychophysiology, if there is to be any chance that their findings will be accepted. Schmidt and Walach's comments also apply equally to the presentiment research (including these current studies) and draw into question any possibility of any findings in the area being accepted as evidence for precognitive arousal.

Despite there being little evidence to support further testing of the presentiment effect, a second experiment was undertaken.¹¹ Although the one-trial methodology has not yet provided any evidence for a precognitive influence (which is vital if we are to learn anything about precognition), it was hoped that by again employing the simple methodology across a relatively large number of participants, that if the presentiment effect is as replicable as Radin at least suggests, some hint of it should be found. Therefore the main focus of Presentiment II was, as with Presentiment I, to measure participants' EDA before the presentation of a spider image and determine if there is a statistically significant fear-based difference.

2.2.2 Method

2.2.2.1 Design

The design of Presentiment II was very similar to that of Presentiment I. Rather than use the raw EDA scores as was done for the previous experiment, in Presentiment II we are interested in a comparable measure based on deviation of EDA over a set period of samples. There was one between-subjects factor: spider-fear (fear versus no-fear). Unlike Presentiment I, raw scores have been retained allowing us to examine the data for a conventional post-exposure response. Thus although we are primarily interested in any potential presentiment influence in the data, we also have a secondary interest in establishing whether there is a conventional post-exposure response.

2.2.2.2 Participants

As with Presentiment I, some of the participants who were recruited had their data rejected by an automated validity check, and thus the experiment was run until there were approximately 30 participants in each fear group. Participants were drawn from the student population at various levels of psychology study at Goldsmiths College and also from visitors to the Museum of the Unknown.¹² There were 27 males and 33 females. Age ranged from 14 to 60 years, with

¹¹It should be noted that the primary motivation for conducting Presentiment II was to randomise a redundant condition variable (some participants not included in the analysis described in this version of the chapter, were not presented with a spider image, but with another neutral image). However as with Presentiment I, no paranormal influence would be expected for participants of either fear group in a condition where no spider image is presented. Had a significant presentiment effect been discovered, this "control" condition would have provided an interesting contrast with which to compare responses (comparing those who see the spider image, and those who do not). Discussion of this aspect of the experiment has been removed for the sake of clarity.

¹²The Museum of the Unknown was one of a series of temporary museums housed on the South Bank of the river Thames in London.

a mean of 26 and standard deviation of 8.3. Participants' spider-fear was recorded: 29 fear and 31 no-fear.

2.2.2.3 Materials and Procedure

The materials and procedure were identical to those used in Presentiment I.

2.2.3 Results

Rather than merely repeat the analysis as conducted for Presentiment I, it was decided to engage in a more useful investigation of the data. Where in Presentiment I, participant raw EDA scores were extracted for the appropriate period before presentation and then pooled according to spider-fear group, in Presentiment II analysis would be conducted upon a standardised measure, akin to that used by Radin (1997b) in his research.

For this analysis a 300-sample period was used, covering the 7 seconds before (140 samples), the one second during (20 samples) and the 7 seconds after (140 samples) presentation of the spider image. The next task was to transform the raw EDA scores into a standardised measure, whereby for each participant's EDA score the first sample was subtracted from the 300 samples, meaning that the scores became a measure of change over the 300-sample period (beginning at zero). Figure 2.2 shows the respective 300-sample period for the two fear-groups. Examination of this plot shows that both spider-fear groups are seemingly responding to the spider stimulus, and that the fear group's is the stronger response.

Because the 300-sample period includes both before presentation and post-presentation EDA, it is necessary to extract certain periods from this data for the purpose of our analysis. From this plot then two periods of interest were identified: the pre-stimulus period (where we believe presentiment might be found) was taken from sample 85–110 (and thus lasted for 1.25 seconds). Next the post-stimulus period was taken from sample 185–230 (lasting a slightly longer 2.25 seconds). Finally, it is important to note that at the start of both periods (sample 85 for the presentiment period and sample 185 for the post-stimulus period) we treat these samples as the first, and subtract the participant EDA at this sample from the rest of the period sample. The result of which is that for all participants each sample period begins at 0 and this means that any variation that might be due to confounding influences does not result in the samples starting at different values. Summary statistics for both fear groups for both before spider presentation and after spider presentation are presented in Table 2.2.

Table 2.2: Summary statistics (means in microsiemens and respective standard deviations) for the two spider-fear groups, before and after spider presentation for Presentiment II

Spider-Fear	Before/After Spider	\bar{x}	σ
Fear	Before	0.06	0.31
	After	0.51	0.90
No-Fear	Before	-0.04	0.12
	After	0.22	0.53

Statistical analysis of the samples is simple. To examine any possible presentiment influence in the data we conduct a two-sample t-test comparing the mean change in EDA (calculated as described in the previous paragraph) for each participant, across the two spider-fear groups (fear vs no-fear) for the *before* scores. The result of that analysis was non-significant: $t(36) = 1.62$, $p = \text{ns}$ (one-tailed). A directly comparable t-test was conducted for the post-

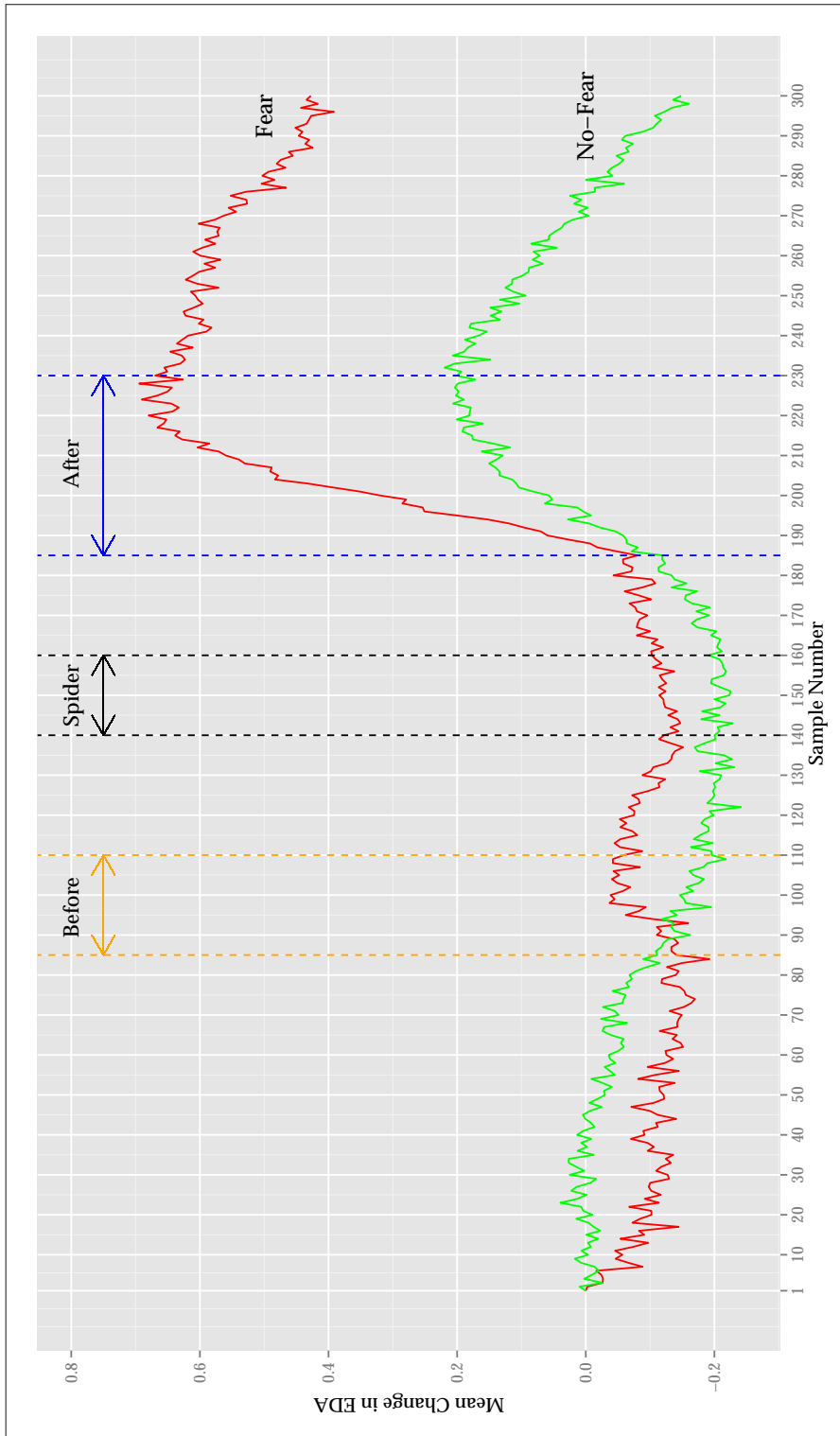


Figure 2.2: Line plot of the mean change in EDA for both spider-fear groups over the 300 samples (Presentiment II). The before period (presentiment) is bound between the orange dashed lines and the after period (post-stimulus response) is bound between the dashed blue lines) and the duration of the stimulus is denoted by the black dashed lines. This plot shows that the spider-fear participants have a larger post-stimulus response than the no-fear participants.

stimulus (or after) responses again across the two spider-fear groups. Despite the graphical evidence suggesting a differential response across the two groups, this t-test was also not significant: $t(45) = 1.5$, $p = ns$ (one-tailed) (although with $p = 0.05731$, it could be argued that there is the suggestion of significance). Clearly then, whilst there is no statistically significant evidence that the two-fear groups have responded in a differential way *before* presentation of the spider stimuli (presentiment), there is graphical evidence and a suggestive statistical result which seemingly shows that the spider fear groups are reacting in a differential way to the image of the spider.

Finally, issues of power must be considered. As was noted for Presentiment I, this study is based on a study by (Bierman & Radin, 1998), who tested 32 participants in a one-trial methodology. Although they did not find a significant result in their data, they do identify a presentiment trend. Our presentiment studies attempted to build on their Study 2 and improve the results. It was hoped that by including a fear element and comparing the behaviour of two fear groups that this would tease apart the presentiment effect (this doubled the number of participants). Thus it must be true that our study is *more* powerful than the previous equivalent Bierman and Radin study, despite our conclusions below that the study has relatively low power.

We observe that the fear group has a response to the spider stimuli over twice that of the no-fear group (the conventional fear-based response) with a difference of .29 between them. As we did not detect this difference as significant in our statistical analysis, it makes sense to ask what sample size we would require for this difference to be statistically significant. Using the 'pwr' package of R, we can calculate what power we have if the forward-in-time spider response effect size is that observed (.27), which gives us a figure of .44. We therefore would need some 82 participants in each fear group to be able to statistically differentiate the spider-fear groups given our observed responses to reach power of .80. This lack of power suggests why the statistical analysis was not significant (although it was suggestively significant).

If this study lacks the power to detect the conventional difference between the two fear groups it is clear that we also lack the power to detect a small difference before stimulus onset. The 'pwr' package of R allows us to calculate the power of the study, if the difference that we see between the two fear groups before spider onset is the consequence of presentiment (.1 difference, a third of the difference seen after spider exposure), which gives us power of .49. To reach a conventional level of power (.80) would have required some 71 participants in each fear group. As with Presentiment I increasing the number of trials may also have helped improve the power of the study, although a multi-trial procedure introduces other problems.

2.2.4 Discussion

The results of Presentiment II are more complete than those obtained for the first study. Although there is some reasonable evidence suggesting that the spider stimulus is eliciting a post-presentation physiological response (the graphical plot and suggestively significant t-test result), neither the conventional nor the paranormal effect have produced statistically significant results ($< .05$). As we concluded that this study is under-powered in regards to the size of the observed difference between the fear groups for the conventional response, it must be even more true for our ability to discern a presentiment difference (as we would expect that response to be related but smaller). Thus as with Presentiment I, we must conclude that the lack of significant presentiment effect is not necessarily due to the absence of such an influence,

but is possibly due to the lack of power in the study to discern such an influence. As was mentioned previously, of course no parapsychologist expects to find significant paranormal results even when power considerations have been met to a level required in conventional psychological research.

Since we have some evidence of a conventional spider-fear based influence, it is an obvious question to ask why this was not revealed statistically. Although increasing the number of participants or increasing the number of trials (and thus increasing the power) may well have helped with this aspect, it appears that the FSQ measure used to discriminate spider-fear from no-fear groups has not provided us with two spider-fear groups who are sufficiently different from one another in regards to their fear of spiders. In short, the fear group is not fearful enough. It is clear that some attempt should have been made to seek out participants with a high fear of spiders and to eliminate or avoid recruiting low-mid fear participants, thus teasing apart the spider-fear aspect. However there is a difficulty in recruiting participants based on their fear of spiders, in that it is possible that participants may have construed that the experiment was related in some way to spiders, which was undesirable in the current study. If participants are informed regarding the nature of the study (or have some idea that spider-based stimuli will be presented) then anticipatory strategies come into play and confound the issue. This is a difficult problem to avoid. In a paper encountered well into the final stages of the thesis write-up, Dalkvist and Westerlund (2006) argue that the results of many experiments, including the presentiment studies, are influenced by just such expectancy effects and that although the theoretical influence of expectancy and arousal *should* be zero, it is in fact larger than the hypothetically accepted influence. (It should be noted that the one-trial methodology was used specifically to avoid eliciting any anticipatory response that might be misconstrued as paranormal.)

A further related methodological issue must be whether any of the participants turned their heads, or closed their eyes, in response to the image of the spider (or perhaps even, throughout the whole experiment). Although none of the participants should have been aware that they were to be presented with a spider image, and participants should also have not been explicitly aware of exactly when the images were going to be presented, it is true that there exists the possibility that some participants either kept their eyes closed during the experiment (expecting the worse) or viewed the stimuli through half-shut (or hand-covered) eyes. Even if participants did not do this, there is also the possibility that the most spider-fearful participants quickly averted their eyes on presentation of the spider image. Although we might not expect such an avoidance to completely remove the appropriate electrodermal response to the spider image, it could be argued that such behavioural responses may reduce the intensity of the stimuli (and with it the electrodermal responses of the participant). At the very least participants should have been instructed to keep their eyes on the screen at all times and not to avert their gaze and a video taken of the session to remove any participants who did not comply with these instructions. Ultimately it will be for future researchers to investigate whether participant stimulus avoidance does influence responses within this type of experimental setting.

Some overall conclusions from the two experiments may be drawn. The presentiment effect was not replicated in the two experiments conducted. Methodological procedures involving physiological measurements of the type described are more complicated and fraught with technical issues which make large-scale testing for paranormal phenomena difficult. Using the one-trial methodology was an attempt to mitigate many of the potential confounding

factors that might result in a type I error, and yet we conclude that without a way of *increasing* the difference between the two fear groups, this is still a methodology that is far from ideal (we have noted above that ensuring that the fear group is composed of high-fear individuals would be the first thing to change if the one-trial methodology is retained). It also seems clear that despite the rather grand claims made for it by researchers such as Radin and Bierman, that in reality psychophysiology is not a clean measure of response and appears to be a poor candidate for demonstrating a parapsychological effect. This is because of the inherent difficulty in collecting the data, in standardising the testing environments and ensuring that the data are as clean as possible (all things that are not outside of the abilities of competent parapsychological researchers) but when these difficulties are combined with small influences and low-powered studies, we see great potential to spend an awful lot of effort with either no expectation of results that have the power to convince researchers outside of parapsychology, or false-positives which again, do little to impress non-parapsychologists.

In the time since data collection (over a decade) the presentiment effect has continued to be investigated by a variety of authors with mixed results. For example, Broughton (2004) (who was discussed in the introductory chapter in regards to his ideas about what psi might be for) has investigated the presentiment effect from the perspective that it might be an evolved adaptive response to threat in the future (although that particular study was also non-significant) and there have been new adaptations, for example using a startle-reflex type methodology (Spottiswoode & May, 2003). An important paper by Mossbridge, Tressoldi, and Utts (2012) describes a meta-analysis of some 26 presentiment-type experiments conducted between 1978 and 2010 (they do not include the two current studies, though this is presumably because they have only been published in a single conference proceedings and not in the manner they have been presented in this chapter). The authors argue that there is a significant overall effect ($z = 6.9, p < 2.7 \times 10^{-12}$), with a small effect size (.21). The experiments included use a range of biological markers (including EDA). They report a conservative estimate of 87 unpublished reports that would be required to reduce the statistical significance of their analysis to chance levels. Ultimately, the authors conclude that there is some good evidence for a seeming presentiment effect in the results of a number of similar experiments, and that whilst there are a range of explanations for such an influence being found (for example sensory cues or multiple statistical analyses) it is possible that the results themselves are indicative of something more controversial (psi perhaps) and thus they urge further experimentation into the influence by competent independent laboratories. We concur with this suggestion.

Given the difficulties described in this chapter, and without the elusive evidence being sought, no further presentiment research will be conducted and the conclusions remain tentative. It will be left for other researchers to categorically prove or disprove the presentiment effect veridical or artefactual. Although it might be taken that the two studies are low-powered and thus “scientifically useless”, it seems fair to recognise that we have learnt a variety of things. Psychophysiological measures appear to be a promising but technically difficult response measure to deal with, and given issues of power and the size of the encountered effect it is not surprising that we will steer away from this particular paradigm. Regarding the use of spider-based stimuli the overall conclusions are mixed. Although no paranormal effect has been found, spider images did provide a less ethically problematic target stimulus and also allowed participants to be compared across their reported fear groups. This introduced an interesting comparison within the research and while it has not been successful in the first two experiments, it will be used again in the experiments to come.

Retrocausal Influences in Stroop-Based Tasks

3.1 Time-Reversed Interference I

3.1.1 Introduction

Although no evidence for time-reversed physiological arousal had been found in the first two studies (described in the previous chapter), it was hoped that by concentrating upon a similar area with equally strong claims made about it, that some evidence for a paranormal effect would be discovered. It is unusual for researchers working outside of parapsychology to seemingly find and publish evidence supporting a paranormal hypothesis, and given the negative opinion that normal psychology holds about parapsychology, it is certainly noteworthy that Klintman (1983, 1984) claimed his precognition studies were published as a consequence of finding an unusual effect in his data, rather than an established parapsychological researcher discovering an anticipated paranormal effect.

As a mainstream psychologist, Klintman (1983, 1984) was interested in cognitive interference tasks based upon a paradigm developed by Stroop (1935) as part of his PhD dissertation and since that time, the Stroop colour-word test has been widely investigated within psychology (Jensen & Rohwer Jr, 1966). The test itself usually consists of three distinct tasks:

1. Colour naming: Participants are presented with coloured rectangles or blocks (usually red, green, yellow or blue) in a random order, and asked to name the colours as quickly as possible.
2. Word reading: Participants are presented with colour-words (again “red”, “green”, “yellow”, or “blue”) usually printed in black ink, and are asked to read aloud the colour-words as quickly as possible.
3. Colour-Word naming: Participants are presented with the same list of words, but the word is printed in a different colour ink to that of the word itself (e.g. the word “blue” might be written with green ink). In this condition, participants are instructed to identify the colour of the ink (not the colour-word that is written) as quickly as possible.

Generally speaking, research has shown that participants find 2 the easiest, followed by 1, and 3 most difficult of all. In fact it is well established that the Colour-Word naming task can produce quite dramatic levels of frustration, with stammering and nervous laughter being

common. Stroop interpreted his own findings as evidence for an automaticity for reading stimuli (that despite being instructed to ignore the word, it was still begin subjected to processing) and that the results perhaps reflected an underlying structure by which mental information is stored and processed. MacLeod (1991) reports that over 400 studies have shown that the Stroop effect is reliable and replicable (qualities often lacking in parapsychological research) and it is now well established that tasks involving interference can influence behaviour.

Publishing many years after Stroop, Klintman (1983, 1984) was interested in a related interference phenomenon, where participants first had to identify the colour of a rectangle (coloured red, green, yellow or blue) and then read aloud a written colour-word presented in white (again “red”, “green”, “yellow” or “blue”). Based on the large body of previous research, Klintman had predicted that participants would name the written colour-word faster if it was congruent with the previous coloured rectangle: for example, the word “green” would be identified faster if it directly followed a green-coloured rectangle, than if the word and coloured rectangle were incongruent, for example, where a green rectangle was followed by the word “blue”. In an unusual move for this type of experiment, Klintman decided to not only concentrate on the reaction time to the second stimulus (referred to as RT2) but to use the reaction time to the first stimulus (the coloured rectangle) to help him establish a more sensitive baseline for each participant.

Unexpectedly, Klintman (1983, 1984) reported that in addition to the expected conventional interference effect of the reaction to the second stimulus (S2) being influenced by the colour-congruence of the first stimulus (S1), that there seemed to be an influence of S2 on the reaction to S1, something which Klintman argued, goes against current understanding of causality (because that influence appeared to be a colour-congruence effect, and trial congruence is not established until S2 presentation). Importantly, although the effect was similar to that found in conventional interference experiments, it was reversed in the temporal dimension, that is, the future (congruence established at S2) influenced the past (the reaction time in response to S1). Klintman claims that he was at a loss to identify what interference could explain this effect, although he admits some familiarity with the parapsychological literature and decided what he had found could well be precognition. Using colour projectors and a voice-key, Klintman embarked on an investigation of this effect, which he termed, *time-reversed interference* (hereafter referred to as TRI). A graphical illustration of the general TRI hypothesis can be found in Figure 3.1.

It is important to understand that Klintman (1983, 1984) established the TRI effect through a calculation of a correlation between two variables, D1 and D2. D1 was calculated from the reaction time scores to S1 (congruent reaction time–incongruent reaction time, marked on Figure 3.1 as a–b) and similarly D2 was calculated from reaction time scores to S2 (again congruent reaction time–incongruent reaction time, marked on Figure 3.1 as c–d). Because D1 scores should be approximately zero (because colour-congruence is not established until S2 onset), Klintman argued that any significant correlation between D1 and D2 was impossible (given current understanding of causality), and was thus evidence for precognition. It should be noted that this is somewhat a simplification of Klintman’s hypothesis because he does identify two groups of participants whom he refers to as *facilitators* and *inhibitors*, who essentially respond in the opposite manner to one another. So-called facilitators react in a conventional manner to S2, that is, their reaction times to S2 are faster on congruent trials than incongruent trials and thus Klintman argues that their TRI responses will mirror this: facilitators will react faster to S1 in congruent trials than incongruent trials. Inhibitors react in an idiosyncratic

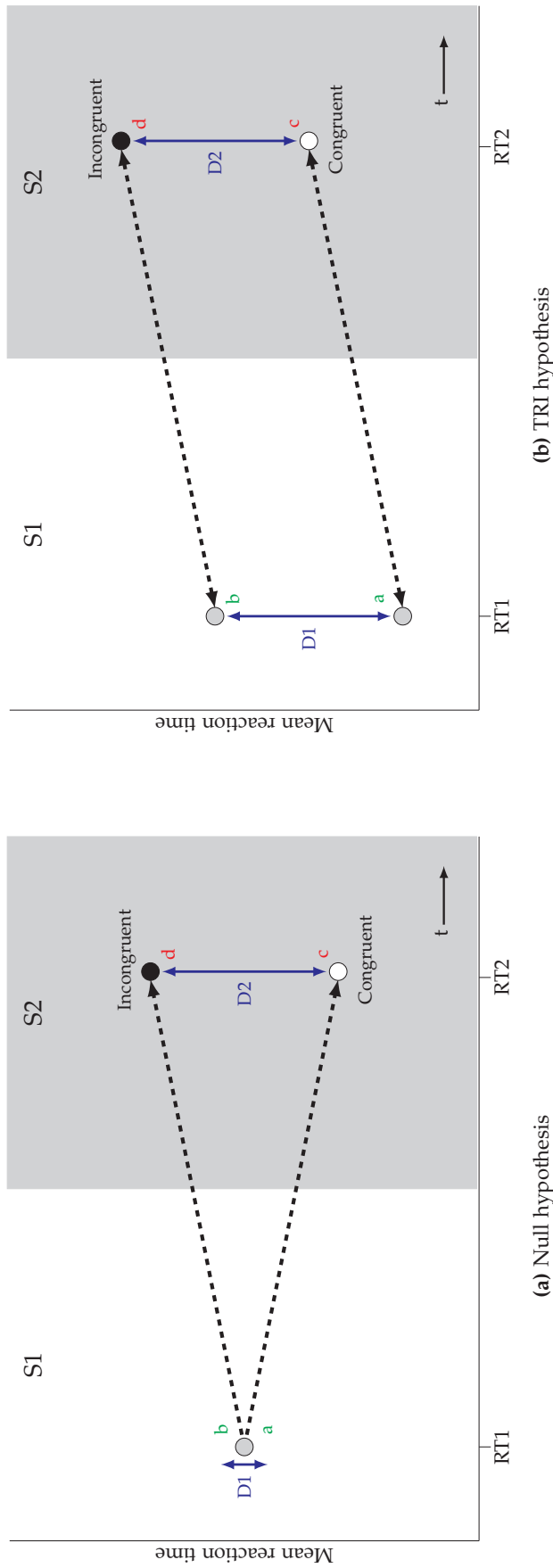


Figure 3.1: Graphical representation of null and TRI hypotheses. Considering (a) first we see that there is no differentiation of reaction times to S1 (the white section). D1 is calculated as the difference between RT1 scores on congruent and incongruent trials (a–b), where congruence is not established until S2 presentation. After S2 presentation (the grey section) there is clear Stroop-based differentiation of reaction times, depending on the congruence of the trial: incongruent trials will result in slower reaction times to S2 and congruent trials faster reaction times to S2 (RT2). D2 is calculated as the difference between RT2 scores on congruent and incongruent trials (c–d). This is a normal interference phenomenon. (b) shows the TRI hypothesis, which predicts differentiation of RT1 scores based on the congruence of the trial, established at S2 onset. (Although D1 and D2 are shown here with equal magnitude, this is for illustrative purposes and we would expect D1 to be relatively smaller than the corresponding D2). It should be noted that although one might expect there to be a significant difference between a and b given a TRI influence, that all other researchers into the TRI effect have established it through a significant correlation between a–b and c–d (marked here as D1 and D2). This graphic has been adapted from Radin and May (2000).

manner: they are thus slower on congruent S2 trials and faster on incongruent S2 trials, with this response mirrored at S1, thus they react slower on congruent trials than incongruent trials. Klintman makes this distinction based on what he calls 'earlier observations' and suggests that this individual difference is due to 'the individual's tendency to inhibit or, alternatively facilitate perceptual/cognitive activity in the afterphase of an initial percept' (Klintman, 1983, p. 133). Klintman does not provide much more detail about this aspect of the methodology but compares this splitting of participants into facilitators and inhibitors as similar to the familiar parapsychological comparisons of extrovert versus introvert or sheep versus goat (e.g. Schmeidler & McConnell, 1957; Palmer, 1977).

Although D1 and D2 were calculated in the manner described above, rather than use all collected data for the calculation, Klintman (1983, 1984) used a subset of the trial data collected for each of his 28 participants with specific inclusion criteria. Klintman argued that to use all collected data would risk inflating correlations due to changes that influenced both RT1 and RT2. In short, Klintman used just two pairs of reaction times out of the 36 trials that participants individually contributed. The two pairs were formed thus:

$D1 = RT[1](C) - RT[1](I)$: The difference in RT to S1 between trial 1 and the first trial differing from trial 1 in congruence (congruent trial minus incongruent trial).

$D2 = RT[2](C) - RT[2](I)$: The difference in RT to S2 between the first trial, not included in D1, and the first subsequent trial (not included in D1), with the same colour and differing in congruence from the former; (congruent trial minus incongruent trial). (Klintman, 1983, p. 23)

To be clear, the paranormal hypothesis of Klintman (1983, 1984) predicts that if $D2 < 0$ then $D1 < 0$, and if $D2 > 0$ then $D1 > 0$. (Participants with a $D2 < 0$ are categorised as facilitators and those with a $D2 > 0$ are inhibitors. Because this distinction is made on the result of two data points (congruent RT2 minus incongruent RT2) it is also clear that Klintman cannot be certain this is a trend that holds for all the participant's trials.) In his first experiment, Klintman tested 28 undergraduates over 36 trials and found a significant paranormal relationship between D1 and D2 and thus Klintman felt he had confirmed his informal observations made during earlier investigations.

Although settling on the paranormal explanation as the correct interpretation of what was going on in the experiment, Klintman (1983, 1984) does raise a number of methodological factors worth repeating. Firstly he notes that the source of randomness in the experiment, which involved implementing an alternator switch between two projectors, each filled with a random order of stimuli, may have in fact contributed far more alternations than might be expected in a random series. This is dismissed by Klintman as not being of sufficient strength to have influenced the data in the way seen. Secondly Klintman raised the issue of individual differences in participant performance. Specifically he claimed that although there is a degree to which the TRI effect may be seen as a universal interference effect merely reversed in the temporal dimension, that there may be a degree of variation in the levels that this effect is shown by different participants. Accordingly, Klintman hypothesised that not everyone will respond in the same way in the experiment.

Although the first TRI study reported in Klintman (1983, 1984) used a Stroop-based task (as described above), for his only other TRI studies (II–V, reported in the same two papers as Study I) he significantly changed the methodology and removed the colour-congruence aspect. It is interesting that the methodology implemented in his studies II–V has never been

replicated independently, and only the colour-congruence TRI effect has been investigated by other researchers (and Klintman has not published further on the topic). This may be because not only is the methodology quite complicated (as compared to that used for the colour-congruence task) but Klintman's justification for changing his methodology is also quite complex. An attempt will be made to describe both aspects of the research (the change in methodology and the justification for the change) although both are only of minor significance to the research described in this chapter (as it concentrates almost exclusively on Stroop-based colour-congruence).

To most easily understand why Klintman (1983, 1984) changed his procedure quite considerably, it should be regarded that a trial in his follow-up research essentially consists of two TRI trials, as they are in the first study. What do we mean by that? The follow-up studies (II–V) generally followed this procedure:

1. The participant is asked to produce a 1–2 second interval on a digital counter (the counter has 1 microsecond accuracy and the screen is covered). When done, the participant is asked to remove the cover over the last digit, read it aloud, and then decide if that digit is odd or even.
2. The participant then turns over a card from a pile. The cards have numbers printed on them (0–9) and are in a randomised order. They read the number silently, and decide whether the card was the same or different in regards to the odd/even property to the number obtained on the digital counter (e.g. if the counter produced a 9 and the card showed a printed 5, they are both the 'same' in regards their odd/even property). The participant was to respond aloud with 'same' or 'different' depending on the outcome (and mark this on a record sheet).
3. After 3 seconds the participant is again asked to produce another 1–2 second interval on the digital counter, read it aloud and again decide if it is odd or even.
4. The participant then turns up another card (from a different pile, containing random cards also printed with numbers on) and again has to silently decide if the card and the digital counter digit are the 'same' or 'different' in regards to the odd/even property (and mark this on a sheet).

This describes one trial. Klintman (1983, 1984) argued that a spreading effect is important in these digital counter studies. The 'spreading effect' is an idea that Klintman claimed was based on early observations, where he believed that TRI can spread to other, related, events in the testing system. In the case of this procedure he argues that the outcome of 4 will influence the interval produced at 1. He makes the argument that if TRI can spread through the system, that it must be harder to achieve trials where the two odd/even decisions are *both* equal (where chance dictates that the outcome for the first counter-card comparison will be 50/50 and for the second will also be 50/50). If a TRI influence is in the data, Klintman predicts that occasions when the two odd/even decisions are equal will be half as likely (from .5 to .25) because those occasions when an 'incongruent' outcome is obtained on the second comparison will spread backwards through the system and disrupt any equality on the first comparison.

Klintman (1983, 1984) justifies the significant shift in methodology by arguing that the button pressing procedure is capable of detecting smaller effect sizes, by amplifying a small effect through the use of temporal feedback. That is, the accuracy of the digital counter means that the time scales involved are very small and well beyond human capabilities to influence,

but that the response patterns *should* be demonstrably large as he expects .25 'same' trials versus .75 'different' trials (versus the 50/50 predicted by chance). With this new and rather complex methodology, Klintman argued he had stumbled upon a previously unknown precognitive interference effect and that it was both replicable and robust. As a whole, the research is both entirely optimistic in its interpretation of the findings and to future success, given the historical precedent within parapsychology. That the effect is some kind of experimental artefact is dismissed in the face of the presented evidence.

Despite the variety of interesting evidence and claims found in Klintman (1983, 1984), perhaps the most compelling question is why, if the TRI effect is that rare thing, a consistent and reliable paranormal effect, did Klintman *not* publish any further on the topic? His two papers were seemingly his first and last contributions to the field and yet beg a large question. The most obvious answer is that all future experiments conducted by Klintman ultimately failed to show any TRI effect, or that the influence remained but a normal explanation became apparent to explain it away. Perhaps the source of any artefact was eliminated and the effect along with it. If such a circumstance did indeed occur, this would highlight the great need for researchers to not only publish their non-significant research, but to publish notes to fellow researchers if any great source of bias is detected.¹

For a number of years this significant paranormal effect was seemingly ignored in the literature, until Camfferman (1987) published the first independent replication. Camfferman had been impressed by the claims made in Klintman (1983, 1984) and attempted to improve the methodology, whilst staying as near to the original research (Klintman's first experiment) as possible. Camfferman tested 40 participants, who each contributed two runs of 35 trials each. One run was a direct replication of Klintman's procedure, with a coloured rectangle (colour-block), followed by a colour-word, whereas the alternative run involved a reversal of that procedure, where participants were asked to name a colour-word, followed by a colour-block, thus providing an interesting contrast. Rather than using a projector and alternator as Klintman had done, Camfferman used a computer to present stimuli (and thus used a computer-based RNG to randomise the order), but still measured oral responses through a voice-key. Camfferman eliminated those trials where there was an erroneous response, mostly due to extraneous sources of noise. Camfferman felt obliged to analyse his data in the same way as Klintman had previously done (despite the large loss of data) but also decided to apply a similar statistical method that used much more of each participant's trial data:

$$D1 = \text{mean RT1 congruent} - \text{mean RT1 incongruent}$$

$$D2 = \text{mean RT2 congruent} - \text{mean RT2 incongruent} \text{ (Camfferman, 1987, p. 18)}$$

Camfferman (1987) found no evidence for a time-reversed interference effect in his data using the statistical technique employed by Klintman (1983, 1984). Using his means-based analysis, he did find a TRI-type influence for the colour-name run, but *not* for the name-colour run. To be clear, this analysis also categorises participants into facilitators and inhibitors, but is based on the means (i.e. mean RT2 congruent minus mean RT2 incongruent) and not just the two data points as used by Klintman. Interestingly, Camfferman found that in the

¹I have attempted to communicate with Holger Klintman on a number of occasions to ascertain the exact reason for his not continuing this research. With no answer to date, the truth of the situation can only be speculated at. Marks and Kammann (1980, p. 35) make this statement in regards to a similar situation: "If scientists are not prepared to allow free and open inspection of their data *after publication of the putative findings*, then the findings themselves are brought under question. One of the goals of science is to let evidence be known. Science is not a secret society; it is an open forum." No criticism of Klintman is implied, other than that it is regrettable that the reason he did not pursue his research further is not a matter of public record.

name-colour run *all* participants had a $D2 < 0$ (facilitators) and thus was evidence for a strong Stroop influence in the run, and yet there was no evidence of a corresponding TRI. Thus, although it is not readily apparent why some people are quicker on the incongruent trials than congruent trials (inhibitors), it is seemingly a consequence of the procedure (colour-block followed by naming a colour-word). Ultimately Camfferman concluded that a fast RT1 will be followed by a fast RT2 and a slow RT1 will be followed by a slow RT2 and that this was a consequence of participants' general level of alertness at any given moment, rather than a temporally reversed interference effect, and he provides some substantial subsequent analysis within the paper, seemingly confirming his conclusions.² Camfferman was happy with this explanation of the results and did not continue any further investigation into the alleged paranormal phenomenon.

Describing the effect as established is certainly unfounded in the face of the mixed experimental findings, however this is very much how it is portrayed in the book, *The Conscious Universe* (Radin, 1997a). Despite the lack of any significant developments in the field for nearly a decade, Radin was very much impressed with Klintman (1983, 1984), reporting that RT1 was faster on congruent trials than on incongruent trials, even though the congruence of trials depended upon the nature of the second stimulus (which was not presented until RT1 had been recorded). That this is a simplification of Klintman's hypothesis should not pass without comment. Radin's evaluation of the area for his readers is entirely simplistic and very much portrays the area in a positive light. Radin neglects to go into great detail about the range of experimental and statistical techniques employed by Klintman, for example the use of negative temporal feedback loops and does not even mention Camfferman (1987) and his alertness hypothesis as a normal explanation for the results. That there is a potential difference of opinion about the results is not suggested and leaves the reader with an unjustifiably positive opinion of the area.

No further research was published on the TRI effect until Radin and May (2000). Radin and May concluded that the previous research was certainly suggestive and that the experiment was worth replicating, not only because of Klintman (1983, 1984) but also due to the success of other allegedly retrocausal effects (such as the presentiment effect as discussed in Chapter 2). Obviously building on Radin's interest in the area, they reported four experiments conducted to test the paranormal TRI hypothesis. Radin and May use a different method to the previous authors in the area for determining a TRI response in their data. Essentially Radin and May standardised participant responses and looked for a correlation between RT1 and RT2 scores across the four colours separately. Radin and May also did not split participants into facilitators and inhibitors (this is presumably because they did not have participants who responded in such a manner). In their first three experiments, for the first stimulus (S1), the participant was shown a red, blue, green or yellow coloured rectangle (a colour-block) and was asked to type the letter 'r', 'b', 'g' or 'y' to identify the corresponding colour. Having responded as quickly as possible, stimulus 2 was presented, a colour-word ('red', 'blue', 'green' or 'yellow') and participants were required to type 'y' if it matched S1 and 'n' if it did not. Obviously this means that for a yellow match trial the participant would press the "y" key twice (and this is the only combination with that sequence). Study four was a colour-block, colour-word experiment and more similar to the previous research.

Radin and May (2000) reported a significant TRI effect in each of their four studies.

²Seemingly, because although a precognitive influence should be obvious within this type of data, establishing subtle dependency issues within this reaction time setting is more difficult.

Regarding the only normal explanation for such a result, they summarily dismissed the alertness hypothesis (Camfferman, 1987):

... while Camfferman's conclusion that RT1 and RT2 are related is undoubtedly correct, his assumption that Klintman's observation could be completely attributed to the variations in alertness may have been premature. (Radin & May, 2000, p. 2)

A number of methodological considerations should be mentioned regarding Radin and May (2000). Due to the different locations of the relevant keys that the participants had to press, Radin and May investigated the key pressed as a factor in their research. There seems little theoretical reason for a TRI effect to differ according to the key pressed, that is, why there would be a TRI effect found on the 'b' key as opposed to the 'r' key. As the second replication of the Klintman research in 13 years, the complications introduced by the inclusion of the key press as a variable of interest seems difficult to justify. Though Radin and May argue that there is no reason to assume the TRI effect will be equal in regard to each colour. There must be the concern, however, that by investigating the possible TRI influence on each colour, Radin and May are increasing their chances of detecting a spuriously significant difference (type I error).

In Radin and May (2000), for the first three studies participants were asked to identify the colour of the colour-block (S1) and were then asked to decide whether the colour-word presented at S2 matched or mismatched. This contrasts with previous methodology where participants only had to identify the colour being presented to them (either block or word) and were not asked to decide explicitly whether the colours matched or mismatched. It is true that in Klintman's studies II-V he employed a match/mismatch decision but these were not colour-based tasks.

It seems likely though that Radin and May (2000) used this match/mismatch question in the hope of strengthening the TRI effect, based on the same logic that Klintman (1983, 1984) himself used when justifying the change in procedure from Study I to Studies II-V (that is, that the abstract reasoning would itself lead to a temporal spreading effect). Also worth noting is the number of participants involved in the research: Study 1 had only one participant who contributed 500 trials; Study 2 had ten participants contribute 1852 trials; details regarding Study 3 do not report the number of participants as the experiment was run remotely over a local network, where some 2223 trials were used in the analysis; and finally, Study 4 had two participants contribute 720 trials. Although a low number of participants is not an ipso facto reason to dismiss the significant deviations from chance, it is a concern that rather than establish a convincing result with a credible regime of testing, Radin and May have seemingly completed four small-scale studies, the results of which were mixed.

Although the evidence for a TRI effect is interesting, the overall picture of research in the area is also mixed. Klintman (1983, 1984) conducted only one colour-based study, but claimed that a general TRI effect would be found in any interference task. The robust nature of the conventional Stroop effect supports Radin and May (2000) who claimed to have found another colour-congruence effect, although the alertness hypothesis (Camfferman, 1987) could certainly be an important normal explanation. The main aim of TRI I was then an attempt to demonstrate a time-reversed effect in a Stroop-based task. Radin and May's seemingly successful replication of Klintman's original effect (albeit in a more simplified version) set some precedent for adapting the methodology whilst remaining as close as possible to conditions in the original research. Also the great reliability of the Stroop effect strengthened the reasoning that if this was a paranormal interference effect, it would not be special to the methodologies used previously, but could be shown in a variety of interference situations.

We can see that the TRI effect has been established through a variety of methods. If we summarise the general claim to *reaction times to S1 depend upon the congruence of the trial (established at S2 onset)*, it becomes clear that it might be expected that the reaction times to S1 will be significantly different for congruent trials than incongruent trials (and as Klintman, 1983, 1984 identifies some participants as facilitators and others as inhibitors, we might expect some participants to react in the opposite manner to others). Because congruence is not established until S2 onset, the null-hypothesis would expect no such differentiation of RT1 times based on a yet-to-be established category of colour-congruence. However, because it is true that a significant correlation between D1 and D2 does not require a significant difference between a–b and c–d (as shown in Figure 3.1), it will be important to establish whether such a correlation is found in our data. Thus, TRI I was an attempt to test for time-reversed interference in a Stroop-based task, remaining close to Klintman’s study I, and to analyse the data in the variety of ways used by other authors in the area, not in an attempt to fish for any result that might be construed as paranormal, but in an attempt to better understand the cause of any significant findings.

3.1.2 Method

3.1.2.1 Design

TRI I employed a mixed design where reaction times for naming colours were compared for congruent trials and incongruent trials and compared across facilitators and inhibitors.

3.1.2.2 Participants

Some 40 individuals volunteered, mainly undergraduate students at Goldsmiths College, University of London, 20 males and 20 females, from 16 to 76 years old, with a mean age of 30 years and a standard deviation of 14 years. Participants were separated into facilitators and inhibitors based upon whether their mean reaction times to S2 (RT2) were larger or smaller on congruent trials than on incongruent trials (those faster on congruent S2 trials were the facilitators, and those faster on incongruent S2 trials were the inhibitors). There were 24 facilitators and 16 inhibitors. As the planned analysis involves using the mean reaction times, it makes sense to not use criteria as set by Klintman (1983, 1984) for defining facilitators or inhibitors, but instead denote participants as such based on their mean RT (the method used by Camfferman, 1987).

3.1.2.3 Materials

A desktop PC was used to run the experimental program, which made use of standard pseudo-random algorithms seeded from the internal clock, and attached to it was a voice-key with microphone. The software was constructed by the author in a psychology experiment development package called *E-Prime*.³ The experimental program was thoroughly tested before experimentation to ensure random presentation and accuracy of recording. In addition sessions were audio-taped to aid in the identification of incorrect responses or alternative sources of noise.

³Further details about the E-Prime software package can be found at <http://www.pstnet.com/>

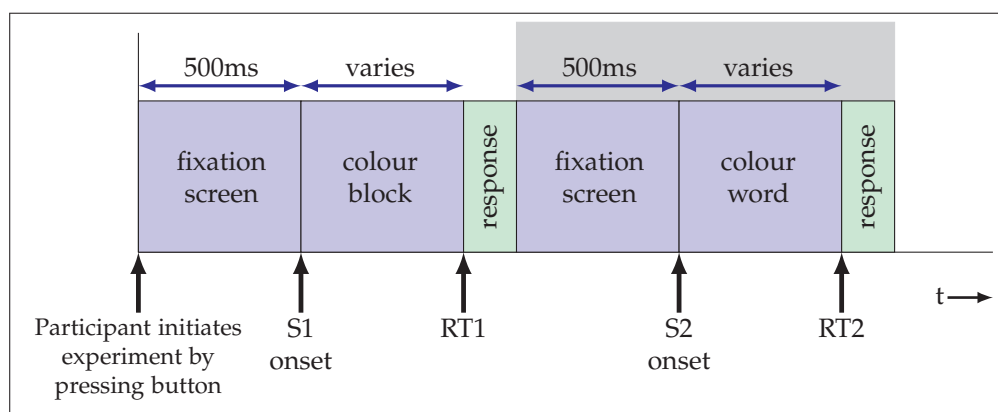


Figure 3.2: Graphical representation of TRI procedure.

3.1.2.4 Procedure

Participants were first directed to the testing computer and asked to read the instructions outlining the procedure they were expected to follow. They were informed that they would first have to verbally identify the colour of a colour-block and then read aloud a colour-word, written in white. When the participant had confirmed understanding of the instructions, five practice trials were initiated. A single trial consisted of a fixation point appearing in the centre of the display, marked with a white-coloured + on a black background and lasting .5 seconds, after which a coloured block was presented in a random colour (red, green, yellow or blue) on a black background. The display was cleared after the participant had named the colour, triggering the voice-key, and a second fixation point was presented, again lasting .5 seconds. This was followed by a random colour-word (“red”, “green”, “yellow” or “blue”) written in white and again presented on a black background. There was a 50 percent chance that the trial would be congruent (that is the colour-block would match the colour-word) or incongruent (where a coloured block was followed by a different colour-word). After the response to S2, the next trial was immediately started (beginning with another .5 seconds fixation screen). For a graphical representation of the procedure see Figure 3.2.

After five practice trials, the participant completed 36 experimental trials, which lasted approximately two minutes. After the last trial the participant was debriefed. A total of 1440 trials were completed (40 participants completing 36 trials each). Before analysis the data were screened to remove inaccurate trials (i.e. trials where the colour spoken was not that shown on the screen) and those trials where the response was either too fast or too slow (the range used was taken from Camfferman (1987), where trials were eliminated if RT1 exceeded 800 ms, or if RT2 was shorter than 200 ms or RT2 was longer than 1200 ms). A total of 1255 trials were used in the statistical analysis accounting for 87 percent of collected data.

3.1.3 Results

Rather than analyse grouped means⁴, the raw data were transformed into a standardised measure based on part of the procedure used in Radin and May (2000):

$$Z_{RT} = (x - \mu) / \sigma$$

⁴The grouped means were used in a previous draft version of this chapter. Although a standardised measure is reported in the final version, the conclusions drawn have not changed.

Thus for each participant, for each trial, the reaction time (RT) was subtracted from the mean (for the participant), divided over the standard deviation for that participant. With the scores standardised we can aggregate them according to congruence (congruent and incongruent trials) and response (RT1 and RT2). Table 3.1 presents these summary statistics. This table seems to show that both responder types are reacting in a Stroop-based manner to S2 (facilitators are faster on congruent trials and slower on incongruent trials, and inhibitors react in a contrary manner). Interestingly responses at S1 are smaller (approximately 3 times smaller than the responses seen at S2) but *are* in the correct direction (i.e. what is seen at RT1 is smaller, but in the same direction as what is seen at RT2). Is this TRI?

Table 3.1: Summary statistics (mean standardised reaction times and respective standard deviations) across responder-type, response and colour-congruence for TRI I

Responder-Type	Response	Colour-Congruence	\bar{x}	σ
Facilitators	RT1	Congruent	-0.20	0.61
		Incongruent	0.21	0.61
	RT2	Congruent	-0.49	0.40
		Incongruent	0.54	0.46
Inhibitors	RT1	Congruent	0.19	0.80
		Incongruent	-0.23	0.73
	RT2	Congruent	0.75	0.46
		Incongruent	-0.71	0.37

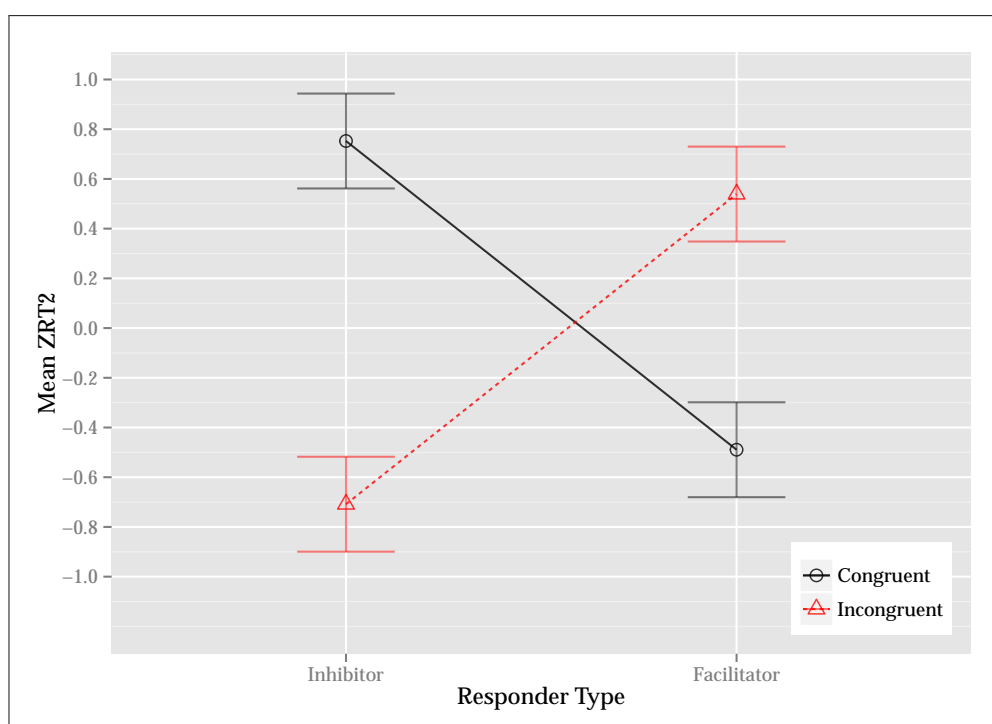


Figure 3.3: Interaction between responder type and colour congruence for stimulus 2

Statistical analysis of this data is simple. Two 2×2 mixed ANOVAs were run, comparing the two responder types (facilitator vs inhibitor) and colour-congruence (congruent vs incongruent); one for RT2 looking for a conventional Stroop based influence, and one for RT1 looking for a corresponding TRI influence. The ANOVA for RT1 was wholly non-significant. There was no

main effect of responder type, $F(1, 38) = 1.5, p = ns$, no main effect of congruence, $F(1, 38) = 0, p = 0$, and no interaction between responder and congruence, $F(1, 38) = 3.6, p = ns$. There is thus no statistical evidence for a TRI influence in this data.

The corresponding ANOVA results for RT2 were that there was no main effect of responder type, $F(1, 38) = 0.03, p = ns$, no main effect of colour congruence, $F(1, 38) = 2.5, p = ns$. There was, however, a significant interaction between responder-type and colour-congruence, $F(1, 38) = 83.7, p < .0001$. This interaction is presented in Figure 3.3. It is easy to see from this plot that the interaction is in essence a conventional Stroop-based result, reversed for those participants who are identified as inhibitors.

Finally then, to investigate this interaction, two paired-samples t-tests were conducted, comparing the two responder-types (facilitators and inhibitors) and their responses to stimulus 2, across colour-congruence. For the facilitators there was a significant difference between their RT2 scores, i.e. they were significantly faster on congruent trials than incongruent trials, $t(23) = -5.9, p < .0001$ (one-tailed). For the inhibitors the opposite finding was true, i.e. that they were significantly faster on incongruent trials than on congruent trials, $t(15) = 7.1, p < .0001$ (one-tailed). Both p-values have been adjusted via Holm's method (Holm, 1979). It should be noted that although participants were divided into respective responder-type *based* on their responses at RT2 (that is, facilitators are those who have a $D2 < 0$ and are thus faster on congruent than incongruent trials), that categorisation is based upon a binary (either/or) categorisation, and not a statistical consideration of the differences. Secondly we do not see the same results for RT1 scores as would be expected in the presence of a TRI influence (although the reaction times for S1 are in the correct direction, but if it is TRI it is much smaller than that reported by Klintman, 1983, 1984).⁵ It is clear then, that we *do* have a Stroop-based influence in the data, but we cannot conclude the same regarding TRI.

Given the seeming lack of TRI effect in the results, it is important to determine whether this study has sufficient statistical power to detect the effect as hypothesised. It should be stressed that although power was considered before experimentation began, it cannot be true that power alone would ensure replication of significant published effects. That being said, Klintman (1983, 1984) describes the size of the TRI effect in his Study 1 as at least as large as 65 ms difference between congruent and incongruent trials. We can ask to what extent this study would be able to detect a difference of that magnitude using the 'pwr' package of R (Champely, 2012) which provides a power test based on the t-test. An effect size of .88 (given the estimated difference in means of 65 ms divided by the observed standard deviation of 74), shows that we have power of 1 and would require a sample size of just 10 to reach a power of .8. It is clear then, that TRI I is sufficiently powerful to detect an effect of the magnitude described by Klintman. If, however, the effect is not as large as that described, then the power of the study rapidly diminishes as a function of sample size.

3.1.3.1 Klintman-style analyses

Although the planned contrasts have seemingly not provided evidence for a TRI influence in the data despite convincing evidence for a conventional Stroop-based effect, it must be stressed that the Klintman (1983, 1984) TRI effect does not require a significant difference between

⁵An earlier draft of this chapter did not split participants into facilitators and inhibitors, which explains the previous inability to identify the conventional Stroop-based effect. Because Klintman explicitly divides his participants according to this categorisation, and because we also find participants who react in an idiosyncratic manner, it is clear that expecting to find either a conventional Stroop effect or a hypothetical TRI effect is unreasonable if the data are pooled in the way they had been previously.

congruent and incongruent RT1 scores (marked on Figure 3.1 as a–b). Instead Klintman argued that he had discovered an inexplicable correlation between D1 scores and D2 scores (that is, a correlation between a–b and c–d). Thus it would be premature to conclude that there is *no* TRI influence in the data, and the subsequent analyses to be presented below argue that the picture is more mixed.

For the Klintman (1983, 1984) analysis, first two difference scores were calculated according to Klintman’s specified procedure (marked as D1 and D2 on Figure 3.1). To remind the reader, Klintman calculates his D1 and D2 scores thus:

$D1 = RT[1](C) - RT[1](I)$: The difference in RT to S1 between trial 1 and the first trial differing from trial 1 in congruence (congruent trial minus incongruent trial).

$D2 = RT[2](C) - RT[2](I)$: The difference in RT to S2 between the first trial, not included in D1, and the first subsequent trial (not included in D1), with the same colour and differing in congruence from the former; (congruent trial minus incongruent trial). (Klintman, 1983, p. 23)

As has already been discussed, Klintman (1983, 1984) proposed that participants would divide into two groups in terms of the difference between the way they respond to S2 i.e. facilitators, faster on congruent trials, will have $D2 < 0$ and inhibitors, who respond contrary to convention being faster on incongruent trials, having $D2 > 0$. It is thus simple to understand what Klintman expected to find in his data given the TRI hypothesis, that is if $D2 > 0$, then $D1 > 0$ and for those who act in the opposite manner, if $D2 < 0$, then $D1 < 0$. It is important to reiterate that the TRI hypothesis as conceived by Klintman is a correlation between D1 and D2. The null hypothesis predicts no such relationship between D1 and D2 scores because colour-congruence is not established until S2 presentation (D2) and thus it should not be possible for D1 scores to correlate positively with D2 scores.

The results of this Klintman-style analysis are presented in Table 3.2. This is a contingency table, where participants are assigned to cells based on the direction of their D1 and D2 scores. The sample size is reduced by 1 (to 39) because of the strict inclusion criteria set by Klintman (1983, 1984), and the exclusion criteria used to remove trials, thus one participant did not have the full set of requisite trial data to be included in this analysis. The highlighted cells are those where Klintman expected his results to load if there is a TRI effect, as they do in the corresponding table in Klintman’s paper. It is important to recognise that the participant categorisation into facilitator and inhibitor in this analysis is based on one pair of reaction times, although in general, when we are discussing facilitators and inhibitors elsewhere (not Klintman) we are discussing a general trend in regard to the means.

Table 3.2: Klintman analysis for TRI I

	D1 > 0	D1 < 0	
D2 > 0	7	7	14
D2 < 0	16	9	25
	23	16	39
Fisher’s exact $p = .30$ (one-tailed)			

Describing this table in words, first let us consider the top row, which because this relates to participants who have a $D2 > 0$ are what Klintman (1983, 1984) refers to as inhibitors (and thus have a slower RT2 on congruent trials than on incongruent trials). There are 14 inhibitors then, and there is an even split (50/50) between those who have a faster RT1 on congruent trials,

and those who have a faster RT1 on incongruent trials. Thus inhibitors provide no evidence for a TRI influence (as given Klintman's hypothesis, the inhibitors should also have a $D1 > 0$). Next we consider the bottom row, which relates to those who have a $D2 < 0$ (the facilitators), who react in a conventional manner (having a faster RT2 on congruent than incongruent trials). There are 25 facilitators. This time we do not see an equal split at D1, but instead there are almost double the number with a $D1 > 0$ (meaning the facilitators were more likely to have a faster RT1 on incongruent trials than congruent trials). This is the opposite of the pattern that Klintman predicted, where facilitators are expected to show a $D1 < 0$.

In relation to this table, Klintman (1983, 1984) reports the group medians for D1 and our corresponding values are for $D1 > 0$: 75 ms and for $D1 < 0$: -116 ms.⁶ TRI is not evident in Table 3.2 which is confirmed by the Fisher's exact test and overall Spearman rank correlation between D1 and D2 values for the 39 participants of $r = .08$, $p = ns$. Given Klintman's criteria there is no evidence for TRI in this data.

Moving on, next a Camfferman-style analysis was conducted. It is conceptually similar to the Klintman-style analysis but instead of the four individual data points used to calculate D1 and D2, Camfferman (1987) makes use of all collected data to calculate D1 and D2 thus:

$$D1 = \text{mean RT1 congruent} - \text{mean RT1 incongruent}$$

$$D2 = \text{mean RT2 congruent} - \text{mean RT2 incongruent} \text{ (Camfferman, 1987, p. 18)}$$

These data are presented in Table 3.3 and are comparable to Table 3.2. Again, if a TRI effect is present in the data, we would expect to find the cells loading not randomly, but in the diagonals (highlighted) and for this Camfferman-style analysis we find that the data *do* load in these cells. Again, describing this table in words, we first consider the top row. This again relates to those who could be described as inhibitors (of which there are 16). Rather than the 50/50 split we saw for the Klintman table, this time we see the pattern of scoring that Klintman (1983, 1984) predicted: those who had a $D2 > 0$ also have three-times the likelihood to have a $D1 > 0$. The bottom row again relates to the facilitators (of which there are 24) and again we see a loading of cells which shows that those who had a $D2 < 0$ were twice as likely to have had a corresponding $D1 < 0$.

Table 3.3: Camfferman analysis for TRI I

	D1 > 0	D1 < 0	
D2 > 0	12	4	16
D2 < 0	8	16	24
	20	20	40
Fisher's exact $p = .01$ (one-tailed)			

The Fisher's exact test is significant, as is the Spearman rank correlation between D1 and D2 values for the 40 participants, $r = .33$, $p = .04$. Both the Spearman's and the Fisher's test are in-line with Klintman's hypothesis (though it is important to note that this is not based on the subset of data specified by the Klintman, 1983, 1984 exclusion conditions). Camfferman (1987) reports a similar finding in his own paper. It appears as though there *is* now some evidence that D1 scores are seemingly influenced by the colour-congruence of the trial (established at S2).

⁶Following Klintman's paper we restrict ourselves to presenting only the D1 median results.

3.1.3.2 Camfferman's "Alertness" hypothesis

Although we have apparently replicated some of the findings of Camfferman (1987) and somewhat confirmed the Klintman (1983, 1984) hypothesis (albeit with the grouped data), Camfferman himself is not satisfied that Klintman's explanation for the correlation between D1 and D2 is correct and is sceptical as to whether Klintman's claim of a precognitive TRI influence is even necessary to explain significant correlations between D1 and D2. Camfferman thus engages in an extended analysis of his data which he argues supports a more plausible explanation. Following Camfferman we find that D2 is more often negative than positive (24 participants showed $D2 < 0$ against 16 with $D2 > 0$, highlighted in orange), $z = 1.78$, $p = .04$, which Camfferman argues suggests that a conventional cognitive interference effect is present, i.e. reaction times in general are quicker in congruent rather than incongruent trials, producing a negative D2. With the observed preponderance for participants to have a negative D2, if time-reversed interference is present, most participants will be expected to also show a negative D1. Looking at Table 3.3 we can see that exactly 50 percent (20/40) of D1 values are negative (highlighted in green) and no further analysis is needed as this observed result is exactly what is expected by chance. So if we have interpreted the D2 relationship as being due to cognitive interference, we do not see a corresponding influence on the D1 results. Camfferman found a similar pattern in his data and we concur that this does not support Klintman's assertion of a precognitive paranormal influence.

From one perspective then, we have seemingly discovered a significant TRI effect as described in Klintman (1983, 1984) but on the other hand we concur with Camfferman (1987) in that if most people have $D2 < 0$, then we do not find that relationship at D1. Camfferman does a more extended analysis of his results to understand how it might be misconstrued that D1 scores are related to trial congruence established at S2 onset. Specifically, as Klintman feels D1 and D2 are independent, any correlation between D2 and D1 involves a precognitive influence of congruence established at S2 travelling backwards to influence responses to S1 but Camfferman is interested in whether there is a normal explanation for this correlation.

In an attempt to establish a forward-in-time coherence between the D1 and D2 scores Camfferman (1987) continued analysis of his data, which will be replicated here. We recall that there is a 50/50 split at D1, where participants have an equal probability of having a $D1 < 0$ and a $D1 > 0$. If we look at Table 3.3 again, we can see that of the participants with $D1 < 0$, 16 out of 20 (printed in red ink) have a $D2 < 0$. However, for the participants with $D1 > 0$, only 8 out of 20 (printed in red ink) have a $D2 < 0$. There is a significant difference between these proportions, $z = 2.58$, $p = .005$. Thus, although participants are equally likely to be faster or slower on congruent trials in response to S1, that direction of the response at S1 is *unlikely* to change when they respond to S2. As Camfferman notes regarding his own findings, although this supports the TRI hypothesis (in that there is an apparent relation between D1 and D2) that the 50/50 split at D1 does *not* support TRI and is more supportive of Camfferman's alertness hypothesis, which will be elucidated with further analysis below.

Camfferman (1987) provides the only other explanation for a TRI effect in the literature which he in fact states is a more plausible one, that of "alertness". Simply put, it is the idea that, contrary to Klintman (1983, 1984), who argues that RT1 and RT2 are independent, they are in fact distinctly related measures. Specifically, a fast RT1 will be followed by a fast RT2 and vice versa, and this is a normal rather than paranormal explanation, relying as it does on the fact that a participant's physiological reactions remain relatively stable as the experiment

progresses (or perhaps more correctly, that because the two responses are recorded very closely in time, some of the variance at D2 is a function of the variance seen at D1).⁷ To further develop this line of argument Camfferman engages in a rather extensive analysis that will be replicated here in an attempt to establish whether Camfferman's argument is credible.

Camfferman (1987) is interested in establishing evidence for a forward-in-time relationship as opposed to Klintman's backwards-in-time relationship and to achieve this Camfferman reconceives of $D1 < 0$ and $D1 > 0$ as proposed in Klintman (1983, 1984). For Camfferman, it does not make sense to consider D1 scores in relation to trial colour-congruence (as it is not established until S2). Instead Camfferman argues that the minimum information that we have about individual participant's responses is contained in their D1 scores, that is a $D1 < 0$ means that the participant has a faster mean RT1 for congruent than incongruent scores (and vice versa for $D1 > 0$). With that reconception, Camfferman reproduces the data in a table which both retains the shape of Klintman's table (that is participants who have a $D1 < 0$ are now considered RT1 fast) but presents how all participants react only in *congruent trials*. Importantly, Camfferman then creates relative frequencies which represent the proportions of participants who respond fast or slow at each stimulus and shows clearly the coherence of RT1 and RT2, which Camfferman calls 'alertness'.

Table 3.4: Frequency of slow and fast reaction times for Camfferman analysis for TRI I

	Congruent Trials			Relative Frequencies		
	RT1 slow	RT1 fast	Total	RT1 slow	RT1 fast	Total
RT2 slow	12	4	16	.60	.20	.40
RT2 fast	8	16	24	.40	.80	.60
Total RT1	20	20	40	.50	.50	1

As reaction times to S1 should not be influenced by trial congruence (established at S2 onset), whether a participant has been designated as having an RT1 fast or an RT1 slow should be random. Like Camfferman (1987), we can also observe that each participant has equal chances to fall into one of two categories: Those who reacted fast and were assigned a congruent trial at S2 (50 percent of the participants) and those who reacted slow and had a congruent trial assigned to them (the other 50 percent), both marked as green. Because we are considering just congruent trials, we can see that some 60 percent of participants (highlighted with orange) show normal cognitive interference i.e. those who were aided by congruence to react fast at S2, whilst the other 40 percent were not, also highlighted as orange. The different responders here equate to our previous distinction of facilitators and inhibitors. Finally, we see that of those who already had a fast reaction time to S1, the majority, 80 percent also reacted fast to S2 (highlighted in blue). Because congruence is not established until S2 onset and because this is just the congruent trials, we know that this is *not* just a colour-congruence effect. In those already slow at S1 only 40 percent were fast in their reaction to S2, also highlighted in blue. As has already been noted above, this proportion is significant $p = .005$.⁸ Camfferman interprets this finding as a measure of a forward-in-time relationship he defines as "alertness" and this is supported by the data so far considered here.

As part of his argument that D1 and D2 are not independent measurements, Camfferman

⁷Klintman does of course recognise this, which is why he favours a calculation of D1 and D2 using the two pairs of reaction times. However, given the results reported here (no significant Klintman effect, but a significant corresponding Camfferman analysis) it appears worth investigating.

⁸Camfferman in private correspondence noted the marked similarity with his published results.

(1987) constructed one further data set to investigate whether he could identify a direct association between RT1 and RT2 within each trial, for each participant. All trials for a participant were pooled and for each trial, each RT1 and each RT2 was labelled as 'fast' or 'slow' as thus: If the RT1 is below the overall median RT1 for that participant then RT1 is labelled as 'fast', otherwise it is 'slow'. The same is used to calculate the 'speed' of RT2. Scores that fall on the median are omitted and the participant median is recalculated.⁹ This technique was applied to the current data and the results are presented in contingency Table 3.5. This table presents relative frequencies and shows clearly the coherence of RT1 and RT2: we find that fast RT1s go significantly more often together with fast RT2s (as compared to chance). A z-test on the difference of the two proportions is significant, $z = 4.24$, $p < .0001$, one-tailed. Camfferman argues that this supports his 'alertness' relationship and we so-far concur.

Table 3.5: Frequency of slow and fast reaction times (RT2 as a function of RT1) for all trials for TRI I

	RT2 slow	RT2 fast
RT1 slow	.59	.41
RT1 fast	.41	.59
N = 556		

Following Camfferman (1987), lastly we split Table 3.5 to see whether this relationship holds across congruence, shown in Table 3.6. Firstly, yet again we can see that there is approximately a 50/50 chance of having a fast or slow RT1 and being assigned to a congruent or incongruent trial. Like Camfferman we can test this using a z-difference score by comparing the 48 percent who are fast at RT1 congruent with the 52 percent who are fast at RT1 incongruent (both highlighted in green). For this analysis we compare the two proportions to chance, $z_{diff} = 1.33$, $p = ns$ (one-tailed) and obviously the same will be true for the RT1 slow proportions.

Table 3.6: Frequency of slow and fast reaction times (RT2 as a function of congruence and RT1) for all trials pooled for TRI I

Congruent Trials			
	RT2 slow	RT2 fast	Total
RT1 slow	.58	.42	.52
RT1 fast	.38	.62	.48
Total	.48	.52	
N = 582			
Incongruent Trials			
	RT2 slow	RT2 fast	Total
RT1 slow	.60	.40	.48
RT1 fast	.44	.56	.52
Total	.52	.48	
N = 530			

Secondly, by comparing 52 percent in the congruent trials which are fast at RT2 with 48 percent in the incongruent trials that are fast (both marked in orange), we calculate the z-difference score using a z-score for each sample compared with chance (to make it directly comparable with Camfferman, 1987), $z_{diff} = 1.33$, $p = ns$ (one-tailed). Thus RT2 does not appear

⁹It is possible to criticise Camfferman's technique in that this analysis is a multi-level comparison, and thus it would perhaps have been better to employ a more specialist statistical technique to test this relationship.

to depend on congruence (and hence cognitive interference is not shown). This contrasts with Camfferman who found a significant difference between these totals. Although there appears to be a lack of evidence of a colour-congruence influence in this data, it is presumably the consequence of pooling the data without respect for responder-type (inhibitors and facilitators). Camfferman has fewer inhibitors in his data (and thus the colour-congruence effect is more marked) and in TRI II there are very few inhibitors and a clear colour-congruence effect.

Thirdly, for congruent trials, a fast RT1 is significantly more often accompanied by a fast RT2, $z = 4.12$, $p < .0001$ and for incongruent trials a slow RT1 is more often accompanied by a slow RT2, $z = 3.27$, $p < .0001$ (highlighted in blue). Again there is a marked correspondence between Camfferman's results and our own. Camfferman (1987) notes that yet again Klintman (1983, 1984) would interpret this as a TRI effect but he concludes that it is much more plausible to interpret this relationship as a consequence of "alertness" which is still evident when data is considered on an individual participant level. Given the repetition of the Camfferman analysis here, we concur with Camfferman's conclusions.

One final note regarding this analysis. It may be argued that by pooling the data across different participants Camfferman (1987) is violating statistical assumptions. Camfferman could have used a bootstrapping technique to investigate whether fast RT1s and fast RT2s are significantly correlated on an individual level. However it is likely that given the age of this research that a multilevel approach or a bootstrapping approach would have been too technologically difficult. It is also true that the approach used by Camfferman remains close to the original technique employed by Klintman (1983, 1984), and thus compares well across the studies. We also believe that one of the strengths of replicating this analysis has been the close correspondence between Camfferman's analysis and our own.

3.1.3.3 Radin & May-style Analysis

Finally, an attempt to analyse the data using the technique employed by Radin and May (2000) was conducted.¹⁰ Radin and May do not exactly follow Klintman- or Camfferman-style analyses as outlined above but their method of analysis is certainly comparable. We are repeating this analysis to allow comparison of our results across all the research in the area but it is understood that this is purely a secondary analysis. Radin and May are interested in developing a construct very similar to D1 and D2 and like Klintman (1983, 1984) are interested in whether there is a correlation between the two, arguing, as Klintman did, that such a correlation should be impossible and is thus evidence of a precognitive, time-reversed influence.

Following the Radin and May (2000) method, first reaction times were normalised for each participant into standard normal deviate scores (a z-score) for each trial for each participant:

$$Z_{RT1} = (x - \mu) / \sigma$$

This is calculated as the RT1 score for a given trial minus the mean RT1 for all trials for the participant over both congruent and incongruent trials, divided by the standard deviation of the participant's RT1 trials (giving Z_{RT1}). The process is repeated on RT2 trials to give Z_{RT2} .

$Z_{RT1.INCONGRUENT}$ is calculated as the sum of all of the Z_{RT1} scores for the incongruent trials divided by the square root of the number of incongruent trials and $Z_{RT1.CONGRUENT}$ is calculated as the sum of all the Z_{RT1} scores for congruent trials divided by the square root of

¹⁰No other investigation other than that reported here has been conducted.

the number of congruent trials (and this is repeated for the RT2 trials). From these scores two differential measures are calculated in what is essentially a z-diff test:

$$\Delta Z_{RT1} = (Z_{RT1.INCONGRUENT} - Z_{RT1.CONGRUENT}) / \sqrt{2}$$

According to the null hypothesis there should be no correlation between ΔZ_{RT1} and ΔZ_{RT2} , for the reasons already stressed a number of times (that colour-congruence is not established until S2 onset). Rather than merely look for a correlation between the two measures, Radin and May (2000) decide to break down scores depending on the colour of the stimulus, correlating ΔZ_{RT1} for each colour with the same for ΔZ_{RT2} .¹¹ It should be noted that the colour is what Radin and May call the *initial* colour-patch, in this context the S1 colour-block. This applies to ΔZ_{RT2} as well as to ΔZ_{RT1} . They justify this by arguing that since different motor movements are required to respond to the different colours, this may effect the scores and any cognitive interference or TRI effects found (also that there is no reason to assume an identical TRI influence on each colour). It may be argued that this is a very low-powered technique and that comparing 2 sets of four means will result in high correlations from time to time but these rarely will be significant. However it is easy to see that for Radin and May to use a multi-level analysis such as an ANOVA, they would have to run many more trials, and thus it would require a great deal more effort. We also cannot run an ANOVA with colour as a factor, because we too lack the requisite number of trials (for each colour).

Although it was argued earlier that Radin and May (2000) have somewhat simplified the Klintman (1983, 1984) hypothesis (that some participants will score in a conventional way, the facilitators, and some will score in an idiosyncratic way, the inhibitors), it is likely that Radin and May do not test for inhibitors because they do not find participants who qualify as such.¹² Given this difference between the current study and Radin and May's research, it was apparent that some strategy was required to deal with the idiosyncratic responders seen in this research. Rather than split the data into two subsets based on responder-type and produce two different Radin-type analyses (where we would expect an opposite direction of scoring in the inhibitor group), we decided to calculate the inhibitor scores as the congruent–incongruent reaction times (the opposite of the calculation employed by Radin and May), thus inhibitor scores and facilitator scores would be in the same direction. This allows us to ignore responder-type and to directly compare our results with those of Radin and May.

Table 3.7 shows the number of trials for the respective colours used in this Radin-style analysis, the results of which are presented in Figure 3.4. We can see from this figure a possible explanation as to why the evidence for a Stroop-based effect was not as strong in the previous analyses as might have been anticipated. Only the ΔZ_{RT2} for the colour red is significant ($p < .05$), although blue is very nearly significant. This indicates that participants were significantly faster on the congruent trials than the incongruent trials (but seemingly only for the red and blue colours). We cannot conceive of a credible explanation as to why this might be so. As for the correlation between ΔZ_{RT1} and ΔZ_{RT2} it is in the opposite direction to that predicted by Radin and May (2000) and therefore does not confirm their findings.

¹¹Radin and May (2000) could be criticised for not reporting the correlation between ΔZ_{RT1} and ΔZ_{RT2} before breaking it down across colour, though they presumably do not, as they argue that TRI is not necessarily significant on each colour separately.

¹²In TRI II, which is a keyboard-based procedure much more similar to that employed by Radin and May, there is a dearth of inhibitors (2 out of 50).

Table 3.7: Number of trials across colour for TRI I Radin analysis

	Congruent	Incongruent
Red	154	143
Yellow	174	164
Green	160	131
Blue	166	163

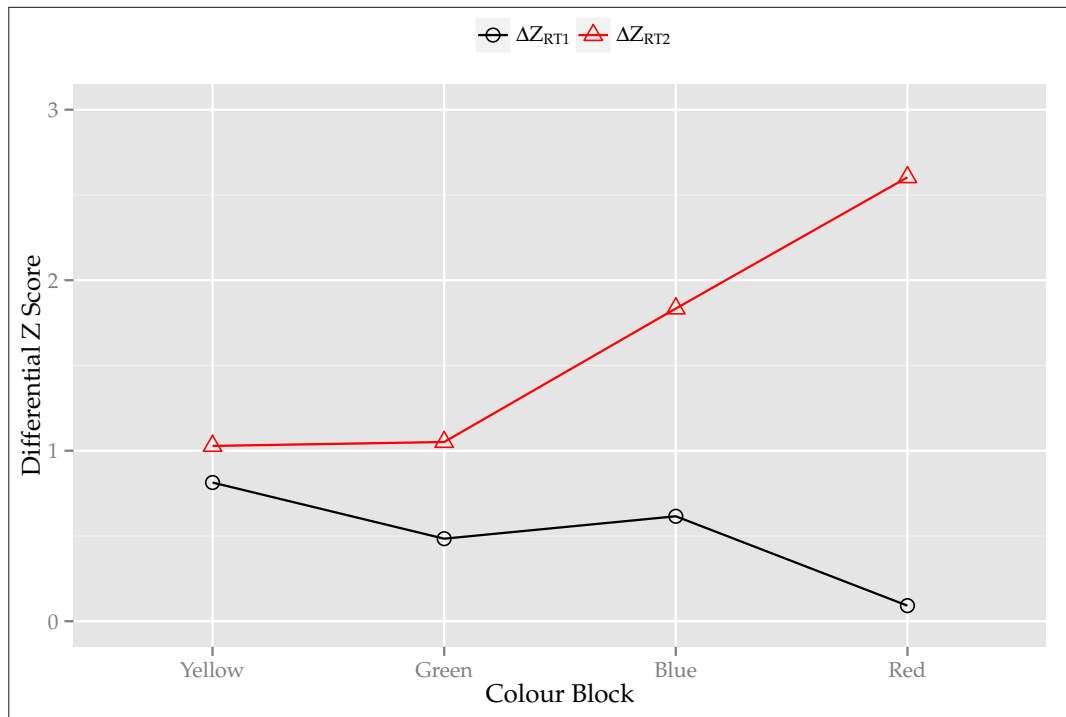


Figure 3.4: Values of ΔZ_{RT1} and ΔZ_{RT2} by colour for TRI I. The correlation between the two lines is $r = -.81$, $p = ns$ (one-tailed). The labelling of the x-axis by colour-block refers to the first stimulus (S1) only.

3.1.4 Discussion

The results of TRI I are as follows: An analysis of variance suggested a conventional Stroop-based effect was evident in the data, but no accompanying TRI-based effect was found. Like Klintman (1983, 1984) and Camfferman (1987) we have participants who react in an idiosyncratic manner to the conventional Stroop effect (i.e. they are slower on congruent trials and faster on incongruent trials) but taking this responder-type into consideration still does not provide evidence for TRI.

Because TRI has been established by the other authors in the area as a correlation between the reaction times in response to S1 and the reaction times in response to S2, it was deemed important to attempt to investigate the data as others had done. Using a Klintman-style analysis (using just four reaction times from each participant), no TRI effect was found. However, using a Camfferman-style analysis (using means instead of the pairs), analysis *does* show a significant TRI effect in the predicted direction. Along with this significant result there also appears to be considerable evidence that the conclusion of Camfferman (1987) regarding the influence of general alertness has also been confirmed in this study. If one compares the current results

with Camfferman's, there is a good deal of correspondence between the two independent studies and thus opinion must be that the same pattern of scoring is a good indication that the paranormal conclusion is probably not correct. Although as with Camfferman's study, the pattern of scoring also supports the hypothesis of Klintman (1983, 1984). It is the plausibility of the hypotheses that dictates which is the more likely explanation.

Finally, the Radin and May-style analysis is provided to allow comparison amongst the main papers in the area and shows that again, although there is evidence for a conventional Stroop-based influence in the data, there is certainly no evidence for a TRI-type effect.

It is quite interesting that we have seemingly found a TRI influence using the method of analysis as used by Camfferman (1987), as opposed to that used by Klintman (1983, 1984). Klintman argues that to use more of a participant's data could lead to correlations that are a consequence of an artificial inflation from influences that have effected RT1 and RT2. Had we found a TRI influence using Klintman's criteria this criticism might be more troublesome. The method of analysis used by Radin and May (2000) is also supposed to take into account individual variation. Since neither method has given us a TRI influence it is hard to evaluate that claim. Given the mixed findings, further experimentation to evaluate the different hypotheses was justification enough to pursue another study.

3.2 Time-Reversed Interference II

3.2.1 Introduction

The results of TRI I were certainly intriguing, and yet came at no small expense. Specifically the voice-key procedure had proved itself to be remarkably similar to the skin-conductance measure used in the presentiment research described in Chapter 2, in that it required little or no conscious input from participants, but attrition and technical difficulties made the data itself costly to acquire. In the case of the voice-key data, each session had to be manually checked to ensure that the voice-key had been triggered by an actual response, rather than some erroneous source of noise (a time consuming process). This fact alone led to the decision that, rather than directly replicate TRI I, the essential methodology would remain the same, but the responses would be manual rather than verbal, allowing reaction times to be accurately recorded using a computer keyboard and thus enabling a larger number of participants to be tested.

Using a manual task instead of verbal responses to investigate TRI has some pedigree within TRI research. Klintman (1983, 1984) himself only conducted one colour-based experiment measuring verbal responses, before moving on to his studies II-V and using manual button presses in testing for the TRI effect. Radin and May (2000) exclusively used a manual task in their four TRI studies and reported finding a significant TRI effect using keyboard presses. In his large-scale evaluation of Stroop research, MacLeod (1991) argued that interference effects can be shown through the manual task, although it is considered not as effective as verbal measures. This suggests that if TRI influences are to be found in Stroop-based tasks, they should not be exclusive to verbal tasks, but should also be found in manual tasks.

One criticism of the Radin and May (2000) study was that their participants had to respond to the initial colour-patch by pressing the key that corresponded to the first letter of the colour-word (e.g. a participant should respond to a green patch by pressing the "g" key on

the computer keyboard). As has previously been discussed, this resulted in participants having to move to different parts of the keyboard for different colour-patches. Because Radin and May did not argue that the TRI effect would be identical on each of the colours, and because of their methodology, they analysed each colour separately, although the general TRI claim being tested was still essentially that D1 and D2 (or in Radin and May's case ΔZ_{RT1} and ΔZ_{RT2}) would be correlated.

Rather than use the methodology used in Radin and May (2000), TRI II favoured a standardisation of the response key, that is, that four keys at the bottom of a computer keyboard were identified with coloured stickers (the "z", "x", ">" and "/" keys). If we contrast the two possible ways in which the manual responses could have been recorded, we can see that each of them has pros and cons:

1. The participant is presented with a green-coloured block and presses the "g" key to respond (the Radin and May method).
 - (a) Each colour stimulus requires the participant to make a different manual movement to press the appropriate key (a con).
 - (b) However, given relative familiarity with a computer keyboard, the participant should have a well established association between the task (the "g" key must be pressed) and the corresponding motor-movement to make the appropriate response (a pro).
2. The participant is presented with a green-coloured block and presses a key standardised on the keyboard by a coloured marker (the method favoured in this experiment).
 - (a) The motor-movements required to make each colour response are minimally different (a pro).
 - (b) However the participant has a minimal period within which to associate the response required with the key that should be pressed (a con).

If we assume that participants are going to require some visual assessment of the keyboard to locate the correct button (i.e. they are not a touch-typist) it might be argued that this somewhat favours the technique used in the current study. However, Radin and May (2000) tested participants over a network, using a variety of computers. In that case it seems more reasonable to use their selection of keys, where presumably it would be difficult to ensure keys be coloured (as they were in the current study).

With this minor methodological difference, the hypotheses remained the same as TRI I. We were to investigate the TRI effect first with a multi-level analysis, and then follow up that analysis by comparing the results of Klintman-, Camfferman-, and Radin and May-style analyses.

3.2.2 Method

3.2.2.1 Design

TRI II employed a within-subjects design where participants' reaction times for identifying colours were compared for congruent trials and incongruent trials.

3.2.2.2 Participants

There were 50 participants consisting mainly of undergraduate students at various levels of a psychology degree at Goldsmiths College, University of London (22 males and 28 females). Participants had a mean age of 29 years (SD 11 years, ranging from 16 to 53). Where in TRI I participants were able to be divided into facilitators and inhibitors based on whether their D2 scores were positive or negative, such a division is not possible as only 2 participants have a $D2 > 0$ (inhibitors), with the other 48 having a $D2 < 0$ (facilitators).¹³

3.2.2.3 Materials

A desktop PC was used to run the experimental program, which made use of standard pseudo-random algorithms seeded from the internal clock. The software was the same as that used in TRI I, except that the input had been changed from an external voice-key, to keyboard responses. The PC keyboard had the letters 'z', 'x', '>' and '/' colour-coded (with a small coloured sticker) to red, green, yellow and blue respectively. Participants were advised to place two fingers from the left hand over the 'z' and 'x' keys, and two from their right over the '>' and '/' keys, though there was no method of ensuring this suggestion was followed.

3.2.2.4 Procedure

Participants were first directed to the testing computer and asked to read the instructions outlining the procedure they were expected to follow. When the participant had confirmed understanding of the instructions, five practice trials were initiated. A single trial consisted of a fixation point appearing in the centre of the display (marked with a white-coloured + on a black background and lasting .5 seconds), after which a coloured block was presented in a random colour (red, green, yellow or blue) on a black background. The display was cleared after the participant had pressed the respective colour-coded key and a second fixation point was presented, again lasting .5 seconds. This was followed by a random colour-word ("red", "green", "yellow" or "blue") written in white and again presented on a black background. There was a 50 percent chance that the trial would be congruent (that is the colour-block would match the colour-word) or incongruent.

After five practice trials, the participant completed 20 experimental trials, which lasted approximately 1.5 minutes, although again this varied due to individual differences in response times. After the last trial, the participant was debriefed. A total of 1000 trials were completed (50 participants contributed 20 trials each). Because of the button input, Camfferman's (1987) criteria for exclusion were not used and only incorrect trials were eliminated. This led to 92.7 percent of the data being available for analysis.

3.2.3 Results

As the split of facilitators to inhibitors was so very uneven (48 to 2), it was not possible to repeat the same ANOVA conducted for TRI I (with responder-type as a factor). The reason for this lack of inhibitors will be examined further below. However, the very lack of contrary

¹³Although this applies to D1 and D2 scores as calculated in the Camfferman-style analysis. To avoid confusion, although Klinton (1983, 1984) categorises responders according to whether their D2 scores are positive or negative (which indicates whether they have faster reaction times on congruent or incongruent trials in response to S2), he bases this categorisation on two reaction times. Camfferman (1987) uses the mean scores for all trials and thus gives a more accurate picture of the participants' general responder behaviour. It is this more general categorisation to which we are referring here.

responders should make an analysis based on that particular categorisation of participants unnecessary. As was done for TRI I, participant scores were standardised. Table 3.8 shows these mean standardised scores across congruence for RT1 and RT2.¹⁴ Examination of this table shows an obvious colour-congruence Stroop based response at RT2. Interestingly the congruent RT2 scores are much quicker than those seen in TRI I, but the incongruent scores are much slower. Although the RT1 scores are a similar magnitude to the results obtained in TRI I, they are in the opposite direction to that predicted by the TRI hypothesis.

Table 3.8: Mean results for TRI II, showing standardised reaction times across the Stroop-based task.

Response	Colour-Congruence	\bar{x}	σ
RT1	Congruent	0.17	0.59
	Incongruent	-0.18	0.55
RT2	Congruent	-1.20	0.76
	Incongruent	1.21	0.78

Statistical analysis of these scores is simple. Two t-tests were conducted, one upon the RT1 scores (testing the TRI hypothesis) and the other on the RT2 scores (testing the conventional Stroop-based influence). The paired t-test for RT1 was non-significant (comparing congruent vs incongruent scores): $t(49) = 2.15$, $p = ns$ (one-tailed). The corresponding paired t-test for RT2 was highly significant: $t(49) = -11.18$, $p < .0001$ (one-tailed). Thus there is good evidence for a conventional Stroop-based influence in the data, but no corresponding TRI influence.

In regards to power, it is not directly apparent what effect size we might expect given Klinton (1983, 1984) asserts that a difference of up to at least 65 ms difference is expected between the congruent and incongruent RT1s, because that assertion was made in regards to the verbal task and not the manual task used here. If we use this 65 ms difference in our power calculation, with the observed standard deviation, we calculate the power of TRI II to be .67, and is therefore under-powered (mainly due to the large observed standard deviations). It might be argued that this is a conservative estimate of the study's power, because although we cannot then be sure that the difference will be as large as 65 ms, it is true that the difference between the congruent and incongruent RT2s is significantly *larger* than that seen in TRI I. Given the TRI hypothesis essentially argues that what is true for responses to S2 is also true for responses at S1, we would thus expect to find a comparably large difference at S1 (and this is not confirmed). Again, this analysis corresponds to the fact that given the TRI hypothesis, we might expect a significant difference between congruent and incongruent RT1s. As has already been stressed in regards to TRI I, the other authors in the area have established the TRI effect through a significant correlation between a–b and c–d, which does not itself depend upon a statistically significant difference between the two scores (see Figure 3.1).

3.2.3.1 Klinton-style analyses

Although we appear to not have found a TRI influence in our data, following the analyses engaged in TRI I (see Section 3.1.3.1), it is of interest as to what the various analyses employed previously show for these data. Is there a significant correlation between D1 and D2? For the

¹⁴It should be recognised that there is a latency period between participants making the appropriate response (pushing a button on a computer keyboard) and that response being registered within the computer system. Although no attempt was made to investigate this latency period, other researchers (e.g. Shimizu, 2002) have reported that the average latency period for computer keyboards is approximately 30ms. Thus any difference smaller than this period could of course be a consequence of this latency period, though there *should* be no systematic influence on the results.

Klintman-style analysis, D1 and D2 were calculated using just two pairs of reaction times, as outlined by Klintman (1983, 1984) (see above). Because just two pairs of reaction times are used in the calculation of D1 and D2 in Klintman's analysis, it is clear that although using a particular participant's mean congruent RT2 and mean incongruent RT2 may categorise them as a facilitator or inhibitor, the same is not true for Klintman's pairs. This is the reason why we have a split of facilitators and inhibitors for this Klintman analysis and not for the Camfferman-style analysis, to be presented below. The results of this analysis are presented in Table 3.9. Again, this is a contingency table, where participants are assigned to cells based on the direction of their D1 and D2 scores. The highlighted cells are those where Klintman expects his results to load if there is a TRI effect. We see no obvious loading in the highlighted diagonals (and it is true that there is a similarity between this table and the corresponding table presented for TRI I).

Table 3.9: Klintman analysis for TRI II

	D1 > 0	D1 < 0	
D2 > 0	9	7	16
D2 < 0	21	13	34
	30	20	50
Fisher's exact $p = .47$ (one-tailed)			

Describing this table in words, again we first concentrate on the top row, relating to the inhibitors. We see a slight preference (2 participants more) for inhibitors ($D2 > 0$) to also have a $D1 > 0$, though the difference is unlikely to be anything other than chance (and a chance 50/50 split is what we found in TRI I for the same contrast). Considering next the bottom row (the facilitators), we see that there are almost twice as many facilitators with a $D1 > 0$ than with the $D1 < 0$ (that Klintman, 1983, 1984 expected). This is what we found for the same data for TRI I.

Although this table appears to be directly comparable to that obtained for TRI I, subsequent analysis (specifically the Camfferman, 1987 analysis) has highlighted an unanticipated problem regarding the interpretation of the Klintman-style analyses. Put simply, the manual task employed in this study appears to have introduced a difficulty (identifying a stimulus colour through a button press) and that this difficulty has resulted in nearly all participants reacting in a conventional manner (Stroop-based), thus inadvertently removing a crucial contrast (between facilitators and inhibitors). This point will be further elucidated as the analyses progress, but it should be noted that in all fairness this problem makes interpreting the Klintman table difficult, because we are essentially missing one side of the analysis (the contribution of the inhibitors). Thus, though it *appears* that this table and the corresponding table in TRI I are very similar (both show a similar pattern of scoring), we now believe that the differences in the task (manual versus the verbal) have resulted in a pattern of scoring which makes the Klintman-style analyses less appropriate for this study. Although less appropriate, the analyses *can* and *do* provide a wealth of information about what is going on in the study.

In relation to this table, Klintman (1983, 1984) reports the group medians for D1 and our corresponding values are for $D1 > 0$: 268 ms and for $D1 < 0$: -189 ms. The overall Spearman rank correlation between D1 and D2 values for the 50 participants was $r = .06$, $p = ns$. The Fisher's exact test is also not significant. Thus, as we found in TRI I, using Klintman's criteria, there is no evidence for a TRI effect (although given the problem identified above, it would be premature to conclude that TRI is absent in the data).

Next we repeat the Camfferman-style analysis (which to remind the reader, is comparable to the Klintman (1983, 1984) analysis, but uses participant mean scores rather than just four reaction times). These data are presented in Table 3.10. Again we would expect the values to load in the highlighted diagonals if there is a TRI influence in the data and this is not shown. This contrasts with TRI I, where the same table *did* load in the appropriate cells. It is clear that part of this disparity is a result of very few of the participants being categorised as inhibitors, thus scoring faster at RT2 incongruent than RT2 congruent (2 against 48).

Table 3.10: Camfferman analysis for TRI II

	D1 > 0	D1 < 0	
D2 > 0	0	2	2
D2 < 0	31	17	48
	31	19	50
Fisher's exact $p = .14$ (one-tailed)			

Describing the table in words again, we first consider the top row. We can see that only 2 participants have a $D2 > 0$ (inhibitors) and of those two, they both have a $D1 < 0$ (though given the first inhibitor being assigned to the appropriate cell, there is a 50/50 chance that the next inhibitor will also share that cell, thus no conclusion should be drawn about this distribution). Considering the bottom row (the facilitators), we see that again we have almost twice as many participants with a $D1 > 0$, than $D1 < 0$, and this is in the opposite direction to that predicted by Klintman (1983, 1984). As raised above, the lack of inhibitors makes interpreting this table much more difficult, because we only expect cells to load in the diagonals in the presence of idiosyncratic responders (that is, we cannot expect the table to distribute in the way Klintman predicted, given the loss of the top row). And to reiterate, we lack inhibitors presumably as a consequence of the task itself, rather than the participant population (which is similar to that of TRI I). It seems likely that the relative difficulty in responding to the stimuli (identifying the appropriate key) has led to all participants reacting in a conventional Stroop-based manner, at the expense of massively increasing the variance of the responses.

The overall Spearman rank correlation between $D1$ and $D2$ values for the 50 participants was $r = .06$, $p = ns$. The Fisher's exact test is not significant. Thus, this analysis also fails to show a Klintman-type finding and thus does not provide evidence for a TRI effect (although again given the problem identified above, it would be premature to conclude that TRI is absent in the data).

3.2.3.2 Camfferman's "Alertness" hypothesis

As we have found no evidence for a correlation between $D1$ and $D2$ in this study, it is still of interest as to whether we *still* have evidence for the prosaic 'alertness' hypothesis (Camfferman, 1987). Following Camfferman's analysis again, we find that $D2$ is more often negative than positive, where 48 participants showed $D2 < 0$ against 2 with $D2 > 0$ (highlighted in orange), $z = 9.20$, $p < .0001$, which is highly suggestive of a cognitive interference effect.¹⁵

With the preponderance for participants to have a $D2 < 0$, if time-reversed interference is present most participants will be expected to show $D1 < 0$. Looking at Table 3.10 we can see that some 38 percent (19/50) of $D1$ values are negative (highlighted in green). And this result is not significantly different from chance expectation, $z = 1.70$, $p = ns$. Again, this can be

¹⁵We have already noted that this congruence effect appears to be a consequence of the manual task.

interpreted as not supporting Klintman (1983, 1984). Although in TRI I the numbers marked in green were an exact 50/50 split, this is not the case here, the meaning of this particular result will become clearer later on in the analysis.

Although we have not established a TRI-type influence in the data, it will again be of interest to follow the Camfferman (1987) analysis and investigate the influence of his alertness hypothesis. Again we can reconceive of the Klintman (1983, 1984) variables $D1 < 0$ and $D1 > 0$ in the same way as we did for TRI I. To remind the reader, for Camfferman it does not make sense to consider $D1$ scores in relation to trial congruence (which is not established until S2). Instead, if we assume, like Camfferman, that the minimum information we have about individual participant responses is contained in their $D1$ scores, that a $D1 < 0$ means that the participant was faster on the congruent RT1 scores than incongruent (and vice versa for $D1 > 0$). With that reconception, we can produce a table that both retains the shape of Klintman's table (that is participants who have a $D1 < 0$ are now considered RT1 fast) but presents how participants react only in congruent trials. This is Table 3.11. Camfferman then converts the same data into relative frequencies, showing the proportion of participants who respond fast or slow at each stimulus.

Table 3.11: Frequency of slow and fast reaction times for Camfferman analysis for TRI II

	Congruent Trials			Relative Frequencies		
	RT1 slow	RT1 fast	Total	RT1 slow	RT1 fast	Total
RT2 slow	0	2	2	0	.11	.04
RT2 fast	31	17	48	1	.89	.96
Total RT1	31	19	50	.62	.38	1

Describing the table in words, first we can see that as reaction times to S1 should not be influenced by trial congruence (established at S2 onset), whether a participant has been designated as having a fast RT1 or a slow RT1 should be random. We can see that there is a slight bias in this comparison: participants who reacted fast and were assigned a congruent trial at S2 (38 percent) and those who reacted slow and had a congruent trial assigned to them (62 percent) (proportions highlighted in green). Because we are just looking at the congruent trials, this slight bias is again likely a consequence of the slight difficulty in responding to the stimulus. Next, because we are considering just congruent trials, we can see that some 96 percent of participants (highlighted in orange) show a conventional Stroop-based response, that is they were aided at S2 by congruent stimuli (compared to just 4 percent who were not). This again suggests that the relative difficulty of the task has righted idiosyncratic responders. Finally, we see that of those who already had a fast reaction time to S1, the majority, 89 percent also reacted fast to S2 (highlighted in blue). In those already slow at S1, 100 percent were fast in their reaction to S2. Camfferman (1987) found the proportion highlighted in blue to be significantly different, where in our data that difference appears to have been swamped by a colour congruence effect.

Again, Camfferman (1987) constructed a final dataset to establish a direct association between RT1 and RT2 on each trial, for each participant. This particular dataset is not beset by the 'problems' identified above, as will become clear. All trials for a participant were pooled and for each trial, each RT1 and each RT2 was labelled as 'fast' or 'slow' as thus: If the RT1 is below the overall median RT1 for that participant then RT1 is labelled as 'fast', otherwise it is 'slow'. The same is used to calculate the 'speed' of RT2. Scores that fall on the median are

omitted and the participant median is recalculated. These data are presented in contingency Table 3.12. This table presents relative frequencies and shows clearly the coherence of RT1 and RT2, or 'alertness'. Thus, regardless of the colour-congruence of the trial, there is a clear relationship between reactions to S1 and S2: Fast RT1s go significantly more often together with fast RT2s, $z = 2.29$, $p = .01$, one-tailed. If a participant reacts fast to S1 they are 56 percent likely to react fast to S2 and if they are slow at S1, they are 56 percent more likely to react slow at S2. This is the same pattern we found for TRI I (and to reiterate, is independent of the 'problem' identified with the task). There is a much smaller number of trials than the total number (1000) because we remove median trials and then recalculate the score.

Table 3.12: Frequency of slow and fast reaction times (RT2 as a function of RT1) for all trials for TRI II

	RT2 slow	RT2 fast
RT1 slow	.56	.44
RT1 fast	.44	.56
N = 364		

Following the Camfferman (1987) analysis, lastly if we differentiate the data in Table 3.12 according to congruence, we arrive at Table 3.13. This table (based on trial by trial responses) elucidates clearly that although the task itself may have introduced a bias (by increasing difficulty of the response) it *has* been very useful to repeat the analyses conducted for TRI I. Describing this table in words, first this time it is clear that there is a 50/50 chance of having a slow or fast RT1 and being assigned to a congruent or incongruent trial.¹⁶ Like Camfferman we can test this using a z-difference score by comparing the 50 percent who are fast at RT1 congruent, with the 50 percent who are fast at RT1 incongruent (both highlighted in green). Obviously this contrast is unnecessary as there is no difference.

Table 3.13: Frequency of slow and fast reaction times (RT2 as a function of congruence and RT1) for all trials pooled for TRI II

Congruent Trials			
	RT2 slow	RT2 fast	Total
RT1 slow	.37	.63	.50
RT1 fast	.20	.80	.50
Total	.29	.72	
N = 369			
Incongruent Trials			
	RT2 slow	RT2 fast	Total
RT1 slow	.76	.24	.50
RT1 fast	.68	.32	.50
Total	.72	.28	
N = 359			

Secondly, by comparing the 72 percent in the congruent trials who are fast at RT2 with the 28 percent in the incongruent trials that are fast (both marked in orange), we calculate the z-difference score using a z-score for each sample compared with chance (to make it directly

¹⁶This clarifies that the slight bias we found in Table 3.11 did not indicate a TRI influence. Even if there is a bias to being slow at RT1, it does not distribute in a non-chance manner in response to colour-congruence for S1 reaction times.

comparable to the reported results in Camfferman, 1987), $z_{diff} = 11.87$, $p < .0001$, one-tailed. Thus unlike the same analysis at TRI I, there is clear evidence that RT2 scores are significantly dependent on the colour-congruence of the trial. Thus, the problem that we have identified (the difficulty with making the appropriate response), has made it *more* likely that participants will show a conventional Stroop-based response and be helped by colour-congruence of the trial.

Thirdly, for congruent trials a fast RT1 is significantly more often accompanied by a fast RT2, $z = 10.1$, $p < .0001$ and for incongruent trials a slow RT1 is more often accompanied by a slow RT2, $z = 8.12$, $p < .0001$ (both marked in blue). Again there is a marked correspondence between the results reported in Camfferman (1987) and our own, despite not finding a TRI-like effect in this data. Thus evidence for alertness remains, but TRI does not. Additionally, it is these cells that *explain* the lack of TRI influence in the presence of alertness. If we recall that alertness explains a correlation between D1 and D2 scores, then here we are presenting strong evidence for alertness, *without* a corresponding TRI-like influence. Does this not undermine our conclusion? If alertness has been found without TRI, perhaps alertness is not enough to explain TRI (as Radin & May, 2000 argue). We do not believe this is true. As we see in the table, 80 percent of participants who were fast at RT1 were fast at RT2 in congruent trials. But of those who were slow at RT2 only 37 percent were slow at RT1 (compared to 63 percent who were slow at RT1 and fast at RT2). We contend this discrepancy is a *consequence* of the difficulty of the task of responding, which *helps* responses to S2 (in congruent trials) and hinders responses on incongruent trials (and this is what the table shows). Because of that fact, we are left with clear evidence of alertness, but a disruption of what would be misconstrued as TRI. Seemingly, what had appeared to be a problem, had in conclusion, provided an even better understanding of what is going on, and why Klintman (1983, 1984) initially concluded his participant responses were paranormal (because his verbal task conflates these issues). It also seems likely that this explains the results of Radin and May. As we have already noted, participant responses in their research were arguably more natural than in the current study (for example the participants responded to a yellow block by pressing the “y” key) and that for participants familiar with using a keyboard, this response was less “difficult” than the method used in the current study (and was thus nearer to the vocal responses used by Klintman).

3.2.3.3 Radin & May-style Analysis

Finally, we investigate the data in a way that compares to that used by Radin and May (2000) (full details are outlined on page 96). To reiterate, it has been argued that other authors in the area have identified TRI through a positive correlation between D1 and D2 scores, and yet we have argued that for TRI II, the Klintman-style analyses are probably inappropriate, because we have a lack of contrary responders (which we argue is because of the task, not the participant population). However, the Radin and May-style analysis provides a final occasion to investigate our data in a much more appropriate manner, given that our procedure is also a manual task, which is practically identical to the procedure used by Radin and May. Thus, though our conclusion *not* to have found TRI in the analyses presented up to now may be criticised as premature, the Radin and May analysis meets this criticism head on.

For this analysis, essentially, a z-score is calculated for each trial, for each participant, and then a z-difference score is calculated for the combined results for congruent and incongruent trials for reaction times to S1 and S2. Z-diffs are calculated for the four different colours and Radin and May (2000) predict that there should be a positive correlation between the



Figure 3.5: Values of ΔZ_{RT1} and ΔZ_{RT2} by colour for TRI II. Colour refers to the initial colour patch (S1). The correlation between the two lines is $r = -.46$, $p = ns$, (one-tailed).

scores calculated for S1 and those calculated at S2. Table 3.14 shows the number of trials for the respective colours used in this Radin analysis, the results of which are presented in Figure 3.5. Although all the ΔZ_{RT2} scores are above a significant level (and thus show that colour-congruence has influenced RT2 scores on each of the colours) there is no significant correlation between ΔZ_{RT1} and ΔZ_{RT2} , and none of the ΔZ_{RT1} scores are significant. In their most comparable study Radin and May present a correlation of $r = .93$, $p = .037$ (one-tailed) between ΔZ_{RT1} and ΔZ_{RT2} (with significant ΔZ_{RT2} scores and non-significant ΔZ_{RT1} scores). Thus, from the perspective of the Radin and May analysis, no TRI influence has been shown.

Table 3.14: Number of trials across colour for TRI II Radin analysis

	Congruent	Incongruent
Red	110	112
Yellow	119	105
Green	109	116
Blue	122	134

3.2.4 Discussion

The results of TRI II can be summarised thus: a multi-level analysis of variance showed that colour-congruence significantly influenced RT2 scores, but not RT1 scores. A Klintman-style analysis failed to show a significant correlation between D1 and D2, as did a related Camfferman-style analysis. Nor was a paranormal effect shown with an analysis based on Radin and May (2000). Not only was a significant conventional Stroop-based effect found but,

there is yet again good evidence that the alertness hypothesis is an important factor to consider (that fast RT2 scores follow fast RT1 scores, and vice versa) and again, there is remarkable correspondence between the results of this study and those obtained by Camfferman (1987). Given these findings it seems clear that it is safe to conclude that RT1 and RT2 are not independent measures, as Klintman (1983, 1984) believed them to be, and that we are finding here inter-dependency interactions in the cognitive task.

The results of TRI II also indicate a major difference between this study and the previous experiment. As has been stressed earlier, it appears that using a manual task instead of verbal responses has made making a response *more* difficult, resulting in participants having to think about the response *and* the appropriate action, at the expense of variability. From the perspective of discovering a TRI influence, it could be argued that this has possibly helped or hindered our ability to discover an effect in the data. That is, it is clear that the verbal responses (used in TRI I) are much quicker and more natural: participants do not have to think about *how* to respond. Yet we found that some participants were consistently worse at identifying colour-words in congruent trials (the preceding stimulus, S1 is congruent with the written colour-word). With the manual task we find that participants overwhelmingly show a conventional Stroop-based effect (they are nearly all faster at responding to S2 congruent trials versus incongruent trials). This is presumably because participants must visually inspect the keyboard to identify the correct colour-coded key to press, to make the appropriate response. Because this response is much more complex than merely verbalising a colour-word, it might be argued that this makes TRI *less* likely to be found at S1. The decision making processes involved in responding to S2 might disrupt the TRI influence (by reducing the automaticity of the response) and thus eliminate TRI from the data. This is one perspective. The other perspective, seemingly held by Radin and May, is that *any* response pattern found for S2 should be accompanied by a similar pattern at S1. Thus if the manual task increases the chance of discovering a Stroop-based influence in the data, it might well increase the chance of discovering a TRI influence in the data, and as Radin and May (2000) report four studies with significant TRI results using a manual task, this seems reasonable. Given this concern, it seems prudent to conclude that though we have not discovered evidence for TRI, and though alertness is of crucial concern to an understanding of responses in this paradigm, it would be premature to conclude that the lack of TRI influence (as reported in Klintman, 1983, 1984) is not itself a product of the methodology, rather than questioning the reliability of the initial observations.

It is with some minor regret that the methodology for Klintman's studies II–V was not investigated further (Klintman, 1983, 1984) and it is hoped that some future research will shed further light on whether that particular methodology can provide stronger evidence for a TRI influence. One final TRI study was conducted, this time however incorporating a novel adaptation which it was hoped would finally reveal a TRI influence using a Stroop-based task.

3.3 Time-Reversed Interference III

3.3.1 Introduction

Although we have little reason to suspect that the results of TRI I and II demonstrate any paranormal-type influence, it was hoped that by incorporating some of the ideas that had

been considered in the introductory chapter (Section 1.3.3), it might be possible to provide the elusive significant paranormal result being pursued. Specifically, TRI III was an attempt to incorporate a danger/threat element into the study, to build upon the secondary question of this thesis, that is, if psi is real, it might be related to danger or threat in the future. It must be remembered that the emphasis of this investigation is primarily parapsychological and thus, despite convincing evidence that TRI *might* be best explained by alertness, there is still the possibility that the procedure itself will provide convincing evidence that cannot be explained in terms of alertness alone. As has already been discussed in the second chapter covering the presentiment effect (Section 2.1.1), there is *some* rationale for using spider stimuli in parapsychological settings. That the procedure had not been successful when testing the presentiment effect only partially dissuaded from pursuing this line of enquiry. Despite the previously unsuccessful TRI results, it was hoped that by incorporating spider-based stimuli into a test of cognitive interference, that a paranormal effect might be found.

As such, TRI III involved significant methodological deviation from the original Klintman (1983, 1984) study, although it does provide an exploratory conceptual replication. Instead of making use of a simple colour-based Stroop task, it was hoped that an emotional-Stroop would produce a more interesting result. It has been found in numerous previous studies (e.g. Richards, French, Johnson, Naparstek, & Williams, 1992) that anxious individuals take longer to identify the colour of threatening words compared to neutral words. Such effects are not found in low-anxiety individuals. The current methodology borrowed heavily from Watts, McKenna, Sharrock, and Trezise (1986), where an emotional-Stroop was developed using spider and neutral words. Spider phobics took longer to identify the colour of spider-related words compared to neutral words. Watts et al. claimed to have demonstrated a very substantial spider-Stroop effect in participants with spider phobia and concluded that an emotional-Stroop is capable of providing a sensitive measure of a large effect.

As the previous spider-Stroop research has demonstrated that spider stimuli retard phobics in their conventional responses, we are thus expecting two specific patterns of responding to S2: one pattern for those afraid of spiders, and one for those not afraid. Table 3.15 illustrates the expected pattern of scoring at S2 to the different trial types. For the no-fear group we expect to see a Stroop-based differentiation at S2 between congruent and incongruent trials (slow on incongruent trials and faster on congruent trials) regardless of emotionality. For the fear group we expect the same differentiation on the neutral trials, but that the spider trials will retard responses. Responses to S1 should be 0 (the null hypothesis) or follow the same pattern of scoring in the presence of a TRI influence. In this experiment then, we are using the spider-fear group not to attempt to *strengthen* any psi-mediated influence (as was done with the presentiment experiment in the previous chapter, and will be described in the next chapter, on psi-timing). Instead we are using spider-fear to retard performance. It is clear that if we demonstrate a pattern of scoring as indicated in the table, no normal explanation, not even alertness could explain it. We would therefore have strong evidence in support of Klintman's original claims for precognitive, time-reversed interference.

It should be noted that processing the stimulus word for emotionality represents an abstract decision making process, somewhat akin to that used in Klintman's studies II-V (Klintman, 1983, 1984), where he argued that such an abstract process might strengthen the TRI influence:

It seemed likely that such interference might not restrict itself to these relatively simple events [Stroop-based colour-congruence] but might indeed become even

Table 3.15: Hypothesised TRI III RT2 results. Numbers are illustrative. The null hypothesis predicts 0 for all scores for RT1, where a TRI influence should follow a similar pattern (but smaller magnitude) than these illustrative results.

	Colour-Congruence	Emotionality	Spider-Fear	
			Fear	No-fear
Congruent		Spider	0	-0.5
		Neutral	-0.5	-0.5
Incongruent		Spider	1	0.5
		Neutral	0.5	0.5

more pronounced if more complex and less stimulus bound events were used, such as for instance the comparing of two stimuli with respect to some abstract characteristic such as 'odd/even'. (Klintman, 1983, p. 30)

Although participants are not being asked to process the words for meaning, the spider-Stroop argues that participants automatically do, and we would thus expect the two spider-fear groups to be influenced differently by the use of such stimuli, both in their conventional responses to S2 and in their hypothetical TRI response.

Due to the more complicated methodology introduced by using the emotional-Stroop, it is difficult to analyse the multi-level data and at the same time remain close to that used in Klintman (1983, 1984) and only the most basic attempt has been made to directly investigate whether there exists a correlation between RT1 scores and RT2 scores in the way previously done for TRI I and II. As such, for TRI III we are interested in whether we can differentiate participants by their responses to S2, according to whether the trial is congruent or incongruent in regards to colour, *and* whether reported spider-fear also influences responses to S2 in regards to emotionality of the stimulus (spider-based or neutral). If we *can* differentiate participants on those dimensions in the conventional direction, it is hoped that a similar relationship holds for their responses at S1, something that would indicate a TRI influence (again refer to Table 3.15).

3.3.2 Method

3.3.2.1 Design

TRI III employed a $2 \times 2 \times 2$ mixed design, comparing 2 within-subjects factors: colour-congruence (congruent versus incongruent trials) and emotionality (spider versus neutral), and one between-subjects factor of spider-fear group (fear versus no-fear). An identical ANOVA analysing these variables would be conducted separately for each of the stimuli (S1 and S2).

3.3.2.2 Participants

Some 54 participants volunteered, mainly undergraduate psychology students at Goldsmiths College, University of London (12 males and 42 females), where ages ranged from 17 to 50 years with a mean of 25 years and a standard deviation of 7 years. There were 20 participants in the spider-fear group and 34 in the no-fear group (as measured by the Fear of Spiders Questionnaire). None of the participants were idiosyncratic responders to colour-congruence (i.e. were inhibitors or had a $D2 > 0$) and as such it is not necessary to compare facilitators with inhibitors.

3.3.2.3 Materials

A desktop PC was used to run the experimental program, which made use of standard pseudo-random algorithms seeded from the internal clock. Again this was the same program as had been used in the first two experiments, but was adapted to meet the needs of the current methodology. The program had been constructed in *E-prime*, a psychology experiment construction package. The PC keyboard had the letters 'z', 'x', '>' and '/' colour-coded to red, green, yellow and blue, respectively. Participants were advised to place two fingers from the left hand over the 'z' and 'x' keys, and two from their right over the '>' and '/' keys, though there was no method of ensuring this suggestion was followed during the procedure.

Additionally, three questionnaires were administered: the State-Trait Anxiety Inventory (Form Y) (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983), the Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961)¹⁷ and the Fear of Spiders Questionnaire (Szymanski & O' Donohue, 1995), where one or more affirmative response categorises the participants as spider-fearful.

3.3.2.4 Procedure

Participants were first required to fill in the Spielberger-State questionnaire. They were then directed to a computer in the testing cubicle and asked to read and acknowledge understanding of the instructions. The participants were then presented with a screen showing coloured noughts, coloured red, green, yellow or blue, presented on a black background and accompanied by its respective colour-word. This was to avoid any confusion about which colour was which. Noughts were used instead of the coloured blocks used in the two previous experiments, primarily to provide a similar visual weighting to the stimuli, since the procedure involved showing coloured words rather than just a colour-word written in white. After the colour presentation, the experiment proper began.

Participants were first given 16 practice trials where coloured noughts (red, green, yellow or blue) were followed by a colour-word ('red', 'green', 'yellow', or 'blue') written in one of the aforementioned colours. These practice trials were non-emotional and participants were unaware of the use of spider stimuli until presentation of the relevant words in the experimental phase. After the practice trials, the experimental phase began. A single trial consisted of a fixation point appearing in the centre of the display, marked with a white-coloured + on a black background and lasting .5 seconds, after which five coloured noughts were presented in a random colour (red, green, yellow or blue) on a black background. The display was cleared after the participant had pressed a key identifying the colour of the noughts. After the response, a second fixation point was presented, again marked with a white-coloured + on a black background and lasting .5 seconds, and was followed by a stimulus word. Importantly the word consisted of either a neutral word or a spider-related word, as used by Watts et al. (1986). The word was presented in a random colour, again red, green, yellow or blue, and displayed until a response was made. After the practice trials, the participant completed 80

¹⁷Although the Spielberger and BDI were administered, they were not used in the analysis, although in the presence of a significant emotional-Stroop effect, they would have been. Summary results from these questionnaires can be found in Appendix A.

experimental trials which lasted for approximately ten minutes.¹⁸ After the last trial was completed, participants were required to complete another Spielberger-State questionnaire, along with a Spielberger-Trait questionnaire, the Beck Depression Inventory and the Fear of Spiders Questionnaire.¹⁹

A total of 4320 trials were completed, and after incorrect responses were removed (a loss of 212 trials) some 95 percent of trials were available for subsequent analysis.

3.3.3 Results

Summary statistics are presented in Table 3.16, showing the standardised means and standard deviations across emotionality, colour-congruence and fear group. These numbers should be compared to Table 3.15 which presents what we *expected* to find (both conventional spider-Stroop and TRI). In short, we expected to find a large conventional Stroop effect at S2 (faster on congruent trials than incongruent trials) for all participants *except* on the spider stimuli and then *only* for the spider-fear group. In those participants who show a significant Stroop-based effect at S2, we expect to find a significant (but smaller) TRI influence in their responses to S1. Given what we have concluded for TRI II (that the task itself has improved evidence for a Stroop-based influence) we would expect to see a strong colour-based result here in this study and this is certainly suggested by the results. However there does not appear to be a spider-Stroop effect. The results at S1 are essentially 0 and this fits with the null hypothesis (and thus there appears to be no evidence for TRI).

There were 1080 S1 trials of each colour (red, green, yellow & blue), and at the onset of each trial there was a 50/50 chance of a congruent or incongruent trial, and a 50/50 chance of the trial being assigned as either spider or neutral.

Analysis proceeded using the 'ez' package in R (Lachapelle & Healey, 2010). Two $2 \times 2 \times 2$ ANOVAs were run, one for RT2 scores (looking for a conventional emotional Stroop-based influence) and the other on RT1 scores (looking for a corresponding TRI-type influence). The result of the ANOVA for RT1 was wholly non-significant, the results of which can be found in Table 3.17. The corresponding ANOVA for RT2 found a statistically significant main effect of colour-congruence and an interaction between colour-congruence and emotionality. The full results of this ANOVA can be found in Table 3.18. Given the significant interaction, a 2×2 ANOVA was repeated, this time omitting spider-fear as a level of contrast. The result of this ANOVA was a statistically significant main effect of colour-congruence, $F(1, 53) = 452$, $p < .0001$. There was no main effect of emotionality, $F(1, 53) = 2.26$, $p = ns$, nor was there an interaction between emotionality and colour-congruence $F(1, 53) = 3.36$, $p = ns$. This significant main effect of colour-congruence is presented graphically in Figure 3.6. Finally, a paired t-test

¹⁸When a paper based on this chapter was submitted to and rejected by the *Journal of Parapsychology* (circa 2002), both anonymous reviewers made the criticism that the use of a large number of practice trials had negated any psi-influence. However, both reviewers had misunderstood that the practice trials were *non-emotional* and thus participants were *not* exposed excessively to the stimuli. MacLeod's (1991) review of Stroop-based research suggests that practice trials and length of study should not impact overall performance to the detriment of the emotional-Stroop effect. For example, Watts et al. (1986) used around 100 words and 100 trials for each participant, a much larger number than used here.

¹⁹As was noted in the presentiment chapter, administering the FSQ after participants have been presented with spider stimuli could mean that responses are biased by the preceding exposure. Because the questionnaire has to be administered at one point (either before or after) and because it was deemed important for participants to be uninformed as to the spider aspect of this experiment, it was pragmatically decided that the measure would be given after the experiment. It was also decided that given the very general nature of the measure (that one or more affirmative score would put the participant in the spider-fear group) that a participant who has undergone the experiment and still does not answer any of the questions in the affirmative is very likely to be the very sort of no-fear participant being sought.

Table 3.16: Summary statistics (standardised reaction times and respective standard deviations) across spider-fear, emotionality, stimulus and congruence for TRI III

Spider-fear	Emotionality	Stimulus	Colour-Congruence	\bar{x}	σ
Fear	Spider	S1	Congruent	-0.01	0.94
			Incongruent	-0.01	0.96
		S2	Congruent	-0.50	0.77
			Incongruent	0.52	0.95
	Neutral	S1	Congruent	0.05	1.04
			Incongruent	-0.03	1.04
		S2	Congruent	-0.42	0.84
			Incongruent	0.44	0.94
No-Fear	Spider	S1	Congruent	-0.02	0.95
			Incongruent	0.02	1.03
		S2	Congruent	-0.41	0.83
			Incongruent	0.48	1.01
	Neutral	S1	Congruent	0.02	1.00
			Incongruent	-0.02	1.00
		S2	Congruent	-0.45	0.78
			Incongruent	0.39	0.93

Table 3.17: ANOVA results for RT1 for TRI III

Effect	DFn	DFd	F	p
spider-fear	1	52	0.35	0.66
colour-congruence	1	52	0.56	0.46
emotionality	1	52	0.10	0.75
spider-fear \times colour-congruence	1	52	0.41	0.52
spider-fear \times emotionality	1	52	0.09	0.77
colour-congruence \times emotionality	1	52	1.60	0.21
spider-fear \times emotionality \times colour-congruence	1	52	0.00	0.98

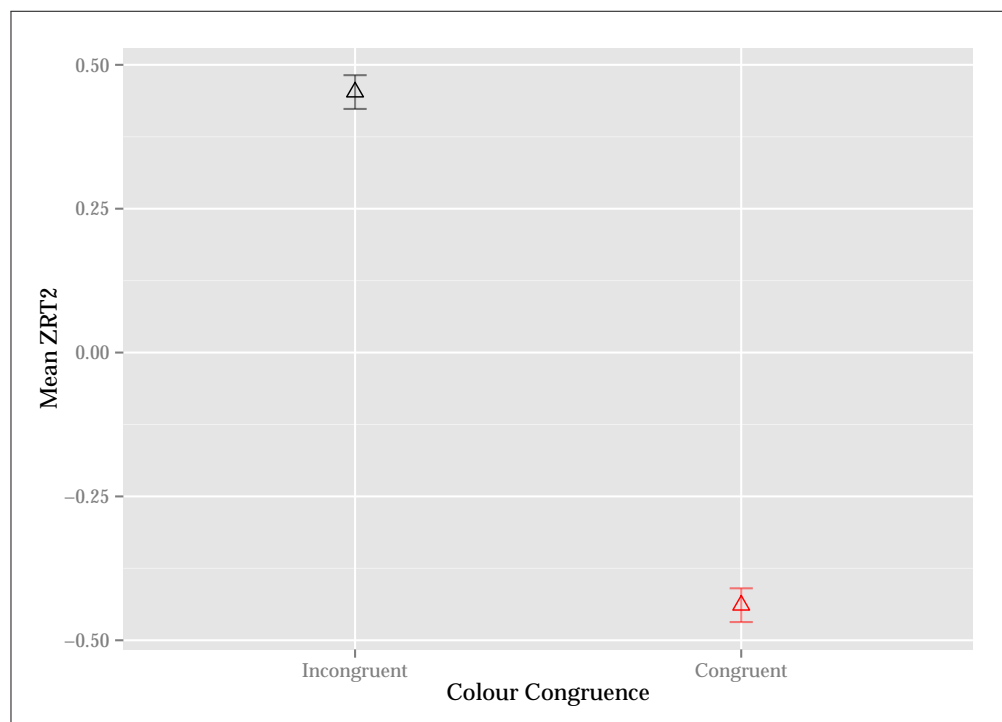
Table 3.18: ANOVA results for RT2 for TRI III

Effect	DFn	DFd	F	p
spider-fear	1	52	3.34	0.07
colour-congruence	1	52	429.34	<.0001
emotionality	1	52	1.41	0.24
spider-fear \times colour-congruence	1	52	0.77	0.39
spider-fear \times emotionality	1	52	1.03	0.31
colour-congruence \times emotionality	1	52	4.07	0.05
spider-fear \times emotionality \times colour-congruence	1	52	0.92	0.34

Table 3.19: Camfferman analysis for TRI III

	Spider Stimuli			Neutral Stimuli		
	D1 > 0	D1 < 0		D1 > 0	D1 < 0	
D2 > 0	0	0	0	1	0	1
D2 < 0	28	26	54	31	22	53
	28	26	54	32	22	54

can statistically investigate this main effect, which shows that there is a significant difference between the congruent and incongruent RT2 scores ($t(53) = -21.27, p = < .0001$). This is a conventional colour Stroop-based differentiation.

**Figure 3.6:** Significant Stroop influence at Stimulus 2

Considering issues of power, we can calculate the size of the difference we would be able to detect at S1. Using the 'pwr' package of R and assuming a power of 80%, a significance level of .05, and data that is paired and one-tailed, then we would have been able to detect a difference of approximately 100 ms between congruent and incongruent reaction times at S1 (based on our observed colour-congruence S2 standard deviation of 295 ms). This is slightly larger than the range specified by Klintman (1983, 1984) (though he was referring to the verbal task). This compares to the conventional Stroop effect of 225 ms seen at S2. Therefore this study has the power to detect a TRI effect of 100 ms (approximately half of the difference seen at S2). We can conclude that this study is somewhat underpowered, though again we are left with questions as to whether the strength of the conventional Stroop effect using the manual task has helped or hindered our ability to discover a TRI influence. Obviously the power calculation also neglects to incorporate the strength of the design, whereby the *lack* of TRI in the fear group for spider stimuli, in the presence of TRI in all other conditions, would be very difficult (if not impossible) to explain in conventional terms.

Finally, as argued previously, despite not finding a significant difference between reaction times to S1 based on responses to S2, the TRI effect as conceived by Klintman (1983, 1984) involves a significant correlation between D1 and D2. Although the procedure of TRI III is more complicated, it is possible to get a glimpse of what the results would look like, if written into a Klintman-type table (as used in the previous two experiments). Table 3.19 shows two contingency tables of D1 and D2 scores, split across the emotionality variable (and here we are looking at mean responses in the manner of Camfferman, 1987). We can see yet again, that there are no inhibitors, which again confirms our suspicion that it is the manual task itself which has caused the large significant Stroop-based effect. Fisher's exact tests are not reported this time, primarily because as has been discussed at length, it is inappropriate to conclude that TRI has not been demonstrated in this table (because we are missing inhibitors). Neither table suggests anything paranormal. Again we see that in both tables there are *no* participants who score as inhibitors for both spider and neutral stimuli. The majority show a colour-congruence effect (highlighted in orange). And again we see that the spider stimuli D1 scores are essentially 50/50 and for neutral stimuli there is a small preference for $D1 > 0$. Although this *might* suggest some kind of differentiation at D1 for stimuli (at S2), the two cells in the two tables (those who have a $D1 > 0$ and a $D2 < 0$) are virtually identical and this again suggests that making the required response has resulted in slower reaction times (as discussed extensively for TRI II).

Given the analysis in Radin and May (2000) involves comparing scores across colours (which in this context would also involve a spider-fear element) it was decided not to attempt any further analyses.

3.3.4 Discussion

The results of TRI III are clear. An analysis of variance found no time-reversed interference effect and despite the lack of an emotional-Stroop effect, a colour-congruence effect was found. There is no evidence that the colour-based interference, for which there is strong evidence in the data, had any influence on participants' ability to respond to the first stimulus. As such this experiment shows no evidence for a precognitive effect of time-reversed interference. The lack of spider-Stroop influence is most likely due to the fact that the participant population making up the spider-fear group comprised of individuals who answered one or more of the Fear of Spiders Questionnaire in the affirmative. Though an anecdote, this measure appears to divide the general public on spider-fear on an approximately 50/50 split. Thus it could be argued, and is probably correct, that the fear group in this study were just not afraid of spiders enough. However, were we to have recruited participants because of their spider-phobia, it would have been difficult to ensure that participants were uninformed about the use of spider stimuli in the experiment, as they were in this study.

As was raised in the discussion of TRI II, it seems clear that the strong colour-based Stroop effect in the data is a consequence of the manual task, where participants are having to visually inspect the keyboard, as compared to the verbal task used in TRI I. Yet again it seems unclear whether the use of the manual task has helped or hindered the discovery of a TRI influence. On the one hand a strong Stroop reaction at S2 should be accompanied with a strong Stroop reaction at S1. On the other hand, by requiring participants to stop what they are doing and look at the keyboard to respond, we are disrupting the automaticity of the responses, which might be argued not to be conducive to a psi-based influence.

What is clear is that any further TRI research should eschew the manual task and attempt as near as possible to replicate the button pressing procedure as used in Klintman's studies (II–V) (Klintman, 1983, 1984). If further testing of that paradigm is unsuccessful then it seems safe to conclude that Klintman was mistaken and perhaps a little overzealous in his initial observations and his subsequent silence on the topic speaks volumes. Until that research has been conducted it appears that it is premature to completely dismiss TRI as impossible, though we have found no evidence that it is anything but a highly unlikely hypothesis.

It was mentioned in the introduction to TRI I that investigating reaction times to the first stimulus is quite unusual within conventional Stroop-based experimentation. Given that, it is not surprising that the effect as described by Klintman (1983, 1984) has not been happened upon by other researchers, and that the 'alertness' hypothesis (Camfferman, 1987) has also been equally ignored. However, conventional researchers who *do* turn to the reactions to S1 should not be surprised to find unexpected correlations between RT1 and RT2, given what we now know about those measures. We *have* found evidence that RT1 and RT2 are related in the manner outlined by Camfferman and the close correspondence between our research and Camfferman's research is one of the positive conclusions of this chapter. For independent research, separated by many years between them, to show such marked correspondence suggests that even if there is some concern about the appropriateness of the statistical techniques employed by Camfferman to establish his alertness hypothesis, that the underlying argument is most likely sound.

In conclusion, there is no evidence for any paranormal effect in any of the three studies. The fact that a colour-congruence effect was found to some extent in all of the studies suggests that the mechanisms of cognitive interference were influencing normal performance. It would be hoped that any normal interference effects might be accompanied by a temporally reversed effect. Instead, these studies contribute three non-significant findings to the TRI literature. This is important because the description of the area in the current edition of *The Conscious Universe* (Radin, 1997a) is now misleading. It is hoped that Radin may communicate the state of TRI in a more accurate manner in the future.

It is worth noting that due to various circumstances a good many years have elapsed since data collection and final publication of this research. The results were at one point unlikely to ever be published and thus highlights another potential factor which could be influencing our view of TRI research. How many other replications exist and how many of those are successful replications? It is quite possible that there are a number of failed replications which have not been published and therefore represent a file drawer problem. One could well ask what became of Klintman's initial enthusiasm for the effect? Why did this psychologist conduct five parapsychological studies, the results of which seemingly impressed him, and then give up?

During the final stages of manuscript preparation, two relevant papers were discovered in the literature that appear to have some bearing on the TRI research. Humphrey and Keeble (1977) found that monkeys reacted faster under red light than under a blue light and Humphrey and Keeble (1978) suggest that the results are a consequence of increased perceptual sampling. They argue that as red light is associated with dawn and sunset, times when predators become more active, primates have perhaps evolved to increase their perceptual monitoring when exposed to such coloured light. There is a similarity between the TRI testing and the methods employed by Humphrey and Keeble. Interestingly, in TRI I red is the only colour which shows a significant ($< .05$) difference between congruent and incongruent trials, and in TRI II the colour red gives the strongest result. Whether this was a consequence of the speed

of perceptual sampling is not clear, but it is worth bearing in mind for future research that primate perceptual awareness may well differ between the stimulus colour used, and perhaps such a difference may account for any unusual result found.

Finally, it is interesting to note that this research was first inspired by Dean Radin's description of the area in *The Conscious Universe* (Radin, 1997a). In the light of the research described in this chapter, Radin's oversimplification of the area can be seen as doing parapsychology in general a disservice. It is not in the public's interest to paint complex experimental findings in a more positive light than is deserved, especially in such a controversial area as parapsychology. Braud (2000) continues this oversimplification in his paper on health and retroactive influences and describes Klintman's findings in identical detail to Radin. It is as though rather than assessing Klintman (1983, 1984) directly, that Braud is directly influenced by Radin's interpretations. That others will engage in such behaviour is certainly probable, as Radin's book represents a layperson's textbook introduction to the area of parapsychology. And yet how many readers will actually independently assess Klintman's research on their own? One would imagine very few, with even fewer actually testing the claims made. As such, readers are misled about the actual state of the research, and may be swayed into believing the paranormal claim, even when the evidence is actually very poor. That the public believe scientists, and that scientists must trust one another when assessing and reporting research, shows an inherent danger in the scientific process. No doubt parapsychologists should engage in much more critical evaluation and espouse much less strong opinions when drawing paranormal conclusions. It is our responsibility to present the research and its claims accurately, so that people are not misled as to the state of research and as such we must take great care when attempting to overturn other well established scientific findings. When presenting parapsychology findings to the outside world, perhaps restraint and understatement are preferable to oversimplification and misleading statements.²⁰

²⁰Hoebens (1986, p. 33) "The authors of such books bear a responsibility, since they must assume that, for the average reader, *this* book will be the most authoritative source of information on parapsychology he will be exposed to. General introductions, especially if written by university professors, decisively influence opinions and beliefs. That is why we may demand that the authors carefully refrain from overstating their case and give a fair presentation of the pros and cons."

Testing Psi-Mediated Timing with a Simple Computer-Based Task

4.1 Psi-Timing I

4.1.1 Introduction

With the lack of paranormal results demonstrated within the first two rounds of empirical testing, a ‘new’ paradigm was sought which would remain close to the original aims and objectives of this thesis, and provide evidence for the elusive phenomena. Specifically, the experiments outlined in this chapter were an attempt to get ‘back to basics’, and to find a simple and easy-to-conduct methodology that might provide evidence supporting a hypothesis of paranormal functioning. An extremely simple experimental procedure was found in the literature which had apparently fallen into desuetude, and appeared as though it would fit well with the testing already completed: the psi-timing paradigm. Psi-timing research involves not conscious guesses from participants, but uses their behaviour, specifically the very timing of their actions, as a means of measuring a paranormal response.¹ This type of response falls between the two effects already investigated. The presentiment effect used physiological correlates to measure unconscious precognition and the time-reversed interference methodology involved measuring the reaction times of volitional responses as a measure of another proposed precognitive influence. In the psi-timing research, conscious processing is minimised, but a behavioural measure is recorded.

Psi-timing research emerged in the 1980s, driven by technological improvements in computing. The use of computers in parapsychology has a fairly long history, with researchers like Schmidt (1969) using them to improve the methodology of parapsychological investigation—by removing the physical props that provided potentially normal explanations for findings—and to further development of novel ways of testing for psi, beyond the card-guessing techniques originally employed. A further reason for using a computer-based procedure is that it can eliminate unconscious bias, although outright fraud may still be possible (see Gardner, 1981). Of all the computer systems used within parapsychology, three are perhaps the most famous. Already mentioned is Schmidt’s experimental work, where over a number of years Schmidt tested a range of random-number generators with his ESP machine. The VERITAC machine, was built as part of a US air-force test of so-called, man/machine interactions (Smith, Dagle,

¹In some ways, Klinton’s (1984) research (Studies II–IV) were essentially a kind of psi-timing research, though they are not often identified as such.

Hill, & Mott-Smith, 1963). The VERITAC system was designed to provide a self-contained, fraud-resistant procedure, where targets were generated by the system and the participants' guesses recorded through button presses. It had three modes (clairvoyant, precognitive and GESP), and was seen by some parapsychologists as representing the future of ESP testing. Some years later, Targ, Cole, and Puthoff (1974) developed a system for automated testing of ESP very much based on the VERITAC system. However, the systems mentioned were merely attempts to digitise the ubiquitous card-guessing techniques and are not directly related to the psi-timing paradigm.

Radin and Bosworth (1985) describes their testing of two specially selected participants using a computer-based ESP task. They argued that many of the spontaneous case reports highlight the fact that some people seem to be able to avoid being at the wrong place at the wrong time² and a similar line of argument has been presented within the previous chapters of this thesis.

To some degree the task involved in the psi-timing experiment is analogous to the real-life reports. All participants had to do in the Radin and Bosworth (1985) experiment was merely press a key on a computer keyboard, when they 'felt' like it was the right time to do so. Rather than have every key press initiate a complex series of calculations (as had been the case for some of the previous research), Radin and Bosworth's study used an array of 5000 pregenerated equiprobable random numbers (0, 1, 2, 3, 4). A pointer would move through the array, taking 83.3 seconds to make a full pass. The participant's key press would stop the pointer as it passed through the array, where a hit was scored if the pointer landed on a 0, else the trial was a miss (the probability of a hit was $\frac{1}{5}$). The authors suggest this methodology is an improvement over others used in that rhythmic button pressing should not influence the scores of hits, as the array is not seeded from the time.

Radin and Bosworth (1985) used two special subjects, both of whom claimed some paranormal ability. Rather than simply test their participants' ability to score a predicted excess of hits, through mediation of their abilities to time correctly their button presses, Radin and Bosworth significantly complicated the analysis of their results, by comparing the raw data to four theoretical models. Model one proposed that psi is goal-oriented, model two proposed psi is goal-oriented but tempered by the nervous system and information processing, model three was based on Tart (1979) and the proposed theory of trans-temporal inhibition and predicted psi is akin to an edge-detection process and model four was based on a holographic model of reality.

What Radin and Bosworth (1985) actually found is very hard to say. Hoppe (1987) attempted to interpret Radin and Bosworth's transformations and equations, which they apply to test the different models. Hoppe published a detailed criticism on the statistics used by Radin and Bosworth, who he described, in a derogatory fashion, as 'psientists'. Hoppe stated that the Radin and Bosworth paper was given to him by a student, who believed it would diminish his long-held scepticism regarding the paranormal, but it in fact did the opposite and was full of, what Hoppe strongly describes as, 'logical and statistical errors' (p. 261). Radin and Bosworth proposed four different models of how the alleged psi-effect may manifest itself in the data and conducted 12 chi-square tests, where five were reported as

²Dossey (2008) reports that one of the motivations for much of Radin's research was the personal experience of a close friend. The friend was an avid gun enthusiast and on one occasion whilst cleaning his guns, felt an unusual sense of impending doom. Although he would often leave one chamber of the gun empty, he felt compelled to leave one further chamber empty. Dossey recounts that a few weeks later, the man's father-in-law attempted to shoot him during a drunken row, and had it not been for the fact that he had left the gun-chamber empty, a fact that is testified as being unusual, the man would have died.

significant, and thus concluded that their research tentatively supported the conclusion that their participants could somehow gain paranormal information about the state of the testing computer. Hoppe criticised Radin and Bosworth for allowing optional stopping, that is, that after the two participants had contributed over 10,000 trials, they could end the testing. Radin and Bosworth argued that the sheer number of trials that participants had to conduct made it unlikely that participants could know when they had an excess of hits, and thus when to stop contributing trials. Hoppe disagreed with Radin and Bosworth, and believed that it was a minor task for participants to mentally note any excess hitting across trials and thereby have an idea what their score was at any given time. Hoppe wrote a computer program, based on the procedure used in Radin and Bosworth's experiment and ran simulated trials on the computer. The program allowed for an approximation of what occurred in the Radin and Bosworth research, and Hoppe found that by keeping track of cumulative number of runs and of the total of any excess hits, the computer program achieved results, which by the standards outlined by Radin and Bosworth, would be evidence for ESP, despite only being a computer simulation.

Importantly, Hoppe (1987) also made a strong claim against Radin and Bosworth (1985), arguing that their statistics were miscalculated and misinterpreted in three ways. First they developed a chi-square equation, which was misunderstood by them and did not do what they imagined it was supposed to do. Specifically the chi-square was supposed to allow for different response patterns and was supposed to allow for testing against the four theoretical models. However Hoppe showed that 48 percent of the data are wrong, with the differences reported in the data being in the wrong direction about as often as they are in the right direction. He also argued that the authors had therefore based their conclusions on erroneous statistical tests and that the equation used by Radin and Bosworth is 'simply wrong' (p. 267). Out of a reported 12 statistical tests, Hoppe recalculated that only one was actually significant, after correcting for errors in the application and calculation of the statistics involved.

Hoppe (1987) also criticised Radin and Bosworth (1985) for dropping 10,000 trials from the data, because of a post-hoc assertion that the participant was attempting to 'psi-miss', that is, instead of trying to get an excess of hits, had been willing to get a dearth of hits. Hoppe argued that to ignore data on the basis of it not fitting in with the hypothesis being tested, smacks of blatant dishonesty. How can a researcher ignore parts of their data, based on a yet, unproven paranormal ability?

In the same journal Radin and Bosworth (1987) rebuked the attack on their research by Hoppe (1987) and there is some discussion between the researchers, in a further two papers, which outlined their differences. Radin and Bosworth ultimately had the last say in the argument, where their conclusion was that they were right and Hoppe is wrong. However Radin and Bosworth did agree with Hoppe, in that they could not be entirely confident that their significant psi-timing results were not an effect of optional stopping and all authors agreed that only with independent replication of the effect using a more controlled methodology could anything further be learnt.

Braud and Shafer (1989), though published after Radin and Bosworth (1985), was actually conducted before. The basic aim of Braud and Shafer's research was to show that participants could use ESP to score a hit, through the timing of a key press response. They developed a computer-based psi task, called PSI MACHINE, which required minimal cognitive processing on the part of the participant. Like Radin and Bosworth's experiment, all the participant had to do, was simply press a computer key when he or she felt that 'the time was right'. If the

participant chose the correct time, they were presented with an audio sound of a bell, if they were wrong they were presented with a buzzer tone, signalling an incorrect response. Braud and Shafer described the experiment as being analogous to automatic writing and therefore to some degree it can be seen as an unconscious test of psi requiring very little conscious input from the participant. Their testing took well over a year and unlike Radin and Bosworth's experiments, made use of many more participants. Importantly, Braud and Shafer employed a number of differences in their methodology, making it a similar, but not identical, experiment. The main difference was the actual task. Where Radin and Bosworth had employed a 'one-step' methodology, with the key press stopping a pointer in a large array of random integers, Braud and Shafer employed a multi-step methodology. A key press, initiated selecting the clock time value as a seed. The seed was used to generate a pseudo-random number between 1 and 6. That integer was the target value. The next key press read the time again, and another integer was obtained from the pseudo-random generator, this was the guess integer. If the guess integer and target integer matched, the trial was scored as a hit. Where the probability of a hit for the Radin and Bosworth research was $\frac{1}{5}$, Braud and Shafer's study employed a probability of a hit of $\frac{1}{6}$. Some 51 participants were tested for a total of 3600 trials, consisting of 100 36-trial runs. Braud and Shafer had prespecified 100 as the total number of trial runs that they were going to collect. Participants could contribute more than once, with 30 providing just one run and twenty-one participating in more than one run.

Braud and Shafer (1989) reported an observed hit rate of 18 percent ($p = .02$, one-tailed), arguing that their participants had scored significantly above the chance expectation of 17 percent and was evidence for a paranormal ability. Along with their experimental data, Braud and Shafer reported three control test runs that were conducted with simulated key presses, where the computer program was adapted to randomly simulate a participant pressing at both times, therefore allowing the randomness of the software to be evaluated. None of the hit rates from the control runs reached significance, and so Braud and Shafer felt confident in concluding that their software was unbiased and the excess of hits from participants was a consequence of a psi-timing effect. Whilst the statistics employed by Braud and Shafer are a lot clearer than those used by Radin and Bosworth (1985), there is still some concern as to the importance of response patterns affecting the data. Radin and Bosworth raised the concern that if participants timed their responses in a way which somehow matched the internal clock, that the participant could potentially implicitly learn a normal means for producing increased hits, in the presence of feedback, although it is hard to evaluate the degree to which any particular response bias might have affected the results of Braud and Shafer's research.

Despite the concerns surrounding Radin and Bosworth (1985), the psi-timing effect was chosen as the next empirical regime to test and a significant paranormal psi-timing effect may well fit neatly into a larger psi-as-death-avoidance framework. However, it is also telling that despite the increasing ease with which this type of experiment can be conducted and data collected, today such an experiment could be run over the internet with minimal technical difficulties and could collect huge amounts of data very quickly, that very few parapsychological researchers have attempted to investigate the effect. As has been seen in previous chapters, Dean Radin has moved on from investigating the psi-timing effect and it gets no mention in his book, *The Conscious Universe* (Radin, 1997a). Are the strong claims, followed by a silence on the part of researchers to replicate and extend the area, an indication that no further research was successful and the paradigm left, in favour of new and more exciting paradigms, like the presentiment effect? It is certainly suspicious that if no effect of psi-timing is found, that this

area will look like the others already investigated: where initial enthusiasm for an effect has been replaced by an inability to replicate and finally stagnation, as researchers flock to pastures new.

As one of the aims of this thesis is to develop and test the usefulness of using evolutionary fear-based stimuli within parapsychological research, it was deemed appropriate to adapt the psi-timing methodology, and make improvements based on what had been incorporated into the previous work with the presentiment effect and the time-reversed cognitive interference. Specifically, two major changes were incorporated into the design of Psi-Timing I, which it was hoped might stimulate the alleged paranormal above-chance scoring. The first major difference would be to take part of the procedure as used by Hoppe (1987), who had employed two simulated key presses to investigate the effect of optional stopping, and use one simulated key press to generate the target in the participant's future. The procedure used would follow this order: first the participant presses a key, when they feel it is right to do so, thus generating the guess integer and this is followed by a simulated key press, which generates the target integer. A comparison of the two integers determines if the trial is a hit or a miss. With this procedural change, the methodology was adapted from a test of contemporaneous ESP, to a test of the precognitive hypothesis. This change was initiated because by incorporating a minor adjustment, the experiment would fit very well with the work undertaken in the previous chapters. Interestingly, there is also some precedent for investigating a precognitive psi-timing effect in the literature.

Vassy (1986) conducted a five-series investigation into precognitive psi-timing. Vassy's psi-timing experiments involved a goal (i.e. for the participant to try to get an excess of zeros) and was an attempt to investigate the effect of complexity of the task on performance. Despite Vassy's overall conclusions that there is evidence for precognition in his data, he fails to highlight that the overall hit rate for none of the studies differs significantly from chance and that the overall hit rate combined is exactly that expected by chance (50%). Vassy (1992) followed up these studies with a further investigation into precognitive psi-timing and reported an overall significant hit rate ($p < .01$) seemingly showing a modest ability for participants to adjust their timing in response to a future event.

What has been conducted in regards to precognitive psi-timing or related phenomena is very mixed and none of the evidence is very strong regarding the proposed effect. However it was hoped that a second adaptation to the methodology might improve the chances of obtaining an above-chance excess of hits. Specifically, Psi-Timing I incorporated the same changes seen in preceding chapters, in that stimuli previously used were replaced with spider-based stimuli. Where

4.1.2 Method

4.1.2.1 Design

Psi-Timing I employed a between-subjects design, where ratings on the FSQ would be used to compare the fear group with the no-fear group. The independent variable was the number of hits scored. A hit was scored if the integer chosen as a consequence of the first keyboard press (the guess) matched the target integer. The total score would be compared to mean chance expectation, and a significant positive deviation from chance would be interpreted as likely to be the result of a precognitive influence.

4.1.2.2 Participants

Fifty participants volunteered (40 females, 10 males) with a mean age of 24 years, a standard deviation of 7.8 years, with a minimum of 17 and a maximum of 50. Twenty-four participants rated themselves as not afraid of spiders and twenty-six reported fear.

4.1.2.3 Materials

The Fear of Spiders Questionnaire (Szymanski & O' Donohue, 1995) was used to discriminate between the fear and no-fear group, and consists of six questions, to which an answer in the affirmative to any or all, designates the participant as spider-fearful. A desktop PC was used to run the experimental program, which made use of standard pseudo-random algorithms seeded from the internal clock. The software was constructed by the author in a psychology experiment development package, *E-Prime*. It was thoroughly tested before use to ensure random presentation of stimuli and accuracy of recording. The stimuli were 12 pictures, half of which were neutral and the other half were pictures of spiders.

4.1.2.4 Procedure

First the participant was recruited and was informed that they were taking part in a short and simple ESP task. Once their age and gender were recorded, participants were given brief instructions as to what they were expected to do. Participants then pressed the space-bar to begin the trials. On the computer screen the participants were presented with text requesting them to 'press the space-bar when you feel it is the right time to do so'. The participant could take as long as they liked, although none took longer than a few minutes for each trial. Once the participant had pressed the space-bar, the program read the internal clock and used it to produce a random number between 1 and 6.

With the participant's guess generated, the software randomly waited for a brief period (between 1–6 seconds) and then a simulated key press was initiated, where the computer clock was used to seed the pseudo-random algorithms and generate a target integer. If the participant's 'guess' and the target integers matched, the participant was presented with a neutral image. If the two integers differed, the participant was presented with a picture of a spider. Instead of presenting the same neutral and spider images on each occasion, a random selection was made between one of 6 different neutral or spider images, meaning that there was a $\frac{1}{6}$ chance of any particular spider or neutral image being presented. This was done to avoid fast habituation to the negative images, and to provide some degree of novelty.

After completion of all 36 trials, participants filled out the Fear of Spiders Questionnaire and were debriefed about the aims of the study. Participants were also given a score out of 36 as an indication of their individual performance in the task.³

4.1.3 Results

Fifty participants contributed 1800 trials, consisting of 36 trials per participant. The combined overall hit rate for the study was 16.1 percent, where mean chance expectation is 16.7 percent. (The probability of a hit was $\frac{1}{6}$, the number of trials was 36 and there were 50 participants.)

Twenty-six participants rated themselves as being afraid of spiders, with the remaining 24

³This score was given as a measure of their psi performance and was part of the incentive to participate. As the specific emphasis of this investigation was in the grouped scores, no attempt was made to maintain contact with any high scorers (otherwise known as special subjects).

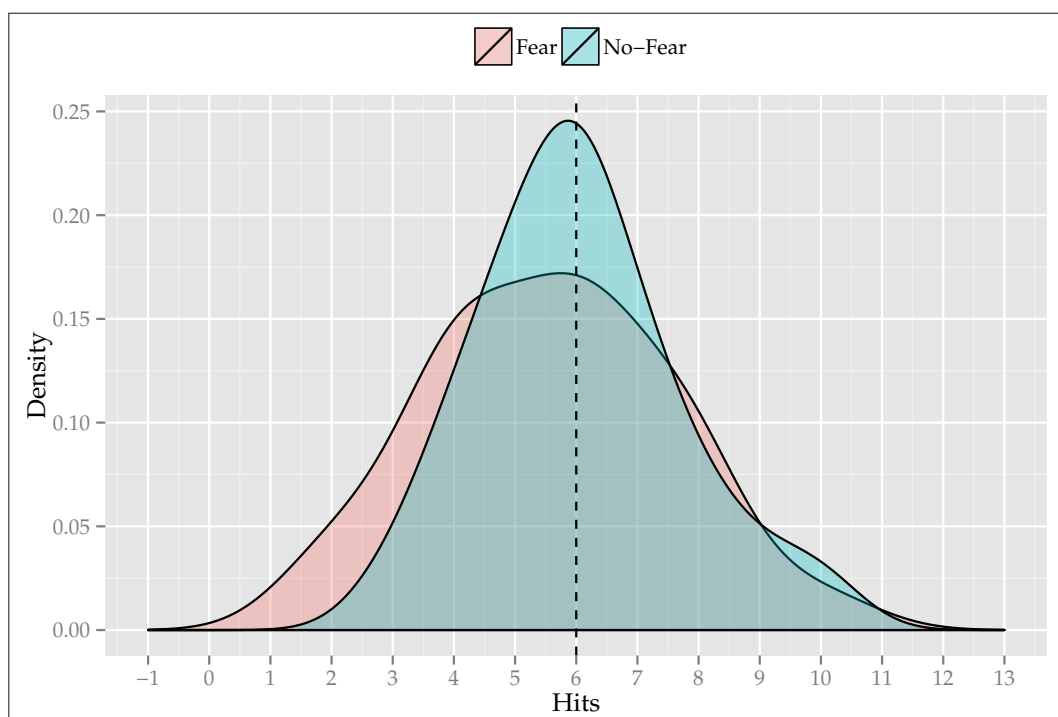


Figure 4.1: Combined density plot of hits for Psi-Timing I, with separate lines showing the hit rate across the spider-fear and no-fear groups. MCE is marked with the vertical dashed line. The slightly taller curve of the no-fear group signifies less variation in the responses and vice versa (though this difference is not statistically significant).

rated as no-fear. The null hypothesis predicted that there would be no significant difference in scoring between the two fear groups, whereas the paranormal hypothesis predicted that there would be a significant difference in scoring between the fear and no-fear group.

The total number of hits obtained in the spider-fear group was 145, and for the no-fear group was also 145. The slight difference in group size results in a slight difference in hit rate: the fear group have underscored from MCE with a deviation of -11 , and the no-fear group have scored above MCE by just $+1$. From the hit rates alone it is evident that there has been no statistically significant above-chance scoring in the individual groups. The combined score was below chance and the spider-fear group have also scored below chance. Although it is a trivial matter to calculate the z-test score, the BSDA package of R (Arnholt, 2010) provides an independent two-sample z-test (comparing the two fear groups): $z = .73, p = .77$ (one-tailed). The spider-fear group did not score significantly above chance: $z = .96, p = .83$ (one-tailed). Figure 4.1 shows a combined density plot⁴ for the fear and no-fear groups, and shows little difference between the scoring patterns in either spider-fear group.

Finally, some consideration must be given to the power of this particular study. Although issues regarding the power or lack thereof of this particular study were considered during the planning stage of the experiment, it is a topic especially difficult when applied to parapsychological research. Specifically, if successful replication were guaranteed merely by increasing the power of a study, unequivocal evidence for the paranormal would already be in existence. Unfortunately this is not the case and no single element can guarantee experimental success, let alone a veridical paranormal finding. As this is the first psi-timing study to investigate a

⁴The density plot was chosen over the more familiar histogram, to allow for a comparison of the distributions of the results in this chapter (see Chambers, Cleveland, Kleiner, & Tukey, 1983).

difference across fear groups it is not possible to directly predict the effect size that is to be expected (given the hypothesis to be tested). The effect size given in Braud and Shafer's study is $d = .2$ (a small effect). If we assume that in Psi-Timing I the no-fear group should score at chance (16.7% hit rate) and the fear group should score above-chance (estimate a hit rate of 20%), then a power calculation allows us to estimate a) how powerful the current study is and b) whether our number of participants is appropriate given the effect size.

We can use the PASS software to calculate the power of a 2-sample z-test (Hintze, 2013), which given the estimated difference in means gives a power rating of approximately .6 for the current study. The number of participants in each group to reach .80 for this estimated effect size is 86. Given the estimated parameters then, this study is slightly underpowered, if the size of the effect is as large as estimated.

4.1.3.1 Control Trials

Two control runs were conducted to provide a test of the computer software. These control runs consisted of 1000 simulated key presses, where the probability of a hit was $\frac{1}{6}$. The key presses were simulated by replacing the pause for participant input with a random wait period (of a few seconds duration) and was intended to test the pseudo-random number generator for bias, and to ensure that the programme was fit for purpose, with no unanticipated bugs etc. The statistical analyses to be presented are repeated later in the chapter with obtained data and provide a comparison.

One block of 1000 simulated key presses was run before data collection for Psi-Timing I had begun. The second block of 1000 simulated key presses was run after data collection for Psi-Timing I. The hit rate for Control Run I was 17.4 percent, which a binomial test (using the `binom.test` function of R) shows is not significant: $p = .52$ (two-tailed). The hit rate for Control Run II was 15.1 percent, which was also not significant with the same binomial test: $p = .19$ (two-tailed).

To allow comparison to the remaining experiments of this chapter (not including Psi-Timing I), the control trials were divided between 40 pseudo-participants and a sum calculated across 25 trials. The probability of a hit remained $\frac{1}{6}$. Figure 4.2 shows a combined density plot for both Control Run I and II. A one-sample z-test (using the BSDA package of R, Arnholt, 2010) gave: Control Run I: $z = .62$, $p = \text{ns}$ (two-tailed, with a mean number of hits of 4.35); Control Run II: $z = -1.33$, $p = \text{ns}$ (two-tailed, with a mean number of hits of 3.78). Thus, there is no evidence of extra-chance scoring within the pseudo-participants of the control runs. In Control Run I, the maximum score was 7 and the maximum score of Control Run II was 8. Neither result is significant based upon the Bonferroni-adjusted probability.

Finally the control data were examined for two purported effects which are found with some degree of consistency within parapsychology: the displacement effect and the periodicity effect. The displacement effect was proposed after a number of early researchers (e.g. Rhine, 1934) discovered that their participants were scoring hits, not upon the target they were 'supposed' to be attending to, but rather ahead by one or two places, or behind by one or two places. It should also be noted that in Soal's famous work (Soal & Goldney, 1943), initial results were non-significant and it was not until the displacement hits were examined that any significance was discovered. Figure 4.3 shows the displacement hits plot for both control runs. The plots can be thought of as showing what the appropriate single density plot would have looked like, had that particular target and response occurred, or had we been interested in that particular combination, rather than the pair that we are specifically interested in (the 1:1 hits).

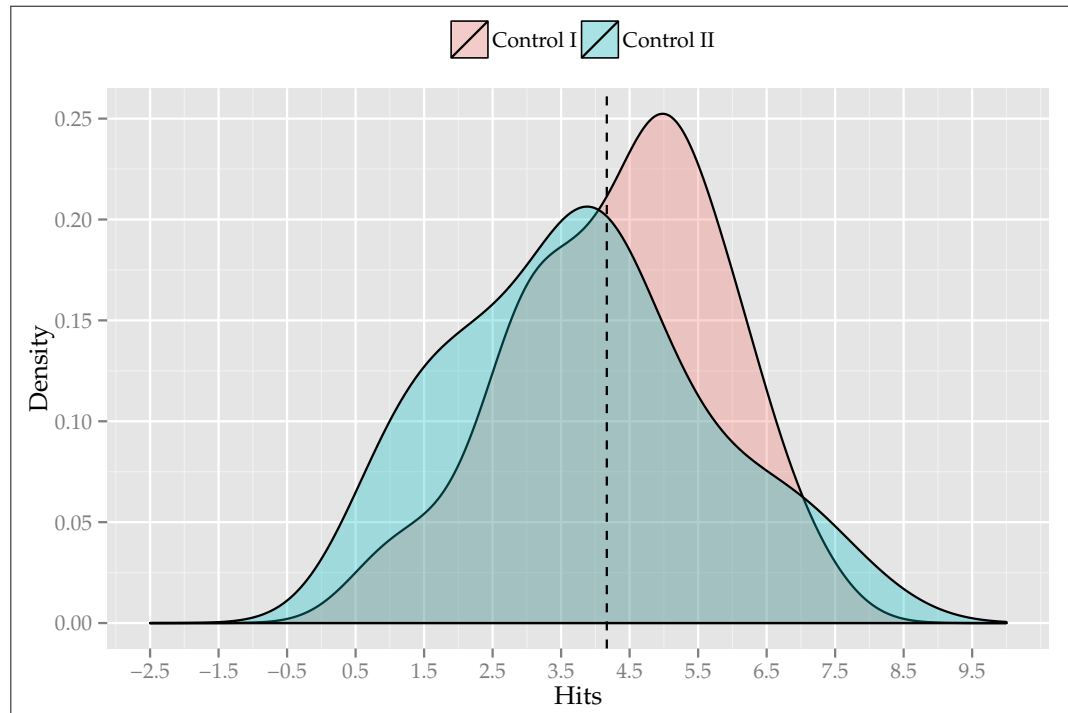


Figure 4.2: Combined density plot of hits for the control trials, with separate lines showing the hit rate across Control Run I and II. MCE is shown via the vertical dashed line.

Although all the lines trend to the normal distribution, there are certain eccentricities within the plots, though these are not interpreted as anything but the consequence of minor chance variation.

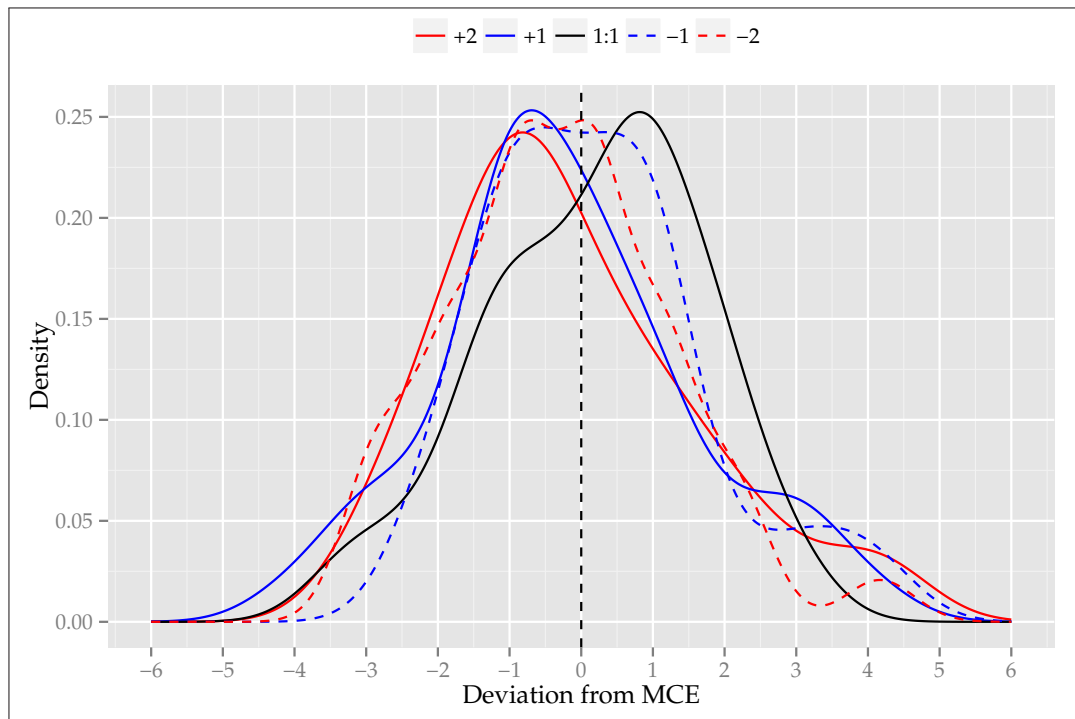
The second analysis is based on details given in Hansel (1966), and is an examination of the periodicity effect on the 1:1 hits. The periodicity effect is very similar to the displacement effect, where it is purported that hits form a U-shape, with high scoring at the beginning and end of a series of trials. Some 40 years ago, Rhine (1969, p. 136) wrote that, “declines and U-curves. . . are probably the most lawful indications of psi performance thus far available.” Figure 4.4 shows the appropriate periodicity plots. It can be seen that there is no apparent trend to a U-shaped curve, as averaged across the pseudo-participants.

Finally, it is important to note that all three psi-timing experiments used an identical software core.⁵ As such the results of these control trials apply equally to the other psi-timing experiments, in so far as they are an indication of the reliability of the software.

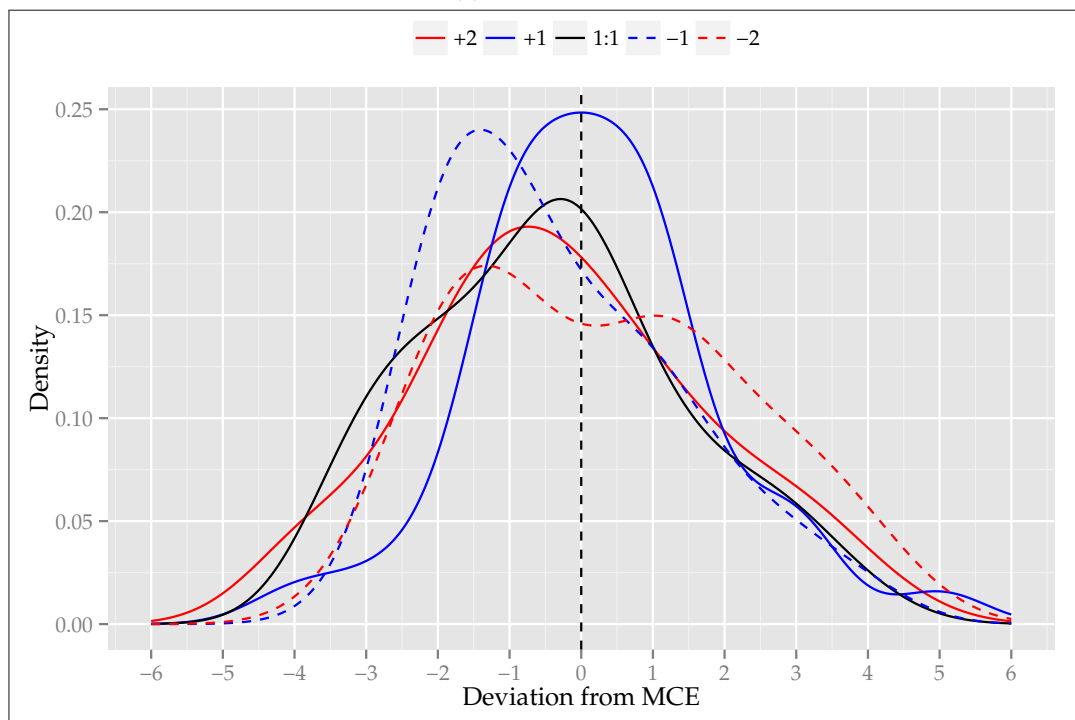
4.1.4 Discussion

The results of Psi-Timing I are clear. The null hypothesis cannot be rejected and, in this study, spider-fear did not influence responses in the precognitive task. The spider-fear group have not scored significantly above chance nor is there a significant difference between the no-fear and fear group. The use of spider-based stimuli has proved no more successful in the psi-timing paradigm than when it had been used previously in the presentiment paradigm and the TRI paradigm. It has also been shown to be no more effective at providing evidence for paranormal

⁵This same software was used in the development of the TRI research and hence the control trials can also be applied to that research.

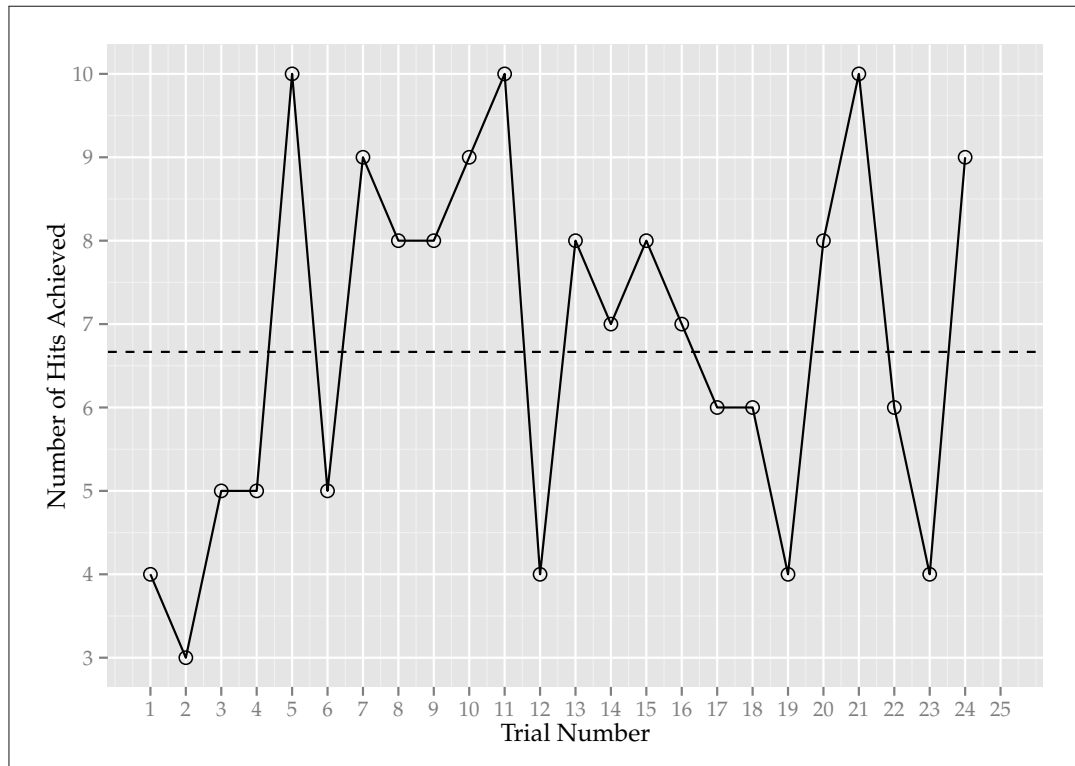


(a) Control Run I

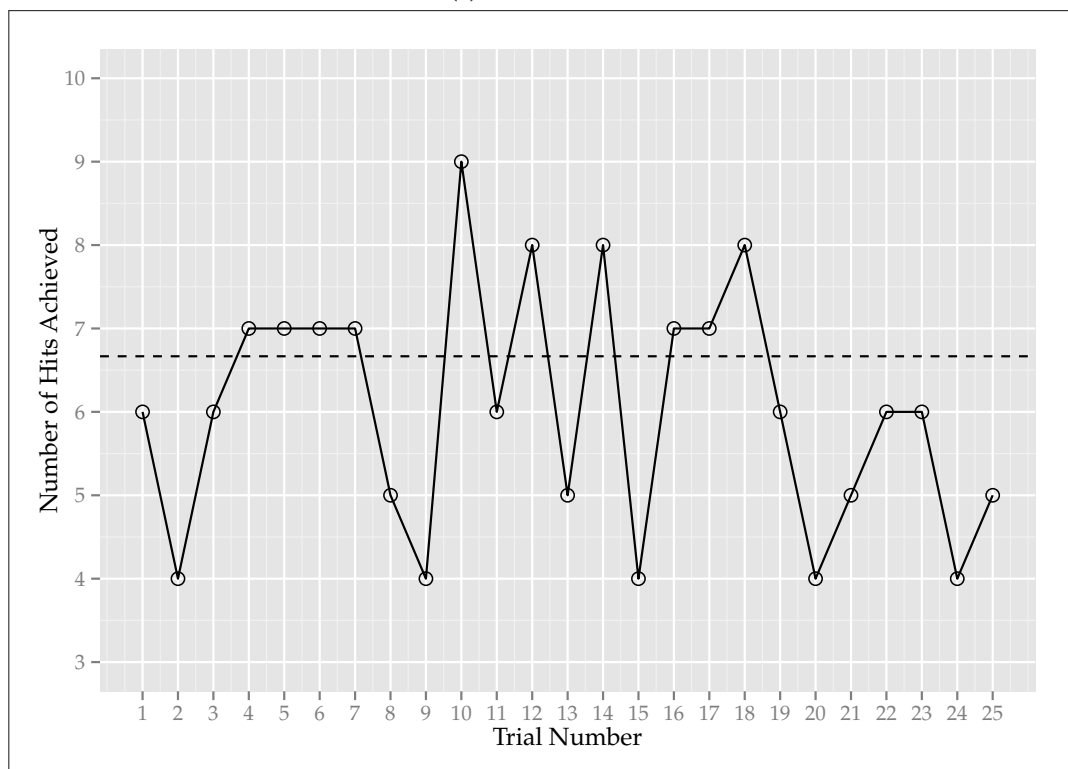


(b) Control Run II

Figure 4.3: Two combined density plots showing deviation from the mean across displacement targets (identified in legend). **(a)** shows the data from Control Run I, and **(b)** shows the results for Control Run II. Due to the differing hit rates (e.g. the +2 deviation has 23 trials) the results are standardised to allow direct comparison. Looking for displacement hits was historically common earlier in parapsychology (for example see Hansel, 1966) and involves investigating that significant scoring has not been displaced to other targets in the testing sequence.



(a) Control Run I



(b) Control Run II

Figure 4.4: Plots to investigate potential periodicity effects. (a) shows the results from Control Run I, and (b) Control Run II. The plot shows the number of hits achieved on each trial number, with the dashed horizontal line denoting chance expectation. Because of the way the data have been constructed (i.e. 2000 trials distributed across 80 pseudo-participants assigned 25 trials where the probability of a hit is $\frac{1}{6}$), no value can have the arithmetic mean value. It is worth noting that this data is an average across pseudo-participants and may neglect individual effects.

phenomena than any of the other stimuli thus employed.⁶ Given the estimated effect size, the study was estimated to be slightly underpowered. But given the procedure it was hoped that a significant difference would be teased out between the two fear groups. The lack of significant difference is thus disappointing.

There can be one confident conclusion. From the control trials and experimental data, there can be confidence that the experimental protocol is free of any major artefacts which might be misconstrued as a precognitive influence.

4.2 Psi-Timing II

4.2.1 Introduction

Because of the absence of a precognitive effect in any of the completed research hitherto discussed, it was felt that a 'back to basics' stance should be taken. It seems hard to continue to research parapsychological phenomena in the absence of those phenomena. Thus, rather than conduct a second replication of Psi-Timing I, Psi-Timing II was to be closer to Braud and Shafer (1989), in an attempt to establish *any* finding that required recourse to a paranormal explanation. Specifically then, Psi-Timing II was to remove the precognitive element from testing and replace it with the original contemporaneous ESP test that Braud and Shafer had employed. It would also eliminate the spider-fear element that had not provided the success that was hoped for, and made use of Braud and Shafer's bell and buzzer methodology.

Psi-Timing II was therefore an attempt to replicate as closely as possible the work of Braud and Shafer (1989). Participants would be asked to press a keyboard button, when they 'felt' the time was right and, if the previous claims for the area are correct, they should be able to time their responses to get an excess of hits overall. The main difference between this experiment and Braud and Shafer's research was a reduction of the probability of a hit from $\frac{1}{6}$ to $\frac{1}{5}$. This was effected only to bring the hit rate in line with the gold-standard of parapsychology (generally speaking 25 trials, with a probability of a hit of $\frac{1}{5}$). This minor difference in scoring was not anticipated to have any significant effect on a real paranormal effect, and Radin and Bosworth (1985) had used the same hit rate in their own research.

4.2.2 Method

4.2.2.1 Design

Psi-Timing II employed a very simple within-subjects design. The independent variable was the number of hits scored. A hit was scored when the first and second key presses resulted in matching integers and a miss was scored when the two integers were different. This independent variable would be compared to mean chance expectation, and any above-chance scoring would be seen as indicating a possible paranormal effect.

⁶This is the last experiment testing the use of spider stimuli as targets in precognitive research conducted as part of this thesis, though I have been involved in further testing using spider stimuli (see Savva et al., 2005; Savva, Roe, & Smith, 2006), although with the same lack of success.

4.2.2.2 Participants

Some 30 participants volunteered (23 females and 7 males). Participants had a mean age of 35.5 years (SD 13.1 years, with a minimum of 14 and a maximum of 69) and were mostly visitors to a psychology open-day.⁷

4.2.2.3 Materials

A desktop PC was used to run the experimental program, which made use of standard pseudo-random algorithms seeded from the internal clock. The software was constructed by the author, in a psychology experiment development package, *E-Prime*. It was thoroughly tested before use to ensure random presentation and accuracy of recording. Two audio sound files were used to indicate success or failure: one a buzzer sound and the other a bell. Each lasted approximately 1 second.

4.2.2.4 Procedure

First the participant was recruited and informed that they were taking part in a short and simple ESP task. After age and gender were recorded, participants were given brief instructions as to what they were expected to do.⁸ Participants then pressed the space-bar to begin the trials. On the computer screen the participants were presented with text requesting them to 'press the space-bar when you feel it is the right time to do so'. The participant could take as long as they liked, although none took longer than a few minutes for each trial. Once the participant had pressed the space-bar, the program read the internal clock and used it to produce a random number between 1 and 5. The screen was then cleared and replaced with an alternatively coloured screen with the instructions 'now press the space-bar, again, when you feel it is the right time to do so'. Again participants could take as long as they wanted, with most acting within a period of seconds. After the second key press, the time was again used as a seed, to produce a random number between 1 and 5. The first key press number was then compared to the second key press number: if they were the same, the trial was scored a hit and a bell sound played, and if they were different, the trial was a miss, and a buzzer sound played. Each participant contributed 25 trials, at the end of which they were given a score and debriefed.

4.2.3 Results

Some 30 participants contributed 750 trials, consisting of 25 trials per participant. The overall hit rate in this study was 23.2 percent, 3.2 percent above MCE. To test whether this above-chance scoring is statistically significant we use the very simple and widely employed one-sample z-test (otherwise known as a critical ratio). This is simply calculated as the observed deviation from MCE, divided by the standard deviation. The BSDA package (Arnholt, 2010) in R provides a one-sample z-test: $z = 2.19$, $p = .01$ (one-tailed). An effect size (h) can be approximated by the formula⁹ $h = \frac{z}{\sqrt{n}}$ which gives .40 (a low to mid figure).

⁷The experiment was advertised with a sign asking potential participants to 'Test their psychic powers', and was part of an effort to encourage a conducive environment given a rather busy testing location. At the end of the experiment participants were given a score as an indication of their 'psychic powers'.

⁸The instructions were: When you successfully score a hit you will hear a bell and when you miss you hear a buzzer. However don't think too hard, let your fingers do the guessing. You can take as little or as much time as you want.

⁹From Braud and Shafer (1989).

Further to the analyses conducted as part of the investigation into the control trials, two post-hoc and informal analyses were performed. The first was based on the historic effect of the displacement effect which, to reiterate, was proposed after a number of early researchers discovered that their participants were scoring hits, not upon the target they were 'supposed' to be attending to, but rather ahead by one or two places, or behind by one or two places. Figure 4.5a shows the density plot for the planned hits, and includes the +2, +1, and -1, -2 displacement hits. Thus, the plot can be thought of as showing what the single density plot would have looked like, had that particular target and response occurred, or had we been interested in that particular combination, rather than the pair that we are specifically interested in (the 1:1 hits). It is interesting that the 1:1 hits is the only line which shows some hint of unusual scoring.

The second post-hoc analysis is again based on details given in Hansel (1966), and is an examination of the periodicity effect on the 1:1 hits. The periodicity effect is very similar to the displacement effect, where it is purported that hits form a U-shape, with high scoring at the beginning and end. Inspection of Figure 4.5b shows no periodicity, and although there does appear to be slightly higher than average scoring across the study (that is, on each trial participants are scoring above the mean more than perhaps would be expected by chance), on closer examination there are in fact 12 trials that have above-chance scores and 13 which have scores at or below chance. It is again worth noting that this analysis is averaged across all participants and thus neglects individual periodicity.

During data analysis, an external statistical observer (AP) noted that four participants have scores that exceed the .05 α level, where the expected number for 30 independent participants would be 1.5, and that one participant, who scored 13 hits¹⁰ has a two-tailed binomial probability (binom.test function of R) of $p = .00037$, which is below even a Bonferroni-corrected two-tailed p-value (.05/30). It may also be argued that this individual shows a degree of periodicity in their scoring pattern. It should however be noted that these analyses are post-hoc and the planned level of analysis was the grouped, not individual scores. It is clear that from the perspective of the planned analysis, Psi-Timing II *has* provided some marginal support for the previously published claim of psi-timing.

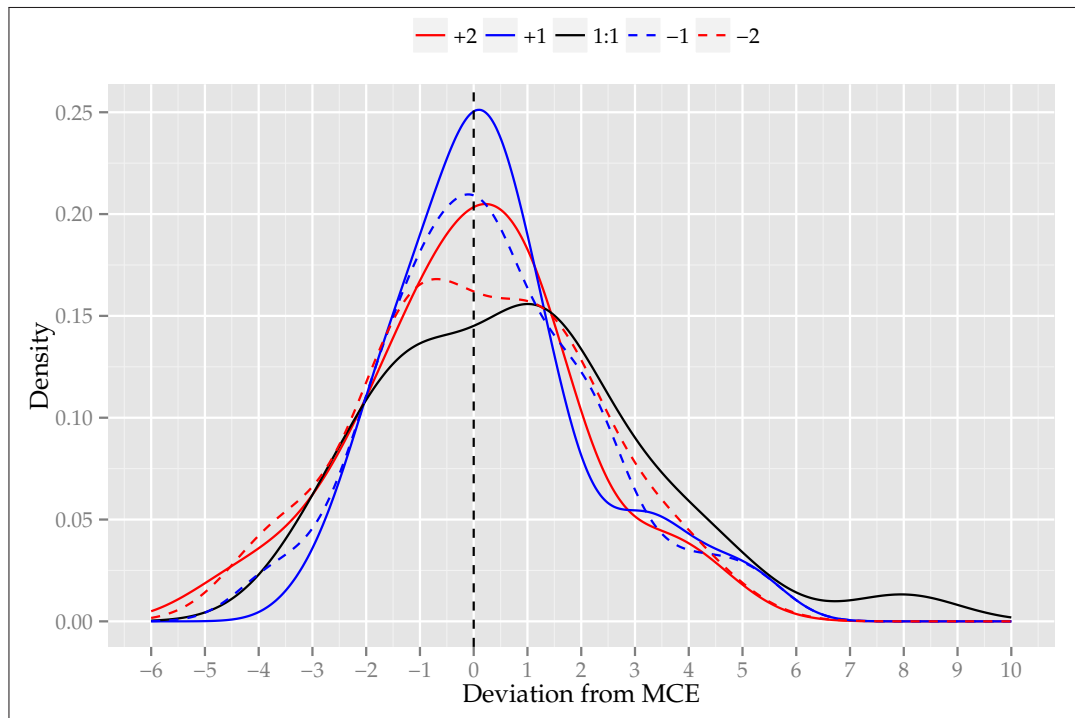
Finally issues of power must again be considered. As an almost direct replication of Braud and Shafer (1989), we can use their found effect size to calculate the power of the current study ($d = .2$, a small effect). The pwr package of R (Champely, 2012) provides a power test based on the z-test normal distribution (pwr.norm.test) which gives a power rating of approximately .29 for the current study. This is a low power. The same package gives a value of 155 participants needed to raise the power of the study to .80. Had a power calculation been done prior to data collection this would have been an achievable number of participants to recruit and it is regrettable that such a small sample size was recruited.

4.2.4 Discussion

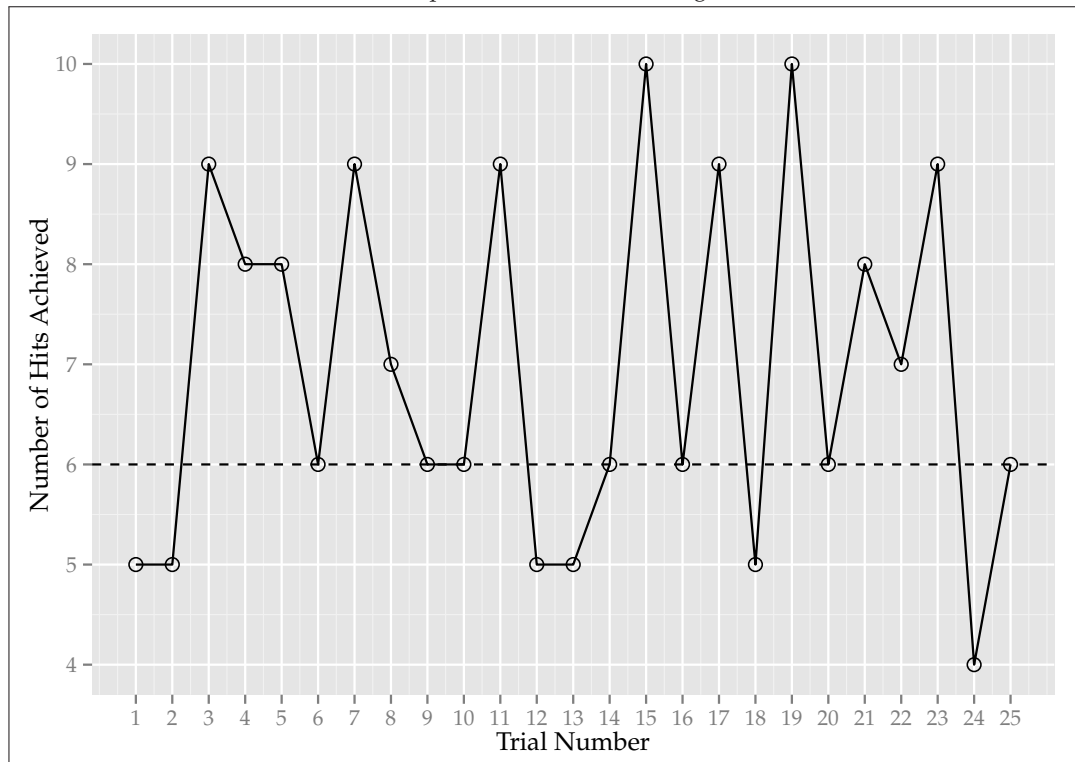
Having employed a very similar methodology to Braud and Shafer (1989), a small psi-timing effect was found. Although not particularly large, this effect appears to be the consequence of a

¹⁰As one of the only potentially paranormal results obtained during this empirical regime, the guess and target sequence is presented:

GUESS:	3	4	1	5	2	2	3	3	4	1	5	5	5	4	1	1	3	4	1	3	5	5	4	5	5
TARGET:	2	1	1	5	2	2	3	3	2	1	4	1	1	5	1	5	3	4	2	1	5	5	4	2	2



(a) Displacement hits Psi-Timing II



(b) Periodicity plot Psi-Timing II

Figure 4.5: (a) Combined density plot showing displacement from chance expectation for Psi-Timing II (b) Plot to investigate potential periodicity effects for Psi-Timing II.

small number of participants scoring above chance, and as was noted in the results section, one participant did score quite significantly above mean chance expectation. Presumably though, not all participants were able to adjust their responses to achieve an excess of hits in the study. However, given that we have concluded that this study has low power, the response to this single result must be muted. Although in the predicted direction and significant, it must also be compared to the other research within this thesis and as such demands to be replicated before it can be taken with any great seriousness.

4.3 Psi-Timing III

4.3.1 Introduction

The suggestively significant finding of Psi-Timing II necessitated replication before any serious consideration of the result. Because no effort was made to enable any high scoring participants to be contacted and retested, it was not possible to subject the high-scoring participant from Psi-Timing II to further testing. Although in retrospect this may appear to be an oversight and the loss of an important opportunity, this testing regime was never intended to rely upon special subjects. As such, Psi-Timing III remained identical in methodological procedure (see page 129), whilst increasing the sample size somewhat. It was hoped that an increase in participant numbers would raise the power of the study and thus, if it replicated the significant finding of Psi-Timing II, would provide good evidence that the result was veridical not artefactual.

4.3.2 Method

4.3.2.1 Participants

Fifty participants volunteered (17 females, 33 males). Aged from 16 years to 48 (mean 32 years, SD 7.4 years). All participants were visitors to the Fortean Times Unconvention, a convention organised by the *Fortean Times*, which is a magazine dealing with the paranormal and supernatural.

4.3.3 Results

Fifty participants contributed 1250 trials, consisting of 25 trials per participant. The overall hit rate was 18.5 percent. This represents an underscoring of hits of 1.5 percent below that expected by chance and as such no further statistical analysis was deemed warranted. Some 2.5 participants would be expected to score above 9 by chance, and only 1 participant does so (with a maximum score of 10). Figure 4.6 shows a combined density plot for both Psi-Timing II and III.

Again, for comparison, that is, to allow us to determine whether or not the suggestively significant finding of Psi-Timing II is very different from chance, Figure 4.7a shows the displacement hits for Psi-Timing III and Figure 4.7b shows the plot of periodicity (it again should be noted that this is an average across all participants and neglects individual scores). Again the results seem to show chance with some minor variation.

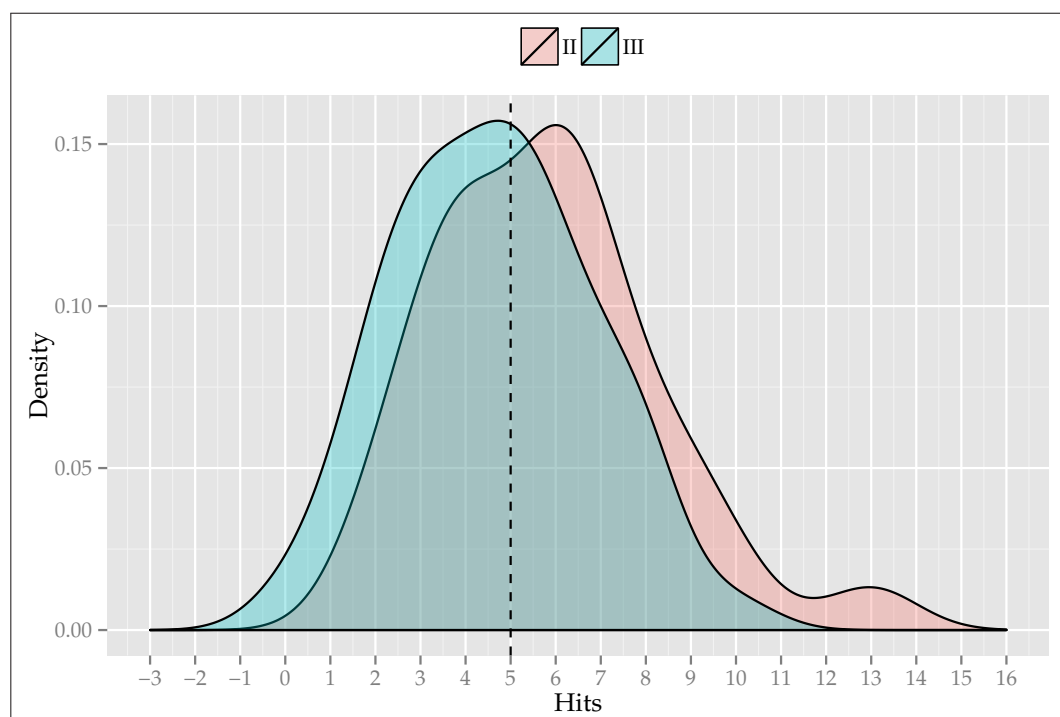


Figure 4.6: Combined density plot of hits for Psi-Timing II & III. MCE is marked with a dashed vertical line.

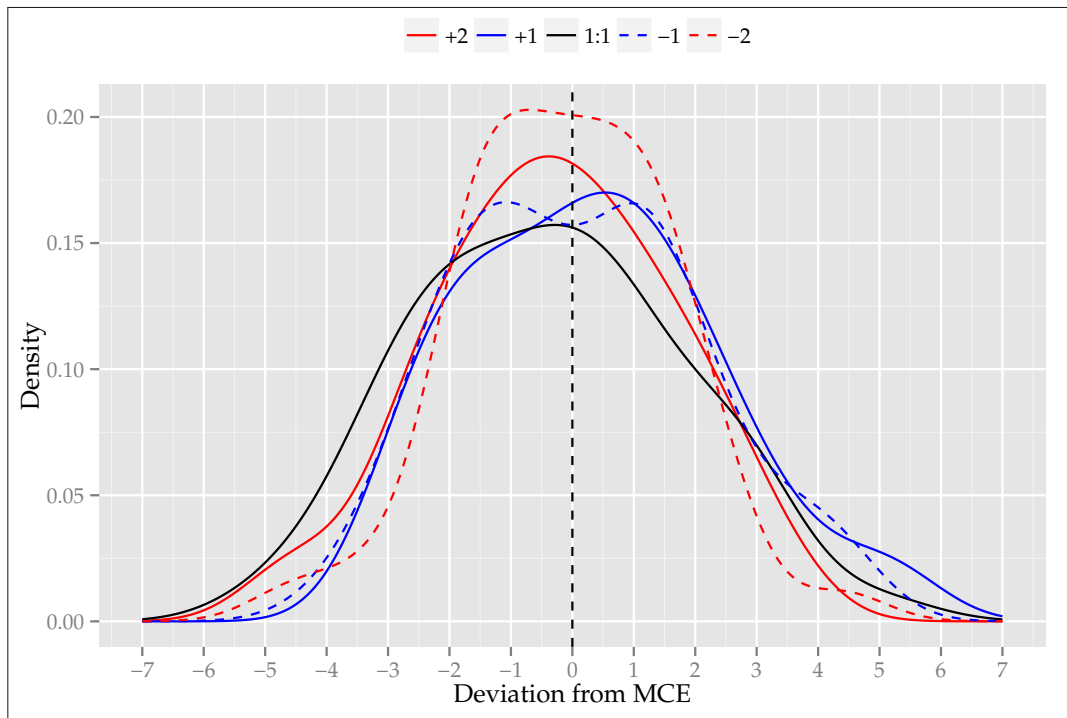
Again issues regarding the power of the study must be raised. The *pwr* package of R (Champely, 2012) provides a power test based on the z-test normal distribution (*pwr.norm.test*) which gives a power rating of approximately .40 for the current study (not as small as for Psi-Timing II, but still on the low side). As with Psi-Timing II, 155 participants would have raised the power of the study to the conventional level of .80 and again it is regrettable that this number of participants was not recruited given the ease with which data was collected.

Finally, related to the issue of power and whether the Psi-Timing II result was a false-positive, Figure 4.8 shows a confidence interval plot derived from z-tests for all the psi-timing results presented above (including the displacement hits). Again, this plot allows a general comparison of the suggestive result found in Psi-Timing II, that is, whether or not the result is remarkable enough to warrant a paranormal explanation.

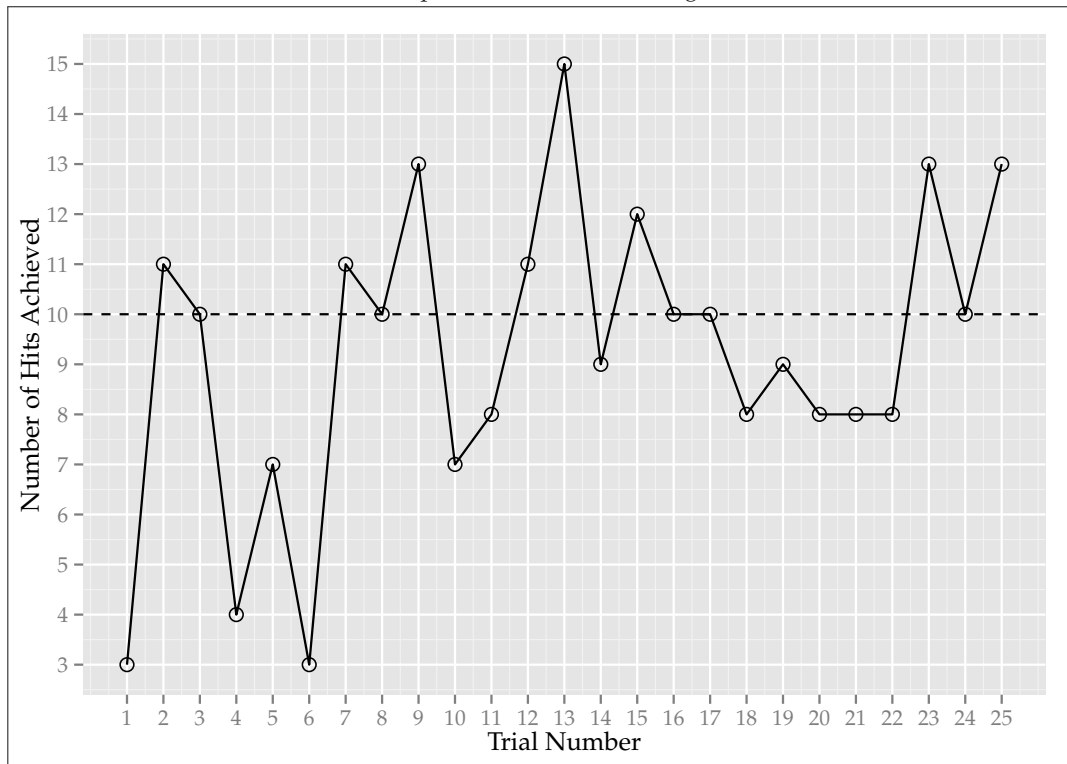
4.3.4 Discussion

The results of Psi-Timing III failed to replicate the suggestively significant p-value of Psi-Timing II. However given the fact that both studies are severely under-powered, the over-riding conclusion from both studies must be that they are “scientifically useless” and thus all we can safely conclude is that no strong psi-timing effect is in evidence. The statistically significant result obtained in Psi-Timing II is anomalous (though the low power of that study must prevent the leap to the conclusion that it was a consequence of paranormal functioning), but only a regime of testing employing much larger numbers of participants can effectively test a paranormal explanation.

Overall there is very little evidence for any kind of psi-timing effect (precognitive or otherwise). Despite the strong claims by the earlier researchers, the three experiments described



(a) Displacement hits Psi-Timing III



(b) Periodicity plot Psi-Timing III

Figure 4.7: (a) Combined density plot showing displacement from chance expectation for Psi-Timing III (b) Plot to investigate potential periodicity effects for Psi-Timing III.

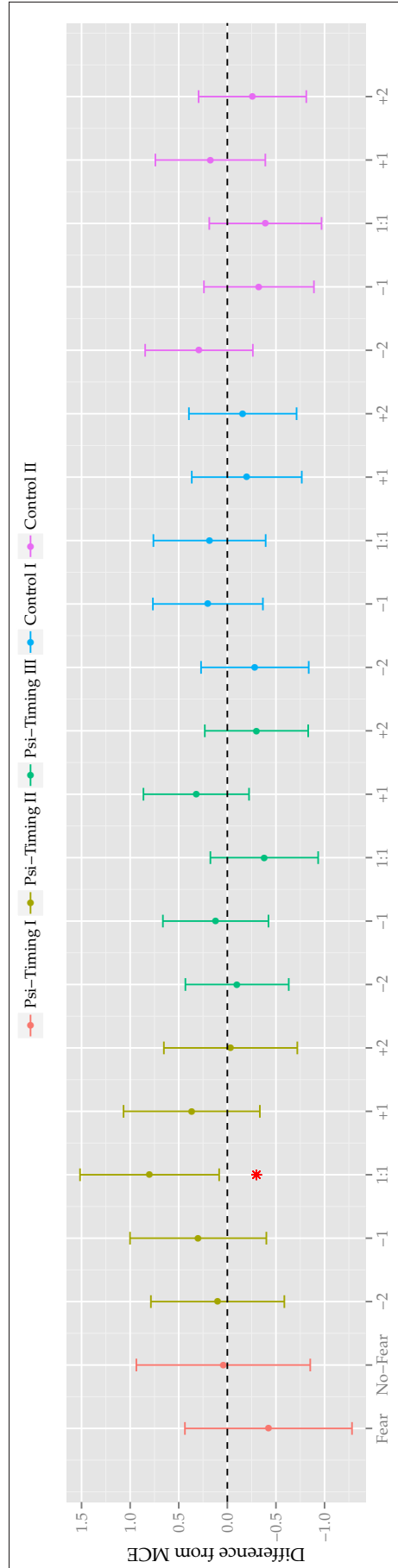


Figure 4.8: CI plot for all Psi-Timing & Control Trials. Presented are the means and confidence intervals calculated from z-tests. +2 and +1 refers to the number of hits ahead of the intended target (plus two and plus one respectively) and conversely -2 and -1 refers to the number of hits scored behind the intended target (minus two and minus one respectively). The one significant p-value (from the z-tests) is marked with a red asterisk (p-values have not been adjusted because the intention in presenting this data is to show what would have been found had the contrast of interest been different and note Holm's method (Holm, 1979) would reduce all the contrasts to non-significant at the .05 α level). The planned analysis was only to be conducted on the 1:1 contrasts and for Psi-Timing I the contrast of interest was the difference between the two fear groups. Due to the different hit rates and number of trials, the values have been standardised to allow comparison across the same scale.

in this chapter provide little support for the belief that people can paranormally adjust their behaviour to achieve a more positive ends. The reasons why no convincing effect has been found, have to some extent already been covered in this chapter and in the previous two chapters. The criticisms made by Hoppe (1987) of statistical errors and misunderstanding certainly seems one of the better reasons for the few reports that do suggest a psi-timing effect. And it is the null hypothesis (that there is no paranormal effect to be found) which again provides the best explanation for those studies that provide no support of an effect. We are left then, with yet another area where the previous paranormal claims are not supported.

As for the participant that scored significantly above-chance in Psi-Timing II, Martin Gardner recounts a relevant observation:

A distinguished mathematician at Keio University was explaining elementary probability to junior high school students. To illustrate a point he had his assistant toss in the air eight poker chips, each one red on one side and white on the other. All eight fell with the same color up! You can imagine how shaken the professor must have been. The probability of such an event is $\frac{1}{2^8}$ or a bit under .008. (Gardner, 1998, p. 15)

It is thus a final conclusion that between wholly expected and wholly unexpected results, there is a grey area, and we must use intuitive and statistical assessments as to whether the result we have witnessed is the product of chance, or the product of a biased process (whether that bias be paranormal, or some other influence). Without further testing it is impossible to completely deny the possibility that the high-scoring participant was not mediating their timing responses through psi, but equally, given the set-up of the experiment, it is not impossible that chance could produce the same results.

Finally, we can only speculate that the reason that the psi-timing paradigm has fallen into desuetude is because it has not yielded the experimental results that other scientists demand of parapsychologists. Given the near ubiquity of computing technology today, it is with some regret that such a simple methodology has not provided more success.

An Indirect Test of an Exceptional Precognitive Claim

5.1 Do Independent Judges Agree with the Putative Correspondence?

5.1.1 Introduction

In the summer of 2002 there arose an opportunity to work with a person who claimed an exceptional precognitive ability.¹ David Mandell (hereafter referred to as DM) was at the time a 69-year-old retired art lecturer, who claimed that throughout his life he had experienced frequent precognitive dreams and that when he could, he had produced a picture, painting or sketch of the dream as a record of the premonition. DM claimed to have produced over 200 pictures over a period of years and had corroborating photographic evidence supporting this claim for a number of his best “matches”.² Because of the lack of empirical evidence for precognition that had been demonstrated with the earlier research in this thesis, the opportunity to investigate a specific claim of accurate and extensive premonitions was deemed an invaluable source of evidence to examine.

Before engaging in a specific investigation of DM, a further review of dream precognition was undertaken (in addition to that material already covered in Section 1.2). There are a number of relevant studies into precognitive dreaming in the literature, mostly inspired by the influential book by Dunne (1927), *An Experiment with Time* (see page 18). Briefly, Dunne’s fame at the time is hard to appreciate now in the 21st Century, but it is not an exaggeration to describe him as something of a celebrity. And for the purposes of this discussion it is worth pointing out that Dunne recorded his own dreams and then searched the daily papers for news events that might match his dream content.

Despite the novel theoretical explanation proposed by Dunne (1927) to explain his own apparent ability to dream the future (in that an infinite number of internal observers “experienced an infinite number of “times”), it was the fact that he experienced so many dreams that apparently contained foreknowledge of future events, which was inspirational to parapsy-

¹CF was approached by a media production company and some of the work from this chapter was commissioned to appear in a Channel 5 documentary, *The Man Who Paints the Future*, broadcast in 2003.

²As will be discussed later in the chapter, we would not refer to DM as a “psi-star” (people who claim to be able to score consistently above chance in experimental tests) in the same vein as Pavel Stepanek (Pratt, 1973), Bill Delmore (Kelly & Kanthamani, 1972), Uri Geller (Randi, 1982), Malcolm Bessent (Honorton, 1970), etc.

chology. If Dunne could find numerous occasions of apparent precognition, why not anyone else? Dunne's particular methodology of investigating dream content has been used a number of times over the following years, although never with the same dramatic results that Dunne claimed. Besterman's (1932) work was seen as a direct empirical test of Dunne and Dunne was himself consulted for advice in determining the best way to proceed. Besterman recruited 20 volunteers to record their dreams over an 8-week period, collecting a total of 265 dreams, covering nearly a thousand written pages. All of the dreams were reread once a week to allow for identification of any life event that might match the dream records. Besterman claimed to have found only two experiences which seemed precognitive, out of the much larger sample of dreams and even these were only moderately interesting. As Dunne believed the age of the sample had biased the experiment against success (the average age is reported as 45–50), the method was repeated with undergraduates. Again, despite the extensive rereading of dreams over a period of weeks, only a small number of apparently precognitive dreams were identified (33 out of 113). Out of the 33, only two were rated as impressive and Besterman himself stated that neither of the cases were evidentially conclusive.

Bishop (1941) is one of a number of people who have published investigations of their own dream content. Again inspired by Dunne (1927), Bishop recorded 12 dreams over a number of years, which she felt were premonitory. On analysing the dreams as a set, Bishop concluded that most of the dreams were trivial and suggested that she was somehow picking up minor events from the coming days. Unfortunately, Bishop reports that whilst the dreams were recorded before the event it is regarded as matching, there is no proof of this, in that no one can confirm the veracity of her claims. However, given the trivial nature of the event and match, it would be a more likely argument that, given the mundane nature of the dreams, some event would have eventually occurred which would have somehow matched the dream content. This particular argument is pursued more fully later in this chapter.

Schriever (1987) reported an analysis of some 28 years worth of dreams sent to him courtesy of a Mrs M (no relation to DM) who claimed frequent precognitive dreams. A total of 3000 dreams were filed, although only dreams clearly dated were used, leaving 115 dreams (or approximately 4 percent of the total). It is not reported why Mrs M was not told to clearly date her dream reports given the effort she made in submitting so many dreams to Schriever. Out of this set, Schriever identified some 48 as well matched, and hence suggestive of precognition and the remaining 67 as not being precognitive. No significant conclusions were made regarding the general nature of the matching dreams, although perhaps it is telling that Schriever remarked that Mrs M seemingly had no criterion for distinguishing precognitive dreams from non-precognitive dreams and it was only the finding of an event which matched the dream, which resulted in this paranormal attribution regarding the source of the dream. That is, when an event happened that matched a dream, the dream was then identified as precognitive. Also seemingly important was the fact that Mrs M refused to entertain any alternative explanation regarding the link between dream and event and constantly maintained the paranormal hypothesis as the only explanation for her experiences, suggesting perhaps a paranormal example of Festinger's cognitive dissonance (e.g. Festinger, 1957).

Finally Sondow (1988) attempted to investigate her own dreams and identify paranormal content over an extended period. Based again on Dunne (1927), Sondow argued that for a dream to be considered precognitive it must contain at least two different and infrequent elements not usually found together and this is done to reduce the effect of spurious, coincidental hits. This requires a probability judgement on the part of Sondow in determining just how likely or

not any particular event is. Out of a total sample of 943 nights in which a dream was recalled, 96 were classified as precognitive, though none are spectacular matches. Dunne raised the concern that given enough time a match will occur by chance, and he restricted himself to two days in the future when looking for residual content. Despite this particular constraint it seems that none of the authors would dismiss matches falling outside of this time-frame if it is impressive enough, and hence this is a plastic restriction, specified to give the illusion of control, when in fact it is much less restrictive if required.

Premonitions of Disasters and Other Events

Although this thesis has primarily been concerned with premonitions that are directly associated with the percipient, it is true that there is a large class of experiences that fall into what is best described as disaster detection. As mentioned in the introductory chapter (Section 1.2.1.1) psychiatrist Barker investigated premonitions associated with the Aberfan coal-tip disaster and concluded that some individuals act as “human seismographs” for disasters: not directly affected by the event, but seemingly picking up information days, weeks or months before the event occurs. He suggested that such people might suffer from what he termed ‘pre-disaster syndrome’ and that although these people often lacked specific insight into the form that the disaster would take they would often develop

intense mental and physical unease for a variable interval beforehand. Their symptoms include severe anxiety, oppression, depression, headaches, lack of concentration, choking sensations, difficulty in breathing and feelings of constriction and suffocation. Their distress is relieved after the disaster has occurred or upon hearing news of it. (Barker, 1968, p. 165)

Although there are numerous examples in the literature of disaster premonitions (and some of those were covered in the introductory chapter) a small number of interesting cases will be examined further here. A particularly striking example of someone who seemingly suffered from so-called pre-disaster syndrome is the story of David Booth (Lowe, 1980). Booth claimed that for 10 consecutive nights he experienced the same dream:

I was standing beside this one-story building. . . and now I’m looking away from the corner of the building and I’m looking out over a field and there’s, like, a line of trees going down and I look up in the air and there’s an American Airlines jet, a great big thing, and the first thing that strikes me—that always struck me—was that it just wasn’t making the noise it should be for being close, you know?

Then it starts to bank off to the right. And the left wing goes up in the air and it’s going very slow. It wasn’t like slow motion. It was just going slow, and then it turned on its back and went straight down into the ground and exploded. . . . (Lowe, 1980, p. 106)

Booth claimed that the dream was unlike any he had experienced previously in that it was indistinguishable from reality and that remembering the dream afterwards was like remembering a real experience and not a dream. Booth’s daily experience after the dream followed Barker’s pre-disaster syndrome. After experiencing the same dream for approximately a week, Booth contacted the local office of American Airlines (AA) to inform somebody with authority about his dream. Booth was unable to get any response from AA and so contacted officials at the Federal Aviation Authority (FAA) and managed to speak to Ray Pinkerton who was assistant manager for airway facilities. Pinkerton made notes on Booth’s dream: “He sounded so sincere, I just couldn’t slough him off. He sounded truly concerned.” Pinkerton reported the dream to Jack Barker, public affairs officer for the FAA, who complained that

the dream lacked specificity and was similar to any of the hundred dreams that had been reported to him over the 25 years working in the aviation industry. Booth reported the dream to Pinkerton on May 22nd. On the afternoon of Friday 25th, American Airlines flight 191 crashed during take-off (from Chicago's O'Hare Airport bound for Los Angeles), with the loss of 271 lives. As the plane took off, it lost an engine, rolled to the left and crashed into the ground. It was the worst domestic crash in U.S. history. Booth believed that this was the crash that he had seen in his dream. An account of Booth's experience was reported in the papers across America.

Whilst an interesting and dramatic case, which seemingly has excellent corroboration of it happening before the event (in the form of the two federal government employees), there are a number of "facts" that must demote the value of this case in regards to its trustworthiness. Firstly, though there is no authoritative source, it appears that in the intervening years since making his prediction David Booth has gained a reputation for other predictions and they seem to have been increasing in outlandishness. A recent book "written" by him entitled *Code Red* appears to be a heavily plagiarised attempt at predicting a Book of Revelations-like apocalypse. There is some evidence then that David Booth has been using his reputation as a "government proved psychic" to make money. Did David Booth hatch a plan in the late 1970s to make a false plane crash prediction in the hope that this would come true and enable him to make a living as a false prophet? The second aspect of David Booth's story that severely undermines his genuineness as a percipient is that research appears to confirm that he is the youngest hijacker in U.S. history. On Monday 10th of November 1969, a 14-year-old boy named David Booth, attempted to hijack a Delta Airlines jet, by threatening a young woman with a butcher's knife. Again, although no authoritative source exists confirming that this David Booth and the David Booth of premonition fame are one and the same, research appears to suggest they are. This also raises the question as to whether either of the government employees were aware of this connection? Thus, although at first examination David Booth's claim appears to be a striking case, the suspicions raised about his mental state and truthfulness must reduce this claim to one that is interesting but hardly evidential. (A similar case is recounted in Inglis, 1990 in regards to the Air India flight 182 bombing in 1985. A retired schoolmaster in the UK dreamed he was looking out over the ocean. Up in the sky he witnessed an aircraft break up and watched hundreds of passengers fall from the sky into the water. The man apparently had the dream at least three times. Inglis reports that the man discussed the dream with two friends. Some two weeks after the dream, terrorists exploded a bomb aboard Air India flight 182, killing over 300 people, and the man recognised the event as the one depicted in his dream.) Whatever the veracity of Booth's story, other similar accounts can be found in the literature.³

Although in David Booth's account he appears to be a "witness" to the plane crash, not all premonitions involve such a direct association with the event. Some premonition experiences relate to news reports or newspaper headlines regarding the event itself, rather than experiencing the event as it will happen. Hearne (1989) calls this the premonition of media-announcements and he estimates that approximately 1 in 50 premonitions is of this particular type.⁴ Hearne (1982a) reports an interesting experience in relation to one of the worst peace-time accidents in Britain, the 1974 Flixborough disaster. At approximately noon on the first of June 1974, a 24 year old female living in Cleethorpes was watching television

³Evans (1983) found that 10 percent of respondents to his questionnaires reported dreams of air plane crashes.

⁴Keith Hearne was involved in the same television programme on DM, although there was no collaboration on the examination of DM's claims.

alone, when the words 'Newsflash' appeared on the screen and a male voice announced that there had been an explosion at Flixborough and that several people had been killed and injured. A visiting couple entered the room shortly afterwards and the percipient relayed news of the events to them. At 4:53 PM that same day, a chemical plant exploded outside the village of Flixborough, South Humberside. The resulting massive explosion was described as catastrophic ('of warlike dimensions'), resulted in the deaths of some 28 people, and injuries to hundreds more. Approximately 2000 homes were damaged. Interestingly, had the disaster happened during a working day there would have been approximately 500 workers on site and with it, many more fatalities. Later on the percipient and the couple watched news reports of the disaster, noticed the apparent temporal discrepancy and came to the conclusion that an ostensible precognition had occurred.

Although the woman could not remember what she had been watching at the time of the newsflash, as this was before the proliferation of television channels that we have today, it was easy for Hearne (1982a) to get written confirmation that none of the channels had shown an unscheduled newsflash at the time claimed. Hearne argues that this case strongly supports the precognitive hypothesis mainly because at the time of the precognition nobody was seemingly concerned about the possibility of an explosion (a telepathic explanation) and he also argues that to a non-technical person clairvoyant perception of the state of the plant was unlikely to be understandable. Also given the unusual specificity of the experience, chance coincidence is regarded as highly unlikely. What Hearne fails to acknowledge with this case is that it is reliant on the memory of the percipient and two witnesses. Unlike many of the experiences in the literature (and including the majority of DM's) this experience occurred on the day of the disaster. It is not reported whether a newsflash occurred *after* the disaster had occurred that matched the percipient's experience, however it seems possible that this case is one of misremembering the actual time of the event, rather than a precognition of the disaster itself. Also, although Hearne states that nobody at the plant was concerned about the risk of a possible explosion, the Flixborough disaster has been the subject of a number of public enquiries and the accident itself was caused primarily by a repair job (and one that was intended to be temporary). With such a repair must come the possibility that the repair will not be successful (and one of the couple who was a witness in this case was an engineer, though unconnected with the Flixborough plant). Thus, as with the Aberfan disaster, it is possible that there was a rumour about the plant, that might have seeded any concern within the local community. (Obviously this is a much less likely explanation than the misremembering hypothesis.) It should also be noted that though Hearne does not state so in his paper, in his 1989 book *Visions of the Future* (Hearne, 1989), he describes the percipient of the Flixborough case as (p. 9) "the niece of a good friend". It must be offered as an alternative explanation of the case that Hearne has fallen for a joke or ruse by his "good friend" and perhaps one that it is difficult to admit, now that Hearne has reported it to a wider community (difficult because if the Flixborough account was initially a joke between close friends, it would now embarrass Hearne if it were exposed as such).

There are other premonitions of media announcements in the literature—and as will be outlined below, DM would often notice a distinct similarity between his dream imagery and subsequent newspaper headlines. Cox (1982) reported an auditory newscast, in what he describes as an unusually well documented case. Another unusual aspect to this case is that the percipient is none other than Cox himself (and Cox will be discussed further in the next

chapter).⁵ Cox reports that in October 1968 he was working on a PK machine and as was customary had the radio on in the background. He states that at approximately 1:05 PM he heard a newsreport of a fatal road traffic accident in the local area. He heard the same report repeated a further three times, the first two hours later and the last time at 6:00 PM. Cox reports that this last time raised in his mind the thought that it was unusual to hear news reports and for the information to not be updated. The next day Cox reports that he was travelling along the route of the accident and happened to talk to the bus driver, who recounted that he had been a witness to the aftermath of the accident. The bus driver apparently told Cox that the time of the accident was some 11:00 PM. The next day Cox telephoned the radio station and enquired as to the newscasts, upon which he was informed that it was impossible that he could have learnt of the news before 11:00 PM as no such newscasts had been made regarding the accident earlier than that time. Although Cox kept his experience essentially to himself over the preceding years, the publication of the Flixborough case instigated his reporting it in the literature.

One other similar case is given in Blackmore and Hart-Davis (1997). Mike Veshecco (district attorney in Erie County, Pennsylvania) and his wife were watching television on May 5th 1980, when unexpectedly the words “stabbing in the city” or “local man stabbed” flashed up on the screen. Both Mr Veshecco and his wife attested that as a consequence of the words on the screen they believed that there had been a local stabbing that evening. Coincidentally, two friends of the Veshecco’s (one of whom was also in law enforcement) also claimed to have seen the words on their television. As both men expected to be called in to deal with the event, they were reportedly surprised that no call actually came. The next day it was revealed that a fatal stabbing had occurred in the local area (though a few hours after the words had been seen on the screen), and that when the words had been seen, the man was alive. Although not as well documented as the other two cases, it appears that some premonitions can appear to be indistinguishable from normal news reports. This obviously raises the interesting question as to whether there are cases where the percipient does not recognise the temporal displacement of the experience and that learning of the event later, the unrecognised discrepancy does not lead to recognition that something unusual has occurred? This may in fact explain the rare nature of this type of experience. They are more easily not-recognised than other premonitory experiences.

Returning once more to the percipient of the Flixborough premonition (Hearne, 1982a), it is not the only precognition reported by this individual. Hearne reports that she has also experienced another explosion premonition (though this time it is unaccompanied by corroborating witnesses, as she did not report it to anyone else). It is reported that while taking a walk on the morning of 24th June 1980, the percipient felt compelled to take a different route than the intended one, and that a strange idea entered her head—that it would be awful if a gas explosion occurred in one of the houses. At lunchtime that day there was an accidental explosion of a gas-cylinder, though no one was hurt. This is obviously more like the kind of death-avoidance experience that has been the main consideration of this thesis. Hearne (1982b) adds a further interesting note to this case. In that paper he outlines the experience of another resident of Grimbsy and her premonition of the sinking of a submarine *Artemis*. Although there is nothing that raises this premonition above the many that could be reported, it is interesting that she too reports having passed by the house at which the gas-cylinder

⁵Hearne (1989) gives an account of a presentiment that he himself experienced. Cox is therefore not alone in having a first-hand experience of the phenomenon under investigation. For the record, LS has had no paranormal experiences, precognitive or otherwise.

explosion occurred shortly before it exploded. Hearne notes this, but does not draw any specific conclusions and alas, we find ourselves in a similar position.

Hearne (1982c) gives another account of three premonitions from a single percipient (not either of the percipients already discussed in relation to Hearne). As with previous premonitions the three are high profile events (two assassinations, one successful, and the other a ship sinking).⁶ Interestingly the percipient (Mrs Barbara Garwell) is reported as the seventh child of a seventh child. (It was noted in the introductory chapter that Hearne favours a theory of precognition that suggests it is a female trait and one that is passed down through the female line. There are also some traditions, like that of Scottish second-sight, that argue that the seventh child of a seventh child has certain special abilities in regards to “seeing the future”, Cohn, 1999; Campbell & Hall, 1968). It is reported that Mrs Garwell suffered three head injuries in her youth and that it was felt that these injuries in some way precipitated the precognitive experiences. Also common is the reported fact that the premonitions were in some way special and that they had a quality that was associated with premonitions. Like DM the percipient claimed to have experienced many precognitive dreams throughout her lifetime, though the dreams that Hearne was reporting were just from that year. Interestingly each of the dreams occurred some three weeks before the event that the dream was said to match. For example, the percipient dreamed of a ship sailing on an ocean. She described two coffins descending slowly down a type of ‘rope gang-way’. The percipient identified this ship as one that two friends would shortly sail on to South Africa. The percipient did not report the dream to the friends for fear of causing upset and because she was not certain whether the dream directly related to the friends. Three weeks later the ship carrying her friends did suffer an accident (a fire broke out) and *three* people died. One person jumped overboard and thus only *two* bodies were taken ashore via the gang-plank. Hearne thought that all three of the premonitions were impressive in detail and that although there was some concern about the role of chance coincidence, that the cases were suggestive to him of the idea that precognition may be an inferential process facilitated by ESP, rather than necessarily a direct knowledge of the future.

Hearne (1986) followed up this report on Mrs Garwell by conducting a pseudo-experiment with her over a single year. Hearne argues that one of the major criticisms of single case reports of premonitions is that they are potentially drawn from a much larger pool of possible premonitions and thus have been preferentially selected (and are a chance occurrence) rather than be a single, unlikely occurrence. To overcome this objection all of the percipient’s premonitions were collected over a one year period. Hearne employed two blind judges to evaluate the 52 accounts, who rated the correspondence between the premonitions and events published in a local newspaper up to 28 days ahead of the premonition. Hearne employed “control” data, by providing the relevant newspapers for the year before (for one judge) and the year ahead (for the other). Judges were not told what year the premonitions were from and thus Hearne was interested in comparing whether the premonitions were specific to that year, or could just as easily match another year. Statistical analysis was suggestive of a link between the event and dream, but according to Hearne did not reach overall significance. Three of the reports (out of the 52) received a rating of 6 or higher (on a scale of 0–8) and were matched according to the correct year. One of these three is given here (it was rated 8):

⁶Stevenson (1970a) argues that political assassinations are not an ideal topic for studying precognition, primarily because prominent politicians are a high-risk subject in regards to assassination and their prominence makes them the subject of a disproportionate amount of thoughts and dreams. Thus Stevenson argues that such political assassinations might be over-represented in apparent precognitions, whilst still being the consequence of chance coincidence.

As I got into bed, I closed my eyes and saw the outside of a castle. From the castle came about eight choir-boys or altar-boys. There were a lot of people together and much confusion. In the centre of these people I saw a figure in white. It was a person of State—e.g. the Pope. (Hearne, 1986, p. 380)

The event that this was rated as matching was an attack on the Pope, three days after the premonition, when a man attempted to stab the Pope with a bayonet (at the Fatima Shrine in Portugal). (If we recall, Stevenson, 1970a recommended not including assassination premonitions in assessing the question of precognition.) Hearne (1986) reports that this premonition was accompanied by an unusual feeling of conviction. Although there is certainly some correspondence between the event and the premonition, and the closeness in time, there again appears to be a lack of any specificity regarding the premonition. This suggests that this technique of testing premonitions (by comparing events in one year, with events in a control year) may not be the best method for investigating the issue. It is highly unlikely that a comparable event (for example involving the Pope) will occur in the relevant month in the control year. Does that strengthen the claim that the report is a premonition? In our investigation of DM we favour a slightly different, if also not perfect method. Finally, Hearne (1989) reports that Mrs Garwell continued to have precognitive experiences after his test with her concluded.

There are numerous similar single cases of precognition and case collections in the literature with more or less similar accounts (e.g. Wood-Trost, 1981; Hearne, 1989).

The Premonitions Bureaux

Dr J. C. Barker (previously discussed in relation to the Aberfan disaster) was not only a collector of precognitive experiences, he was also a consultant psychiatrist and was instrumental in setting up the British Premonitions Bureau in 1967 with Mr Peter Fairely, who was at the time the science correspondent of the London *Evening Standard*. The premonitions bureau was Barker's attempt to establish a precognitive 'disaster early warning system' (Barker, 1967a). His work on Aberfan had led him to conclude that certain sensitive individuals could pick up on major disasters *before* they occurred, and thus Barker hoped to tap into this seemingly haphazard insight into the future and apply a scientific cataloguing to the premonitions. He hoped that such a process would remove the inherent uncertainty found in the individual premonitions and would allow disasters to be averted. In the first year the Bureau received about 500 premonitions and in the first six years, the bureau received over 1200 responses, including some which appeared to be precognitive (Zohar, 1983).

MacKenzie (1974) in his book *Riddle of the Future* gives a dramatic account of the work of Barker and the Premonitions Bureau. Of the 3000 or so precognitions that were filed up to 1974, the bulk of the premonitions came from just six people. One of those was Mr Alan Hencher (who worked for the Post Office) and apparently had a precognitive gift for aircraft accidents. An example of one of Hencher's predictions:

Aircraft Caravelle over mountains. Will be leaving in early morning between Monday and Sunday—it is going over mountains. It is going to radio that it is in trouble. Then it will cut out and there will be nothing. There are 123 or possibly 124 people on board. It is going to crash shortly after take-off. I can't tell exactly where or when it is going to happen. One person is saved in a very poor condition. I have had this feeling for a week, but it has been very strong in the past two-three days... (MacKenzie, 1974, p. 137)

According to MacKenzie (1974), approximately one month later, on the 20th April 1967, a

Britannia airliner crashed into a hilltop near Nicosia in Cyprus. In the London *Evenings News* it was reported that 124 people had died but only 123 bodies were recovered. A later account reported that only four people of the 130 on board survived. The crash happened during a severe thunderstorm and the pilot was attempting to land for the third time when he stuck a hillock and then crashed into another hill, bursting into flames. Mackenzie argues that not only is the premonition a good (if not perfect) match to the future event, that there is some suggestion that the prediction is a good match for the early news reports of the disaster, rather than being a direct perception of the event (perhaps another case of a premonition of a media announcement?). This is somewhat supported by Hencher's description of his experiences:

Sometimes I see things in black and white—sometimes in colour. I never get pleasure in premonitions. Often I get a figure in my mind, as though someone had spoke it. The details seem to come almost as headlines, and they come afterwards. As regards geographical location, I don't get place names. I only see the immediate area and certain features. Occasionally I see words written, or see people' lips moving, but I am not near enough to get a clear picture. (MacKenzie, 1974, pp. 138–139)

It is interesting that Hencher also identifies a head injury with the onset of his premonitions. According to MacKenzie (1974), in 1969 (and after the sudden death of Barker) the Bureau came under the control of the *TV Times* weekly magazine, of which Fairely was the science editor. Blackmore and Hart-Davis (1997) give details of how the premonitions were scored (out of a total possible 15) when they were received. First premonitions were assigned a score out of 5 depending on how unusual the premonition was: a simple plane crash would be rated low, but a more unusual event would get a higher score. Next the premonition would get a score out of five based on the amount of detail it contained: scant detail was given a low score whereas a rich premonition full of detail would be awarded a high score. Lastly the premonition would get a score out of five based on the timing of the premonition: a precise premonition would get a higher score than one with vague, open ended timing.

The year after the British Premonitions Bureau was opened, a U.S. counterpart was opened with the same aims in mind, known as the Central Premonition Registry. It was founded by a newspaper journalist, Robert D. Nelson (with guidance from parapsychologist Stanley Krippner). It is reported that up to 1980 Nelson had received more than 7500 predictions, though only 47 were counted as hits, the majority of which were related to death and disaster. Of those 47 hits, approximately half of them were from six so-called "heavy hitters", one of whom was a professional psychic (Lowe, 1980).

Although no authoritative source exists, it is apparent that the hopes surrounding the founding of the premonitions bureaux have not been fulfilled. The fact that neither institution appears to still exist can be put down to a number of reasons. First the death of Barker will have removed one of the major impetuses. Secondly funding must surely have been an issue. Details of the running of both bureaux suggest that a large number of premonitions were submitted for examination and the filing and checking of this large data set was time consuming and ever-increasing. Given a lack of funds, the task itself was a labour of love and one that would eventually be subsumed by more important life events. Thirdly must be the obvious lack of success of either bureaux in doing what Barker had initially hoped would be the eventual outcome: To prevent a future disaster. There is no written record of any event having been prevented due to the work of the premonitions bureaux. Fourthly must be the fact that although no bureaux apparently exist today, the main parapsychology organisations e.g. the *Society for Psychical Research* still receive premonitions from members of the public. Thus

although there is no dedicated organisation that attempts to predict and avert disasters, there are organisations that can take a record of premonitions.

Before moving onto a discussion of DM's specific claims it is worth discussing one final case of an apparent precognition. Frazier (1980) recounts the dramatic prediction of Richard Newton in December 1978. Newton predicted in writing that there would be a major plane crash on the 11th of March 1979 outside a major population centre in the Northern Hemisphere, killing 46 people and that the aircraft's logo would have red in it. On the 14th March 1979 a jet belonging to the Royal Jordanian company crashed outside Doha (the capital of Qatar) resulting in the death of 45 people. The company's logo was maroon. Although this may appear to be a very good prediction made before the event had occurred, Newton used statistics to make his prediction (for a thesis on the ways that people can convince themselves and others that they have psychic powers).

According to Frazier (1980), Newton examined 26 years worth of plane crash data and calculated a statistical average for when the plane crash was likely to occur. March was the second most likely month after December. He used the average number of deaths associated with crashes in that month. Most flights were in the Northern Hemisphere, 50% of airline logos had red in them. Although Newton's prediction was entirely rational and conscious in origin, Newton argued that other "psychic" predictions are based on similar unconscious processing of information (without the person even being aware), that they are continually accumulating evidence which is used to make rational inferences about what might occur in the future and that these unconscious predictions are misinterpreted as being a consequence of ESP.

Although Newton's prediction is evidence enough that predictions regarding plane crashes etc. can be made entirely upon what has happened in the past, it is unlikely that DM's experiences are entirely a consequence of such unconscious processing (though it is likely that some percentage of the experiences are indeed caused by a more rational pathway).

DM's Specific Claims

As noted earlier, DM was at the time of testing a 69-year-old retired art lecturer, who claimed that throughout his life he had experienced frequent precognitive dreams which would often match the occurrence of a later event, usually a newspaper story or a television news report. As an artist, DM took to painting or sketching his dream as a pictorial record of the premonition. On the advice of another party and in the hope that the ostensible nature of his experiences could be confirmed, DM took to getting corroborating evidence by taking his finished pictures to a local bank (or estate agents) the morning after they had been produced and getting a photograph taken in the bank of himself holding the picture in front of the date and time. A bank or estate agents was chosen because of the perceived trustfulness of the institutions involved. By doing this he hoped to eliminate the criticism that the pictures were merely completed after the event occurred. One example of DM's pictures is reproduced in Figure 5.1 and further examples can be found in Appendix B (along with a hyperlink to a digital archive of all DM's pictures used in this study).

DM's most sensational precognitive experience and one of the aspects that raises DM's claim to the level of exceptional, is associated with the 2001 terrorist attacks on the Twin Towers of the World Trade Center, New York.⁷ DM reports that on the morning of September 11th 1996, he had a dream about the destruction of the Twin Towers from which he produced a picture

⁷Dr Sally Rhine Feather (daughter to both Joseph Banks and Louisa) published a book on premonitions associated with the September 11th attacks and reports that the event generated the largest ever number of responses submitted to the Rhine Lab (Feather & Schmicker, 2005).

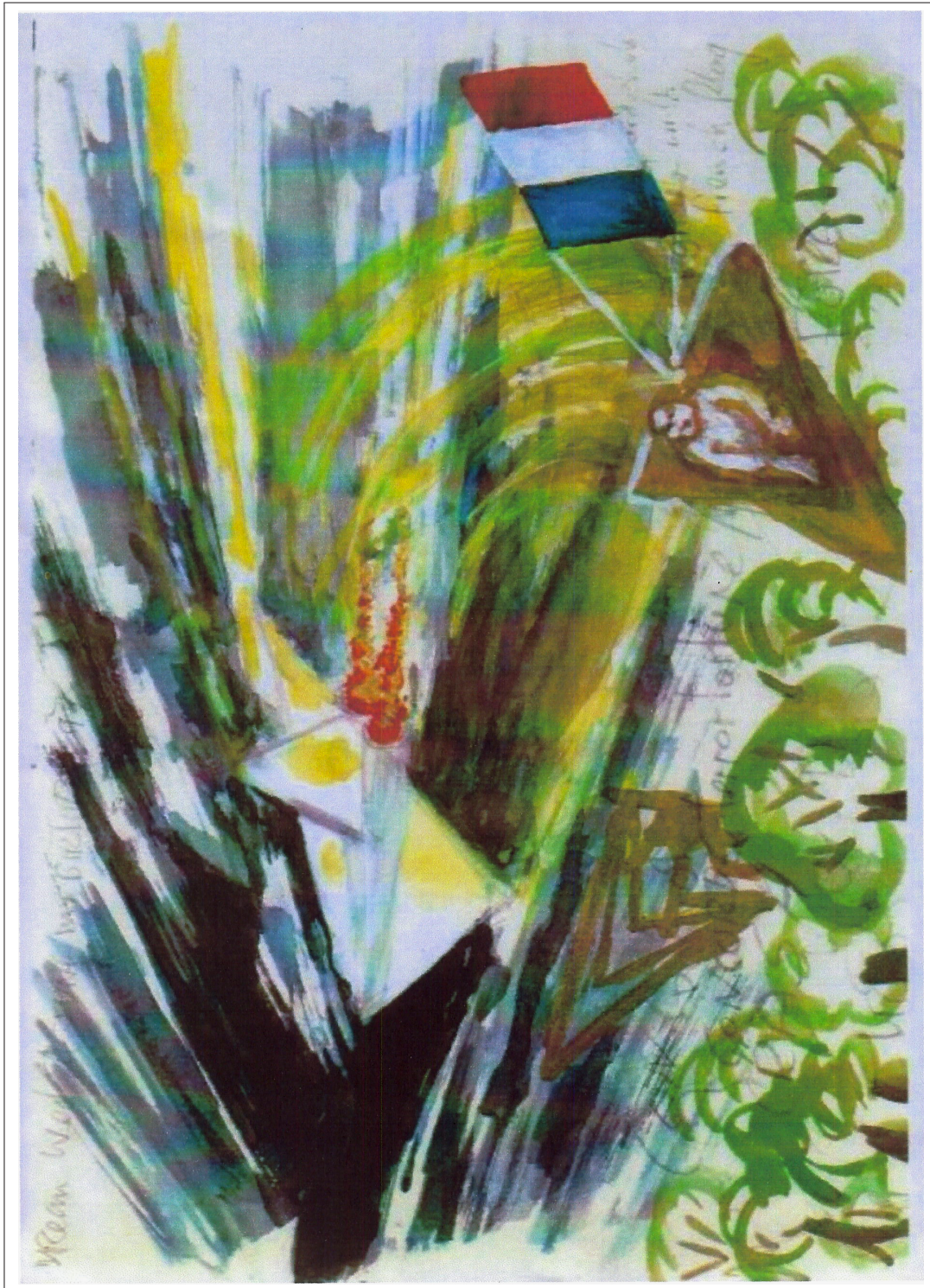


Figure 5.1: "Concorde Crash" painting. Further details pertaining to this image can be found in Appendix B, page 224.

and with it, later that morning he attended a local bank to get the accompanying corroborating evidence. Reproductions of both "Twin Towers" pictures can be found in Appendix B.⁸ Examination of the relevant picture certainly shows what appears to be a sketch of the New York skyline, with the Statue of Liberty particularly prominent and easy to identify. This then identifies the other two buildings as the Twin Towers in New York. At first examination there does appear to be a greater-than-chance correspondence between the picture, the date of the experience and the subsequent terrorist attack that would occur some five years hence from the date of DM's experience.

In addition to the bank photograph, DM also has a number of newspaper clippings that he claims bear a striking resemblance to the relevant picture. Examination of these supplementary clippings does show some correspondence between the newspaper clippings and DM's pictures.⁹ It is difficult to quantify that similarity objectively, but from a subjective perspective some of the supplementary images *do* share features with DM's picture, although it is certainly not unequivocal.

Although DM's claims are impressive and the Twin Towers example particularly so, more detailed examination of his pictures reveals a number of incongruities between the picture and the actual event it is purported to match. Taking the Twin Towers sketch, it appears to depict the events of the devastating attacks on those buildings in the year 2001 but there are a number of criticisms that can initially be raised concerning the degree of actual matching between the picture and the event. Specifically the sketch clearly shows the left hand tower crashing into the right hand tower, where in reality both towers came straight-down independently of each other. Moreover, before the events of September 11th had happened, DM actually believed that the sketch depicted an earthquake. As such, there was then already some concern as to the degree of matching that actually existed between DM's pictures and the events that he claimed matched them.

As mentioned, the corroborating photographic evidence provides an interesting boost to DM's paranormal claim and shows DM in front of a clock showing the date and time. For the Twin Towers picture, the date is Wed, 11th September and although the year is not given, the last time that that date fell on a Wednesday, was some five years earlier in 1996. Does this then prove that DM's dreams are precognitive? Interestingly, the "birthday problem" poses a similar situation by asking, "if there are 40 people in a room, what are the chances that at least two of them share the same birthday?" The answer is 89 percent, a number much higher than most people assume (Diaconis & Mosteller, 1989). One could easily argue that given the large number of pictures to choose from (over 200 in his lifetime) that it would be more surprising if such a match had not been found between one of the dreams and a future event. In other words, the Twin Towers picture would be almost as impressive without the corresponding date connection and were another picture to have a similar date coincidence, for example had DM painted the "Concorde Crash" picture on an anniversary before the event, then no-doubt that would have been raised as a significant aspect for that picture. Moreover, if the picture had been produced on September the 12th, or September the 10th, would it not still be an impressive near miss (Diaconis & Mosteller, 1989)?

It is important to note that not all pictures and sketches were of the same quality or were accompanied with photographs. Some were of a very general nature and others were

⁸DM produced two images which he felt were connected with the Twin Towers disaster, only one of which was produced on the morning of September 11th 1996.

⁹Copies of the supplementary material used in this study can be found in the DM digital archive, the url to which can be found in Appendix B.

more specific. On the face of evidence provided, three possible conclusions were identified as potential explanations for DM's experiences:

1. DM has a genuine precognitive ability which current scientific understanding cannot explain.
2. DM is a clever con-artist, hoaxer or fraud.
3. The match between picture and event is coincidental.

It must also be noted that DM's claims did not involve any conscious control of his ostensible ability and thus his claim was not seemingly amenable to the type of empirical testing that would provide the best test of his ability. For example, Blackmore (1995) directly tested the well-known "psychic" Chris Robinson, who claimed to have had precognitive dreams of air disasters, traffic accidents, etc. Along with this ability, Robinson also claimed to be able to successfully guess the contents of a closed box and it was this latter skill that Blackmore tested. Over six weeks, objects were hidden in a box (two per week) and Robinson recorded his dreams over the period. At the end of testing Robinson was sent a list of the 12 objects used and a photograph of the object, and his task was to guess which of the objects had been in the box during the course of testing. That particular experiment failed to support Robinson's psychic claims.

Thus developing an empirical regime to determine if (1) was correct was deemed too difficult given time and economic constraints, with little assurance of success were such a regime to be undertaken. There was also no way to experimentally manipulate events in the real world to facilitate DM's precognitive dreams, and there was no time to record dreams and wait years for their fulfilment. Therefore a direct assault of DM's claims was abandoned in the light of the specificity of his particular claim.

In regards to (2), it was not possible to completely rule out the possibility that DM was a fraud or hoaxer. Given the exceptional claim, efforts were made to ensure that DM was not part of a ruse to test our Anomalistic Psychology Research Unit in the manner of Project Alpha, (e.g. Randi, 1983a, 1983b, 1986; Truzzi, 1987). However, given DM's effort at attempting to provide stringent controls over the documentation of evidence (something distinctly lacking with other "special subjects" e.g. Marks & Kammann, 1980), it was concluded that as far as could be assessed, DM was not making any obviously fraudulent claims, had not apparently doctored any of the photographic images, and was essentially sincere and accommodating. Given the physical size of the photographic corroborating evidence, it was not possible to confirm that every element of the picture was original to the day of dreaming, that is, notes could have been added after-the-fact, although the general shape and colouring of the image was enough to confirm the prior existence of the majority of the images. Finally, DM appeared not to be financially motivated or motivated by a desire to gain media exposure. He claimed to be motivated primarily by a sense of responsibility regarding ensuring relevant authorities were warned of forthcoming events. Although DM was 'eccentric', and evidence of his living conditions implied a certain level of personal neglect, he did not report any psychological or psychiatric problems and no independent psychiatric evaluation or assessment was sought.

Finally, in considering (3), this is essentially the null hypothesis: that there is no causal connection between DM's pictures and the event it is alleged to correspond to, and that any correspondence is purely coincidental. Given all the billions of people in the world, all dreaming every night, it could be argued that it would be much stranger if dreams *never*

appeared to come true, just on the basis of chance alone. Aristotle argued that if you throw enough arrows, you will eventually hit the target (and drawing a target where the arrow has hit, is easiest of all.)

DM's huge number of precognitive drawings are comparable to Dunne's own collection of dreams and is certainly unparalleled in providing a visual illustration of precognitive dream content. That some of DM's dreams seem remarkable matches for future events, such as the "Twin Towers" pictures, must be weighed against the fact that many more of his pictures are ambiguous and only generally match a subsequent event. Thus, an alternative line of testing was developed on the basis of the inability to directly test DM. Since DM's claims were specifically linked to a paranormal attribution based around the content of each picture, a regime of testing based on examining to what extent each picture really did match the event that DM had chosen, was developed. This was based to some extent on a method employed by Hearne (1984), where two judges had been used to compare dreams either with an event close in time, or one from another control year. As Hearne had shown that the judges could not accurately match the premonition with its future event, it was hoped that a similar methodology could be used to show that participants would not show obvious preference for DM's matching event, over another similar event. The paranormal hypothesis was thus: that independent judges would rate DM's event as corresponding better to the relevant picture, than another, alternative event. The null hypothesis then was that no such preference for DM's events over an alternative event would be shown. As such this was not a specific test of DM's abilities, but was an attempt to see if judges could differentiate between DM's events and an alternative, and thus agree with the putative correspondence.

Essentially this study employed a technique developed by Marks and Kammann (1980) known as the *delayed control group*. In their own testing of remote viewing and psychic 'performances', Marks and Kammann argue that if participants in their research can recreate without the use of allegedly psychic powers, the kinds of evidence as claimed by the paranormalists to represent paranormal functioning, then the paranormal hypothesis becomes redundant. For example, in their book, *The Psychology of the Psychic*, Marks and Kammann show that a drawing made by Uri Geller purportedly through psychic impressions of the contents of a sealed envelope, was no more impressive, as judged by independent raters, than normal participants asked to draw the contents of a sealed envelope, without looking or guessing—Marks and Kammann actually used the same images used to test Geller.

In this current study, we are interested in whether independent judges show a preference for DM's event over an alternative event (referred to as the ALT), chosen because of its similarity to the picture, as judged by the main experimenter. If the match is as good as DM claims it to be, our independent judges should show a preference for it, since it is the event that DM argues *is* the cause of the picture. However, if DM has selected this event because it matches a previous dream, but there is no causal link between the two events, then independent judges should not show a preference for either event. We favoured this procedure over the alternative used by Hearne (1984) (where judges were given newspapers for the year before or the year after the intended year) because it is clear that DM does not know when the premonition will be fulfilled when he makes the dream image. As such restricting ourselves to a year before or after the "event" (for example comparing September 2000 or 2002), is guaranteed to *not* give us any event that matches DM's dream. By comparing the event to another similar event (one that DM *could* have chosen as the target event) the comparison is fairer. (It is interesting that Hearne, 1982b notes, that in regards to the premonition of the sinking of the submarine *Artemis*, that

there is a similarity between the account and another event that happened *on the same day*—the death of three Soviet cosmonauts as their vehicle was compromised during re-entry. As Hearne asks (p. 286): ‘...was the precognition really about the Soyuz—but portrayed in imagery familiar to the percipient?’ Are multiple endpoints detrimental to precognitive reports?)

It should be remarked however, that given the large number of pictures that DM produced and the time frame that he had to collect matching events means that the alternative event is at a disadvantage: time and economical constraints meant that alternative events were matched as best they could, but better matches may exist.¹⁰ However, as an exploratory indirect test of DM’s claims, these factors were not considered overwhelming, given the rare opportunity such a well documented case represents.

5.1.2 Method

5.1.2.1 Design

This study employed a within-subjects design, whereby participants would supply ratings as to how close both DM’s event and the alternative event matched each picture. A between-subjects variable of paranormal belief was included as a secondary contrast.

5.1.2.2 Participants

Some 30 participants volunteered (18 females, 12 males), aged from 20 to 67 (mean of 34 years and a standard deviation of 12.5 years). The mean belief score, as measured by the Australian Sheep-Goat questionnaire (Thalbourne & Delin, 1993; Thalbourne, 2010), was 9.3 with a standard deviation of 7.5. Although there is some bias to being sceptical (goats) over believers (sheep), there are some high-belief individuals.

5.1.2.3 Materials

Forty of DM’s best pictures were provided along with a short abstract describing the event that DM said the picture was supposed to match. An intensive search of internet news-sites facilitated the collection of 40 alternative news events (hereafter referred to as ALT) which matched the picture to some degree. This process was a subjective one, limited by a number of circumstances. Time constraints meant news sources could not be searched indefinitely and resulted in a number of ALTs being of a lower quality than DM’s event (although given enough time better examples would have been found). A number of the pictures are symbolic or ambiguous and hence there was some latitude in interpretation. ALTs were selected from both before and after the date the picture was produced.

Two sets of test booklets were compiled containing the 40 images, with any significant date information redacted, and the accompanying DM and ALT descriptions. Each booklet consisted of a different randomised order of presentation, to attempt to control for possible order effects. Supplementary material was also provided for a number of pictures (6, 7, 10, 12, 13, 17, 30, and 39) mostly consisting of news clippings that DM insisted bore strong correspondence with his dream content. Elements from the booklet are reproduced in Appendix B including the instructions, the event descriptions and a hyperlink which points to a digital archive containing copies of all the images used and supplementary evidence, such as newspaper clippings. It is important to note that the 40 pictures used were part of a much larger pool and were primarily

¹⁰The main constraint was imposed by the television production company who had commissioned the research.

chosen based on subjective appraisal regarding significance (that they were DM's best matches). Finally, the Australian Sheep-Goat questionnaire (Thalbourne & Delin, 1993) was included at the end of the booklet to measure belief in the paranormal. This is an 18 item questionnaire with responses of yes, maybe and no (scoring 2, 1 and 0 respectively) designed to provide a measure of paranormal belief.

5.1.2.4 Procedure

Participants were presented with test booklets, which first outlined the basic rationale of the study, then given some brief biographical information related to DM's claims and outlined what the procedure involved (reproduced in Appendix B). Half the participants were given Booklet A, and half Booklet B. Both booklets presented the pictures in a random order and reasonable efforts were made to try and prevent participants identifying whether descriptions were one of DM's or an ALT. It must be noted however, that it was felt prudent, given DM's claims, to include the newspaper clippings and to point out the putative correspondence, and this could not be done for the ALT events. Participants were asked to rate the degree to which they felt the pictures matched the event for each description and the picture on a 7-point scale (1 = no match at all and 7 = perfect match) for all 40 pictures. On completion of the rating task, participants completed the Australian Sheep-Goat questionnaire.

5.1.3 Results

Given the difficulty in assessing DM's claim, the planned contrast was to investigate whether independent raters would give the events picked by DM as the match for his dream picture a higher rating than an alternative event that had been chosen because of a subjective similarity to the picture. DM was quite sure that participants would show a preference for his pictures over the alternative match.

Firstly, Table 5.1 shows the responses to the question, "Before today, had you ever heard of David Mandell?", rated on a scale of 1 (definitely not) to 5 (definitely so). It clearly shows that the majority of participants (80 percent) have definitely not heard of DM and only 2 participants (6.7 percent) have definitely heard of him. Next, Table 5.2 shows the responses to the question, "Do you think it is possible for dreams to provide a paranormal glimpse into the future?", again rated on the same 5-point scale. Here we find there is a more mixed response, with approximately 47 percent rating low to definitely not, and a further 17 percent rating high to definitely so. Moving on, Table 5.3 shows responses to the question "Do you think all such ostensibly precognitive dreams can be explained in terms of concepts already accepted by modern science (e.g. coincidental matches and/or people dreaming about future events which they already know to be probable) ?", again on the same 5-point scale. We can see that nobody believes that *no* ostensibly precognitive dreams can be explained by science (in other words, this rating seems to reflect the opinion that science can explain some precognitive dreams), some 23.3 percent have given ratings that indicate a belief that science fully understands precognitive dreams with reference to normal, understood concepts, and 60 percent show a middle-to high rating, reflecting opinion that science has some understanding of the process of precognitive dreaming.

Table 5.4 shows the responses to the questions, "Have you ever had a dream which appeared to foretell the future?" and "If yes, how many such dreams have you had?". There is a 50 percent split between those answering in the affirmative and negative, and of those who

Table 5.1: Frequency of responses to the question, "Have you ever heard of David Mandell?"

Rating	1	2	3	4	5
%	80	6.7	3.3	3.3	6.7
N	24	2	1	1	2

Table 5.2: Frequency of responses to the question, "Do you think it is possible for dreams to provide a paranormal glimpse of the future?"

Rating	1	2	3	4	5
%	13.3	33.3	36.7	13.3	3.3
N	4	10	11	4	1

Table 5.3: Frequency of responses to the question, "Do you think all such ostensibly precognitive dreams can be explained in terms of concepts already accepted by modern science?"

Rating	1	2	3	4	5
%	0	16.7	40	20	23.4
N	0	5	12	6	7

Table 5.4: Frequency of responses to the question, "Have you ever had a dream that appeared to foretell the future and if yes, how many?"

Yes	No	Frequency	N
50%	50%	1	3
15	15	3	3
		4	3
		5	1
		6	2
		12+	3

Table 5.5: Summary results from the Australian Sheep-Goat Scale

Min	1st Qu.	Median	Mean	3rd Qu.	Max
0	3.5	9	9.3	14	29

report a previous precognitive experience, approximately half of those have had 5 or more such dreams. It is interesting to note that 3 of the participants report to have experienced 12 or more precognitive dreams.

Finally, Table 5.5 shows the summary of results of the Australian Sheep-Goat questionnaire (Thalbourne & Delin, 1993). This questionnaire is composed of 18 items scored on a “True”, “?” and “False” scale, with “True” scoring 2, “?” scoring 1 and “False” scoring 0. The maximum score possible is 36 and the maximum score for all ? ratings would be 18. The range of scores obtained suggests a heterogeneous distribution of belief in the paranormal (i.e. there are not just those sceptical of, or believers in, the paranormal).

The overall assessment for these initial results must be that most of the participants have never heard of DM, are open to the possibility that precognition exists and that there is a 50 percent split between those who have never had a precognitive dream and those that have. Given the reasonably small sample size, there doesn’t appear to be a major bias in favour of one perspective (contra or advocating).

The primary means by which we would test whether DM’s event was rated significantly higher than the alternative event was by comparing the results of 40 paired-samples t-tests for each of the pictures (ALT vs DM rating). Figure 5.2 shows the result of this analysis. It is evident from this plot that of the 40 contrasts tested, only 6 favour the ALT event, 3 neither event and the majority favour DM’s event to some degree. Given the large number of comparisons the resulting p-values were adjusted via Holm’s method (Holm, 1979) (using the p.adjust function of R) and the results were: Of the 40, only on seven pictures is DM’s event rated significantly higher ($p < .05$) than the ALT event (highlighted with a red asterisk on Figure 5.2). Thus despite the general trend to favour DM’s event over the ALT, this preference is only statistically significant on some seven pictures. It might be argued that this is significant given the number of contrasts (given the alpha level of 0.5, we would have expected just 2 by chance), however, given that the pictures are DM’s *best* matches out of a larger set of some 200 pictures, and given the wide scope that DM had in selecting matching hits and given the time constraints and other difficulties inherent in developing a database of alternative events, it is obvious that there are non-paranormal reasons for the fact that our participants have judged the alternative just as plausible as DM’s purported match for 82.5 percent of the images. In the spirit of the delayed control group, this implies that there are other events out there that match DM’s images and one only has to wait long enough for a good match to come along.

Table 5.6: The seven pictures where a significant difference between ratings for DM and ALT event was found. The p-values are two-tailed. There were 40 contrasts in total and thus the p-values have been adjusted via Holm’s method.

Picture	Description	t	P
28	Rackham’s knife attack	3.67	.033
37	Two miners killed	3.71	.03
15	Indian earthquake	3.88	.02
7	Heathrow mortar attack	4.18	.009
9	Dead fish in River Irfon, Wales	4.47	.004
35	Twin Towers colour painting	5.65	<.0002
33	Twin Towers sketch	7.65	<.0001

Results for the seven pictures for which DM’s event was rated as a significantly better match than the ALT event, can be found in Table 5.6 (copies of these pictures can be found

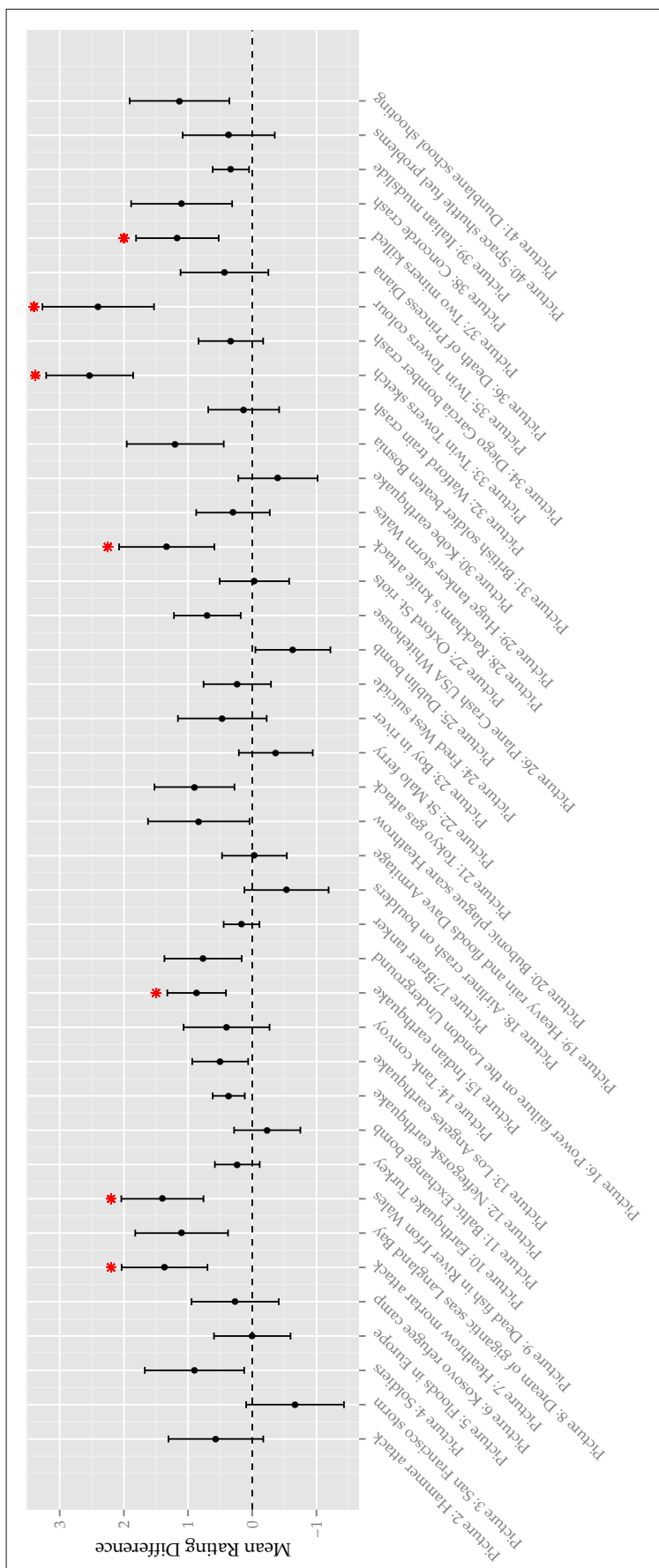


Figure 5.2: Graph showing the mean difference between the rating for DM's event and the alternative, with 95% confidence intervals. The seven pictures where the mean rating was significant after Holm's adjustment (Holm, 1979) are indicated with a red asterisk. A hyperlink to the DM Digital Archive with reproductions of all the images used in this experiment is given in Appendix B.

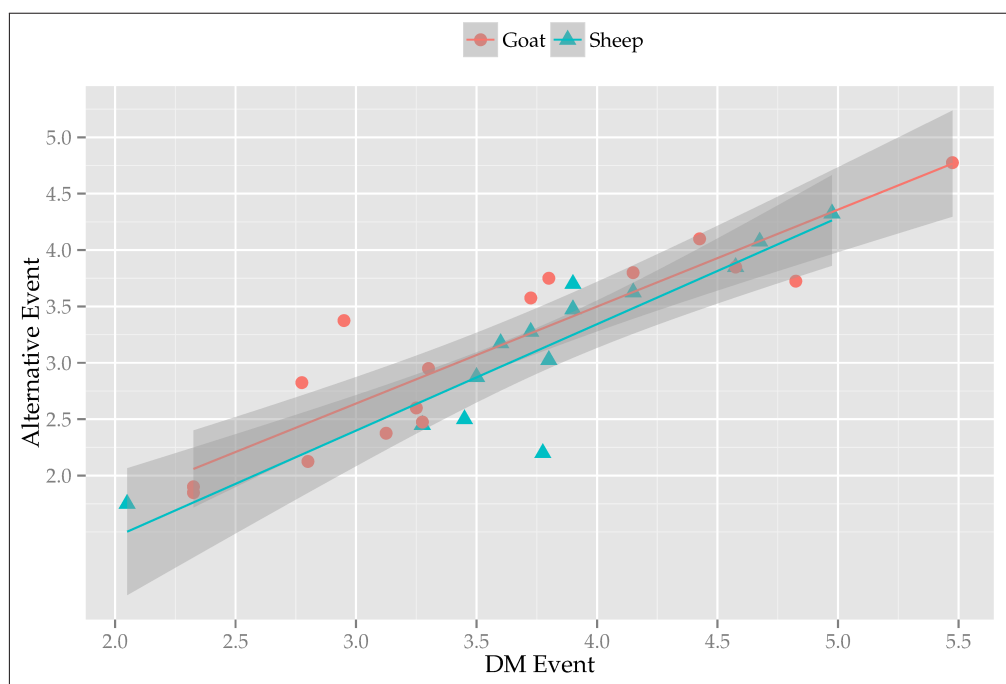


Figure 5.3: Scatterplot showing correlation between ratings of DM's event and the alternative event, with a median split designating the lowest scores on the Australian Sheep-Goat scale as "goats" and the remaining participants as "sheep". An individual regression line is given for both the sheep and the goats and both lines have a greyed area which denotes the 95% confidence intervals.

in Appendix B.3). We can see from these results that both images pertaining to the Twin Towers have impressed the judges, along with a terrorist attack on Heathrow, an earthquake, an environmental accident, a mining accident and a violent knife attack.¹¹ Importantly, only one of the significant differences is for a picture with supplementary evidence (Picture 7, "Heathrow mortar attack") suggesting that the presence of additional supplementary material did not sway ratings in favour of DM. Furthermore, it is interesting to note that one of DM's most specific predictions does not appear on the list. The "Concorde crash" picture, though described as one of DM's best predictions was not rated by independent judges as significantly different from the alternative.

In regards to secondary analyses, interestingly both rating scores (DM's and ALT) were significantly correlated ($r = .89$). Figure 5.3 shows a median belief split for the sheep-goat scores, and thus shows the mean rating for DM's event correlated for the mean rating for the ALT event with an indication of whether the individual is a sheep or a goat. This analysis appears to show that when a given participant has given DM's event a certain rating, the ALT is given a similar rating.

One final secondary analysis involved looking at the possible influence of both belief in the paranormal (using the sheep-goat scale) and the possible influence of set order (where the two sets A and B presented the order of stimuli in two different ways). A $2 \times 2 \times 2$ mixed ANOVA (employing type-III sum of squares) was conducted using the "ez" package of R. The dependent

¹¹Although merely a subjective opinion, I am confident that were we to run this experiment today, nearly one decade since the original judging took place, events could be found that would provide a better match than the ALT events employed. This is due to both improvements in technology and an increase in the number of events to choose from.

variable was each participant's mean rating for the event. There was one within-subjects factor of event (DM versus ALT) and two between-subjects factors of set (A versus B) and belief (sheep versus goat). The results of the ANOVA were, there was no main effect of belief (sheep versus goat), $F(1, 26) = .3, p = ns$ and no main effect of set (A versus B), $F(1, 26) = .1, p = ns$. There *was* a main effect of event (DM versus ALT), $F(1, 26) = 56.7, p < .0001$ (the mean rating for DM's event was 3.68 and for the ALT event was 3.15). Considering the interactions, there was no significant interaction between belief \times set, $F(1, 26) = .0, p = ns$, nor between belief \times event, $F(1, 26) = 2.8, p = ns$, nor set \times event, $F(1, 26) = .42, p = ns$. There was also no significant $2 \times 2 \times 2$ interaction between belief \times set \times event, $F(1, 26) = .2, p = ns$.

This analysis reveals that there is a small but highly significant difference between the ratings for DM's events and the ratings of the ALT event (the difference between the two ratings is .54, but the significance of that difference is $p < .0001$). Neither belief in the paranormal nor the set (that is, order of presentation of the events) appears to have influenced that rating. It appears safe to conclude that DM's event is favoured slightly, and statistically significantly, over the ALT, though in the light of the initial t-tests, it appears that this may merely reflect the advantage that DM has over the ALT in terms of having the first choice of events to select from etc. Of course, others might interpret this as evidence supporting DM's paranormal claims, though again, there appears to be a number of normal reasons why we find a slight preference for DM's event over the ALT (e.g. fact that these are DM's best pictures etc.).

5.1.4 Discussion

Out of more than 200 drawings, paintings and sketches that DM has created over his lifetime, 40 of his best pictures were presented to a small group of 30 participants, who compared DM's image with a brief account of the event that it was purported to match and a description of an alternative event, chosen because it also had some subjective similarity to the image. Of those 40, only seven of DM's events were rated as a significantly better match to the dream-based picture than an alternative. Thus, some 82.5 percent of the alternative events were rated by participants as being as good a match as DM's event. This finding alone suggests that the longer DM collects images and waits, the better the likelihood that a good or very good match will come along.

We can see from Figure 5.3 that there is a correlation between the two ratings: when a participant has felt that DM's event matches a picture, they have also rated the ALT event as matching the picture too. This makes sense. To put it another way, somebody who could see a connection between DM's event and the picture, could also see a similar level of connection between the alternative event and the picture. A further ANOVA reveals that there is a highly significant and yet small preference for DM's event over the ALT. Again this makes sense. DM had first pick of any event that he thought matched the picture. He therefore was likely to pick good matches. The ALT events were left with less good events to choose from, though in the light of the t-test results the ALTs are still as good as DM's event for the majority of pictures. Moreover, this slight preference for DM's event probably reflects the contribution of the Twin Towers images, for which it was impossible to provide an alternative event that was as good a match as DM's event.

Does this mean that DM did have a premonition of the Twin Towers collapse? As we have already mentioned, DM actually expected the event to be an earthquake something that no-doubt would be extremely destructive were it to occur in NYC. Thus DM has multiple-outs.

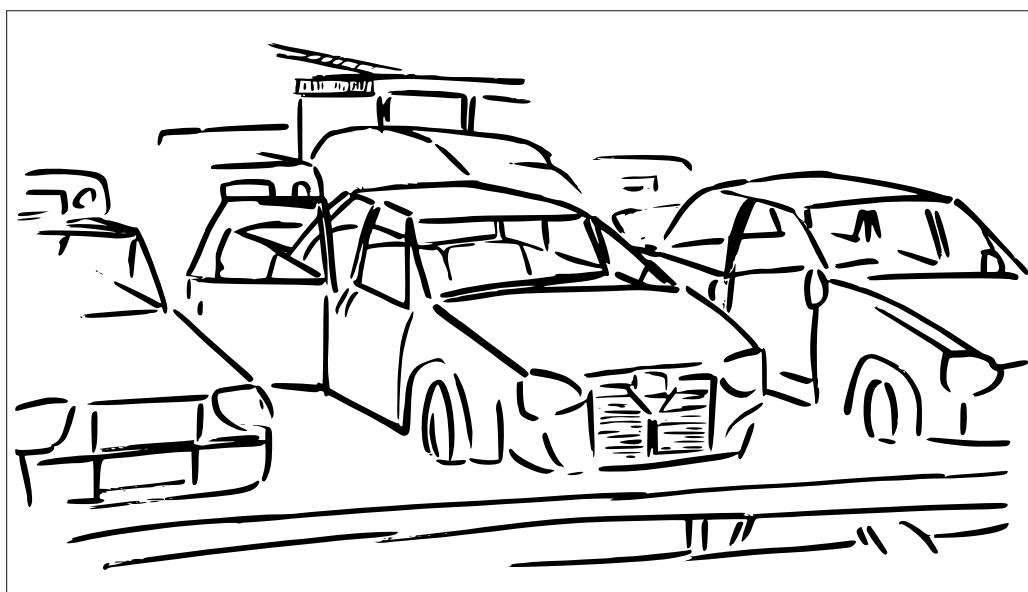
Had an earthquake occurred instead of a terrorist attack, this would also have been regarded as a significant correspondence. It seems likely that as Stevenson (1970a) argued in relation to assassinations (that they may be over-represented because of the prominence of the person involved and the propensity for the person to feature in dreams) that the Twin Towers were a prominent world landmark and were the target of a kind of political assassination. Although it may be argued that this slight preference for DM's events over the ALT may be taken by some as support for DM's precognitive ability, we do not favour that interpretation, given that the preference is small (though highly statistically significant) and yet may possibly be explained by other normal reasons (e.g. participant demand characteristics). It should also be noted that the 40 pictures used in this study were DM's *best* pictures. Were we to have included all 200 or so pictures that DM produced throughout his lifetime, we might expect this slight preference for DM's event over the ALT to reduce.

As also argued, the fact that the date of the Twin Towers premonition was 5 years to the day before the disaster appears to be purely coincidental (though one which is impressive and unlikely, nonetheless), given our conclusions about the nonparanormal basis of the claim. We can back up this argument with a statistical assessment. Since there are two Twin Towers pictures, only one of which was produced on the same day 5 years before the event, the chance of that occurring are $\frac{2}{365}$, which is less than $p = .05$. However, since none of the other pictures used in this study were produced on a comparable anniversary date, we can calculate the chance that one of the pictures used would have been produced on the "correct" day of the year as $\frac{40}{365}$, which is not significantly below the .05 α level. This then supports our contention that the surprising date claim for one of the Twin Towers pictures is the product of coincidence.

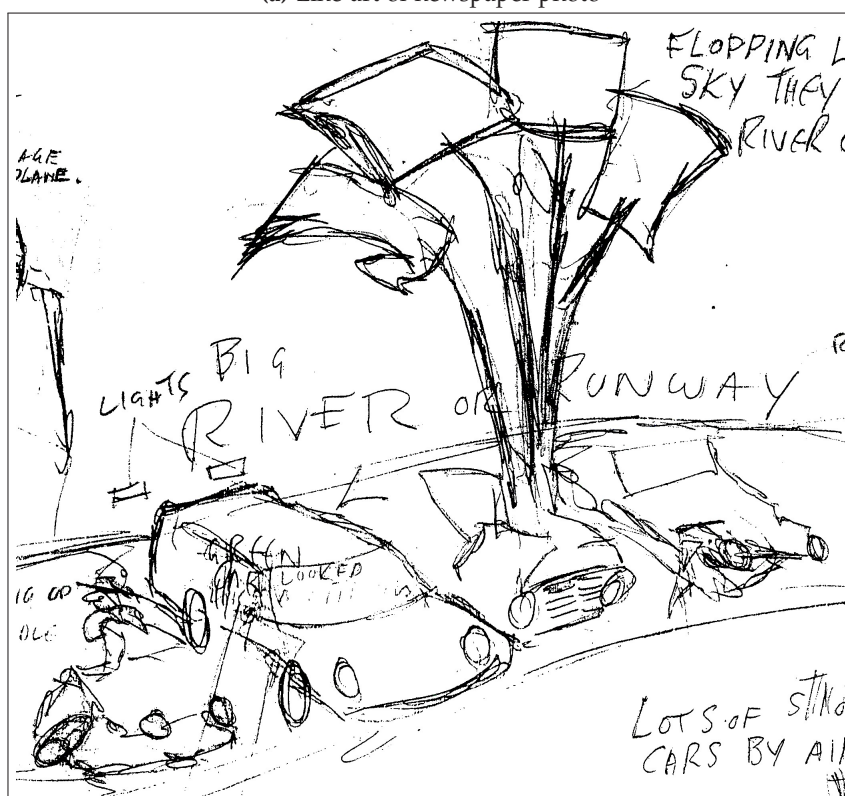
Interestingly, only one of the pictures presented with supplementary material in the test booklet had a DM event that was rated significantly higher than the ALT event. This is the "Heathrow mortar attack" picture, which does share an interesting similarity between the picture and the accompanying newspaper clipping (see Figure 5.4). It seems likely that the explanation for this seeming correspondence is chance coincidence and no further investigation of this correspondence was conducted.

A criticism of note is the fact that the only information participants received about the different events was the information provided in the booklet. Perhaps the mid-scale ratings and non-significant differences between event ratings reflects an ignorance on the part of the participant. Not knowing much about the event perhaps they found it difficult to find the connections, where DM had often learnt about events in the newspapers, or television news programmes, which are often richer sources of information and thus participants were not provided with a means of differentiating between the two events. The fact that only one picture with supplementary evidence was amongst the seven significant differences suggests that participants were not swayed by the extra material.

It also might be criticised that we are understating the significant difference between DM's event and the ALT, mainly because where DM's events are within 10 years of the picture being produced, the ALT event is not restricted to this time frame, and in fact has no restriction regarding when the event occurred (be it before or after the dream was recorded). Thus it might be suggested that the ALT has an advantage over DM's event. Given DM's actual claims, and lack of amenability to empirically testing them, it appears that this is a difficult issue to resolve. However it should be recognised that, of the 7 pictures where there is a significant difference between the two ratings, a good number of the pictures are abstract and indistinct. Perhaps future research should employ a different methodology (asking participants to match



(a) Line-art of newspaper photo



(b) DM sketch

Figure 5.4: (a) shows a line drawing of the photograph that appeared on the front of the *Independent* newspaper accompanying a report of the Heathrow mortar attack (10/03/1994). The quality of the newspaper image makes it difficult to reproduce in print, although a copy can be found in the DM Digital Archive. It should be recognised that the photograph shows the cars in a burnt out state. (b) shows DM's sketch (produced in 1991) which he believed was a premonition of the Heathrow mortar attack. The full image can be found in Appendix B.1. It can be seen that there is some correspondence between DM's picture and the photograph. We presume that this correspondence is coincidental.

picture with event) in an attempt to allay this particular concern.

Finally, it is clear from the results that set order (which gave the events in a randomised order, including the presentation order for DM and ALT events) and sheep-goat score has not influenced the ratings in any systematic way. Although surprising given past research (Schmeidler, 1994), it is a sign that participants attempted the judging task in a fair and unbiased way.

What we can say is that before testing, DM had an advantage in that he had identified a number of events which he felt matched the dreams. Finding alternative matches was not particularly difficult, and given greater time and repeating the research today (with greater access to online news archives etc.) it is reasonable to suppose that the difference between the ratings for DM and ALT would be reduced. The correlation between the two ratings shows that given a choice of DM's event and the alternative event that the ratings are linked—a strong rating for DM is matched with a strong rating for the ALT—and this undermines DM's claim that the dreams correspond solely with a unique event in time. Given the lack of experimental amenability, no further test of DM was sought.

On initial examination DM's claim is very impressive. It is a simple fact that the bank photograph of DM standing with his picture of a disaster occurring to the Twin Towers, dated five years to the day before the Twin Towers *were* involved in one of the most important events of the first decade of the 21st century, is a good example of a coincidence and would be noteworthy without the additional paranormal claim.

The main conclusion of our testing of DM is that DM does not have precognitive dreams of the future, but through the process of *subjective validation*, he has been selectively scouring news sources which might match the great variety of dream based imagery that he has accumulated over a number of years, and that, by chance *alone* he has happened upon events, with some regularity, which seem to him to be a good match for his dream imagery, thus validating in his experience, that he truly does have premonitions of the future. In truth, were DM to collect enough pictures and wait long enough, he would, by chance, be expected to find any number of correspondences between his dream imagery and events that occur in the future.¹² Marks and Kammann (1980, p. 182) argue that subjective validation is one of the key components of reported experiences such as DM's: "The rule is simply—keep searching for similarities until an overall match has been made. . . Once the match is presented it will be hard to see how it could be any other way." Obviously, this is a kind of hindsight bias (Fischhoff, 1975).

As well as subjective validation, there is another potential aspect to this research that has been effectively ignored by all involved. If we examine the so-called "Rackhams knife attack" picture, that seemingly shows a man throwing boulders off a building, an internet search reveals that along with a small number of potential news stories involving boulders and murder, that boulders and murder were part of a plot-line for a popular murder-mystery television drama, shown in the UK. Thus it is entirely possible that the preoccupation with finding newspaper events that match the dreams, led to the ignoring of the possibility that DM was influenced by non-factual sources. To support this conjecture, it is worth noting that a spin-off of the popular (now defunct) television drama, *The X-Files*, recorded an episode where the plot-line involved a jet-liner being flown by terrorists into the Twin Towers, broadcast in March 2001. Although some have inevitably taken this as a sign of premonitions and

¹²Westerlund, Parker, Dalkvist, and Hadlaczky (2006) come to a very similar conclusion in regards to matches between mentation and targets in the Ganzfeld experiment. Having worked on a Ganzfeld project (Smith & Savva, 2004a, 2004b), I can confirm (anecdotally at least) that striking and unexpected matches do occur, though when they are considered in the light of the understanding of the occurrences of such coincidences they do not remain remarkable.

conspiracy, this current research sheds light on the fact that unusual coincidences do occur.

Thus, there are three normal explanations for what DM has experienced:

1. The match between dream and event is entirely coincidental.
2. The match between dream and event is a consequence of normal influences, such as television news reports and popular dramas, and relates to potential or normal inferences regarding what might happen in the future. For example, talking heads may for weeks discuss the possibility of a terrorist attack, before one occurs.
3. The match between dream and event is the result of a repeat event misidentified as a unique event. For example, a plane crash that is in the news for several weeks is forgotten, later dreamed about, and this event is the seed stimulus for a dream, which is then identified as precognitive when another, similar event occurs.¹³

Whilst the scope of the current investigation was limited by the restrictions imposed by DM and thus we cannot say to what extent these influences played a role in DM's identification of his dreams as precognitive, one particular example provides a good illustration of the way that a previous event could well have seeded DM's dream imagery, such that when a similar event subsequently occurred, DM believed his dream was a reflection of the future event, and not the past event. Note that this is not the primary conclusion regarding the mechanism behind DM's paranormal attribution, which we believe is that any match is entirely coincidental and the precognitive attribution a consequence of subjective validation. However the fact that the "Concorde Crash" picture does not appear in the list of significantly rated images, does raise the issue of why a seemingly "perfect" match, was not rated as such by independent judges. Of course, others may interpret some of the results (the slight, but highly statistically significant difference in the ratings of DM's and the ALT event) as supporting DM's initial claim. Without more definitive evidence (stronger statistical results favouring DM) or the results from a direct experimental test of DM's abilities, we prefer to attribute the findings to purely normal psychological mechanisms.

Koncordski?

Whilst preparing the alternative news stories for the study described, an event was discovered that matched the picture that DM had produced, if not better, then at least as well as the event that DM had identified. Although essentially anecdotal in nature, it clearly illustrates the difficulty in evaluating any particular spontaneous claim of precognition.

DM produced the picture which he identified as *Concorde* crashing, on 20th March 1997. The picture was accompanied by corroborating photographic evidence taken in the bank, as outlined earlier. At the time the picture was produced, DM assumed it was a premonition of the crash of *Concorde* at some future date. Examination of the image (reproduced as Figure 5.1) seems to depict a triangular looking aeroplane, flying into a dark patch, with lots of flames or a jet of fire coming out from underneath the plane's tail. At the bottom right-hand corner of the picture is a figure of a man, seemingly sat in a small triangular plane (over tree tops) and flying a French flag. There is some text written on the picture, much of which is not clear enough to distinguish what it says, although the words "Concorde, terror, torture, crash, French flag and glider" can be made out. This cursory examination seemingly justifies DM's interpretation

¹³To be more explicit, in the year 2013 a large number of Boeing 787 Dreamliner planes were grounded because of defective lithium-ion batteries. Such batteries could cause a fatal fire on board and a prediction of such an accident caused by the defective batteries could be entirely inferential.

that the picture represents *Concorde* crashing and that the event is associated with a French plane or glider.

On the 25th of July 2000, *Concorde* crashed, with the loss of 113 lives, just after take-off from a French airport. At the time of the crash, *Concorde* was deemed one of the safest planes on record, having not had a single fatal accident since it first flew in 1969. On hearing of the crash involving *Concorde*, DM made the association between the details of the *Concorde* crash and his picture painted three-years earlier. Certainly the idea of *Concorde* crashing in France seems a reasonable interpretation of this particular painting and were it not for the disparity in the time between picture being painted and the time of *Concorde's* crash, there would be little mystery to explain. At first glance then, the *Concorde* example seems compelling evidence for some kind of precognitive ability on the part of DM. *Concorde* had never crashed before and here was a prediction, three years before the crash, and hence far back enough in time to rule out any normal or ESP (non-precognitive) inferential process, which seems to show ostensible foreknowledge of the event.

Although a number of DM's pictures are of a general nature, and hence could match more easily with any number of plane crashes, train crashes and ship sinkings, some are seemingly far more specific and hence far less likely to be the product of pure coincidence. Surprisingly, in 1973 the Tu-144, a Russian supersonic jet, and rival to *Concorde* crashed at the very same airport that the pilot of the 2000 crash had been attempting to reach. A Russian supersonic jet crashing in France, killing 13, was widely reported in the European press and for a brief period *Concordski* or *Koncordski* made the headlines of the day. Zarakhovich (2000) reporting in *Time* magazine, notes that at the time of the Concordski crash there were rumours of a botched espionage attempt by the French, and that eventually the French government admitted to spying on the Russians with a Mirage III fighter plane, sent to photograph the Tu-144.

In comparing the two events we can identify a number of similarities. The Tu-144 looked very similar to *Concorde* and was an unusual triangular supersonic plane. Both planes crashed in France and in close proximity to each other—in space if not in time. However there are a number of elements which suggest that the crash of Tu-144 was the actual event which inspired the picture. Firstly on the picture itself, is the man sitting in the small plane (glider is written nearby) flying a French flag. In the *Concorde* interpretation this merely signifies the fact that *Concorde* crashed in France. However in the account of the crashing of Tu-144, a small French fighter-plane was scrambled to observe the flight and somehow ended up precipitating the ensuing crash. The fact that a small plane flying a French flag is an important element of the picture, certainly suggests that it is an important element in the story behind the picture. *Concorde* crashed because of debris on the runway, which caused damage to the plane and the subsequent accident.

There is one other suggestive piece of evidence which points to the crash of the Tu-144 as the inspiration for the picture. DM's painting shows a source of fire emitted from directly under the tail of the plane and appears to represent part of the aeroplane's propulsion system. Examination of photographs of both planes and the specifications, suggest that the configuration depicted in the picture is a better match for the Tu-144 than *Concorde*.¹⁴ The engines of the Tu-144 lay much closer to the tail, and thus the centre of the plane, where *Concorde's* engines are further apart and appear as though they are more on the wing. Thus it seems highly likely that DM's picture of *Concorde* crashing incorporates details from the earlier

¹⁴A photograph of the Tu-144 can be found at http://upload.wikimedia.org/wikipedia/commons/f/f1/NASA_107665main_tu-144_drawing.png and a corresponding photograph of *Concorde* can be found at [http://en.wikipedia.org/wiki/File:Concorde.planview. arp.jpg](http://en.wikipedia.org/wiki/File:Concorde.planview arp.jpg) for comparison.

crashing of the Tu-144 and more than likely is based on that event.¹⁵

Thus, one of DM's best premonitory pictures, which on first examination appears to provide strong evidence of precognitive foreknowledge of a disaster, of which DM could not have known anything through normal channels of information, after examination, appears to be a much less difficult to explain experience, from a normal perspective. This is reflected in the "Concorde crash" image not being one of the pictures where DM's event was rated significantly better than the ALT. This case provides first-hand experience of an obviously non-fraudulent claimant, who has been persuaded of their own psychic abilities through the uncanny experiences that they have experienced. In effect, this is the same cognitive attribution process as the lottery winner who believes they got lucky, even though somebody has to win the lottery.

Relating DM's experiences to some of the work covered in the thesis introductory chapter, Saltmarsh (1934) defined some often cited criteria for a perfect precognition and at first glance DM's claims certainly appear comparable to the most impressive reports. Firstly the premonition must have been told or recorded before fulfilment or acted upon in a way to suggest foreknowledge. DM developed a commendable regime of taking the pictures to a bank, the day after the picture was completed, to obtain a photographic record with the date and a copy of the picture. Whether this particular protocol is sufficient to allay the concern that the picture was produced after the event which it is claimed it matches is difficult to assess. Certainly photographic images are not resistant to doctoring and using very limited computing resources one could potentially replace, add or otherwise change a photo to manipulate and deceive. However given the reports from bank staff and DM's motivations, this particular concern is minimal. Secondly, the premonition should include enough details so that coincidence is rendered unlikely. None of the premonitory pictures provided by DM contained conclusive evidence of a foreknowledge of the event and many were in fact vague and required a great deal of interpretation.

Thirdly, Saltmarsh (1934) required that the perfect precognition specify narrow limits for the time of fulfilment and none of DM's predictions included specific time frames for an event's happening. This weakens the overall claim, given that over a long enough period of time any unlikely event becomes more likely to happen.¹⁶ Fourthly, Saltmarsh required that inference could not explain the foreknowledge and here Saltmarsh is concerned with normal inference. Is there any evidence that DM's claims are based on inference? The case can certainly be made that given the previous attacks on the Twin Towers in New York, a terrorist attack on such prominent landmarks might not have been too fantastical to predict. As such then, DM's pictures fail to provide compelling evidence of a paranormal precognitive phenomenon. Given the specificity of his claim, his genuine willingness to be investigated, and his wide range of material provided, make DM an important case within this thesis. DM's claims of precognitive dreams, accompanied by photographic evidence of the picture's existence before the event and the specificity of the dreams, meant at first glance DM appears to be able to provide the elusive evidence for precognition. Certainly the findings of this study cannot rule out this eventuality altogether. However with all of the details of the paranormal claims laid bare (with each event that DM felt matched his picture twinned with another event that it should not have

¹⁵Of course, it is entirely possible that given that the source of the imagery is a dream, it is based on an unconscious confabulation and not on any event past or present that has occurred in the real world. However, it should also be noted that DM claimed to have worked as an aeronautical designer, and as such, may well have been acquainted with the different designs of the two supersonic aircraft. It is regretful that DM was not questioned fully on this matter.

¹⁶Ford (1963, p. 122), "Anything that can happen, does happen."

matched) it became apparent that the initial optimism of obtaining evidence for an ostensible precognitive ability diminished. Instead we are left with an example of how something which looks very much like an impressive paranormal claim can be explained just as well, when restricted to looking for normal explanations.

Strange and uncanny experiences occur. Interestingly, Richard Kammann recounts a very similar experience to that of the mother precognizing her son drowning in a bath (p. 21), this time dubbed 'The Case of the Drowning Daughter' (Marks & Kammann, 1980, pp. 156–157) and, based upon that, develops a non-psychic model of paranormal belief. Briefly, Kammann coins the term "oddmatches" for those occasions when two seemingly unrelated events occur in such a way as to be identified as psychologically connected, such as the aforementioned dreams about children drowning and then the subsequent near-drowning event. The second part of the model, called "Koestler's fallacy", notes that there is a constant and rapid flow of events every day of a person's life, which leads to numerous chances for coincidences and, noticed or unnoticed, coincidences occur.¹⁷ That which seems unusual, with research appears not quite so unusual.

Is it possible to quantify such experiences, rather than merely state that such experiences occur? Alvarez (1965) reports an experience he had where he remembered a known person and then five minutes *later* learnt of that person's death and gives an estimate of 3×10^{-7} that such a thing will occur by chance. In fact he makes the pointed observation that given the population size it is not surprising that such seemingly unlikely experiences are reported. Zusne and Jones (1989) also calculate the chance of thinking about somebody we have known in the five minutes before learning of their death, and give it as one in 30,000 per year, which given the number of humans on the planet suggests that it should not be an unusual event. Interestingly, Flammarion (1922) gives a similar calculation (p. 337) for a case where one man claimed to have seen an apparition of a dying friend just *before* the moment of his [the friend's] death. He gives an estimate of the chance occurrence of such a coincidence as being $\frac{1}{804,622,222}$ (or 1.24×10^{-9}) which makes it a much less likely occurrence than other estimates.¹⁸ It seems likely though that, through purely normal mechanisms, scientists should expect people to report experiences such as DM's and should be able to refer those who feel compelled to explain their experience within a paranormal context to the appreciation of this law of large numbers.

Finally, one last consideration: What if DM were a hoaxer (e.g. Hansen, 1990)? Although no evidence was found to establish a motive of fraud, there are two papers in the literature which point to a mechanism by which strong premonitions may be made in the absence of paranormal powers. Kammann (1983) writing in the *Zetetic Scholar*, gives an account of four prophecies made on a New Zealand radio show, by "psychic" Emory Royce. The predictions are remarkably close to the kind of predictions made by DM, although this may be because the more dramatic events and premonitions are of more interest to us. The four predictions were:

1. The death of the Soviet president, Brezhnev.
2. The sinking of two ships
3. Cancellation of a major construction project.

¹⁷Kammann gives two other related phenomena. The clustering illusion is where something becomes psychologically prominent and is then seen as more frequent than had been previously noticed. An example Kammann gives is a medical doctor who encounters an unrelated cluster of patients suffering from a rare disease. The other phenomenon is the Gremlin illusion, which relates to the expectation that bad things happen at inopportune moments (think Murphy's law). An example might be the telephone ringing or the door-bell sounding, when in the bath.

¹⁸Flammarion concludes (p. 338) 'The apparition of the deceased to his friend is not to be doubted.'

4. Mid-year scandal involving the New Zealand prime minister.

Kammann writes:

It is noteworthy, I think, that these four predictions are not selected out of a much larger list of failed prophecies, but are the total set of testable predictions involved. It is regrettably impossible to calculate an exact significance level or "chance probability" of four events occurring with a goodness-of-fit equal to, or better than, these four actual events, but a subjective estimate places that probability well beyond the .0001 level, and some might go much further. (Kammann, 1983, p. 20)

Readers of this article are certainly persuaded that something decidedly unusual had occurred. A follow-up article appeared quite a number of years later (Kammann, 1987) and revealed that, although the predictions themselves had been 'genuine', that Emory Royce was none other than Richard Kammann himself, and that the study had been an attempt to incorporate the *delayed control group* into a demonstration of precognition, that would illustrate that normal mechanisms can be used to produce subjectively impressive premonitions. What methods had Kammann used to produce his predictions? The main technique involved using chance or coincidence and the "minor" methods (inside information, trickery, and data-selection) were deemed to have made little overall difference to the final prediction. Kammann revealed how he made his predictions:

I was having a battle in the newspapers with a local astrologer who believed I had pulled off a fraudulent stunt on a TV program. . . The day after this astrologer gave a rather dull radio interview, I was invited to the same studio to comment on his performance. The night before, I thought it might be fun to offer some generalised predictions in the style of Jeanne Dixon and such folk, and I cooked up a few prophecies off the top of my head. It was my hope that if *even one* of these prophecies found a matching world event, I could later claim a success. I was also confident that if nothing happened, the radio folks would forget all about it, so there was no threat of embarrassment.

. . . The idea that I predicted a "naval disaster" and the sinking of two ships is nonsense. I was mainly thinking of the Three Mile Island nuclear accident which I free associated, "man-made disaster. . . structure bathed in or surrounded by water. . . poisonous gasses spreading some distance and seriously affecting life. . . possibly nuclear war. . . war scare." This I felt could also cover major oil spills, collisions at sea, and other situations only vaguely visualised . The "war scare" covered all military variations but could easily be dropped out. (Kammann, 1987, p. 35)

Kammann's methods for the other predictions are similar to this. Kammann's conclusions are clear:

Competent students of the paranormal are well aware of the combined effect of ambiguity, coincidence, and data selection in the manufacture of alleged miracles of precognition. . . Skeptics might be intellectually satisfied that alleged miracles of prophecy can be explained away in terms of subjective validation, coincidence and data selection, but it is also necessary to go the extra step and demonstrate that these factors are adequate to produce a few "equivalent miracles" for the world to inspect. (Kammann, 1987, p. 37)

Thus, our work with DM helps provide a similar illustration of the complex psychological mechanisms that are involved in the construction and maintenance of the belief that what one has experienced is paranormal foreknowledge of the future. Richard Kammann and Uri Geller:

two polar opposites, though both undoubtedly 'knowing' frauds. DM sits between the two; sincere yet most likely erroneous in his attributions.

It is without doubt that cases such as that of DM, will continue to be reported in the literature. Cardena, Iribas, and Reijman (2012) in a paper on art and psi, describe the fact that a famous Romanian artist painted a number of portraits in which one of his eyes was damaged or lost, only for one of his eyes to be actually lost (due to a completely unforeseen event) some years later. Although we may favour the chance coincidence explanation, ultimately dramatic cases do suggest the possibility that precognition may occur.

It is fitting that as the last study in this thesis using a human participant, so much has been discovered about the real nature of people's paranormal experiences. As the spontaneous cases are such a driving force in encouraging continued investigation of paranormal claims, it is fortuitous to have had the opportunity to investigate so impressive a claim in quite such detail. The overall conclusion that normal psychological processes can explain all of the fantastical elements in DM's story is a now familiar conclusion but one from a more naturalistic (or human) perspective. Most of the participants in parapsychology experiments do not claim to dream hundreds of premonitions and nor do they attempt to obtain proof that their dreams really are about the future. Yet despite these fantastical claims, DM's beliefs cannot be accepted as evidence for ostensible precognition in light of the evidence.

Insect Death-Avoidance

6.1 Anpsi test and Selective Breeding with Laboratory Populations of *Tribolium*

6.1.1 Introduction

None of the results obtained in the research undertaken for this thesis have required recourse to the paranormal hypothesis of precognition. Despite the intriguing spontaneous case collections which certainly show that people have odd experiences, empirical testing has failed to confirm the parapsychological claims thus investigated. Moreover, testing in the last chapter demonstrated that even a seemingly impressive claim of precognition can weaken under scrutiny. The secondary aim of this thesis was to investigate the effectiveness of employing fear-based methodology, to test the hypothesis that psi experiences are mediated by some kind of evolved death-avoidance and this also has not been supported. Replacing conventional stimuli with fear-based stimuli has proved no more successful at eliciting any kind of paranormal effect, than those usually employed. The implications of evolution to the investigation of parapsychology were discussed in Chapter 1. No evidence has so far been found and there is no support for conceiving of precognition as a pre-sense. However, despite the null findings in the previous research, there was one last experiment which would be a logical continuation of the previous research and represent a last, best test of the pre-sense theory of psi (within the confines of this regime).

The ideas presented in William Cox's papers (Cox, 1956a, 1956b) are some of the most important to this thesis, as they were influential in developing the evolutionary death-avoidance ideas.¹ Cox first reviewed the precognition literature, specifically the spontaneous case reports, and concluded that there seems to be some suggestion of an ability to avoid a future fatal event, which is a similar conclusion to that outlined in the first chapter, after a similar review. Cox (p. 99) was interested in whether: 'In advance of a sudden misfortune, a wholly subliminal precognizance can prevent the percipients' involvement.' Ultimately this is the question that lies at the heart of this thesis. Is there a need for a paranormal explanation for those times when people seem to precognise the future? Cox argued that there is seemingly evidence for an unconscious precognitive ability to be at the right place at the right time, or to avoid being at the wrong place at the wrong time. This is not an unusual conclusion and has been discussed a number of times within this thesis. Cox then presented his evaluation of train

¹Gardner (1998) notes that if Cox is remembered at all, it is for his efforts at proving Uri Geller genuine, or as Gardner (1981, p. 143) calls him, a "Geller-gawker squawker."

crash data. He argued that train crash data are very amenable to the purpose of investigating a paranormal effect, whereas ship sinkings are too infrequent and the demand exceeds capacity. Bus and tube passenger details were not regularly kept at the time. Only the railways recorded data that could be analysed by Cox for the purpose of investigating this question.

Cox (1956a) approached 35 railroads, out of a potential 131 in the U.S. and Canada, and obtained their crash records, which consisted of some 28 crash incidents. Cox reported finding a small but significant tendency for there to be *fewer* people in a train carriage on the day of a crash, than on other days preceding the accident. This difference was less marked in the so-called Pullman carriages (where people were more likely to book ahead) than in the coach carriages (where foot-passengers were more likely to purchase a ticket at the time of travel). This difference was seen by Cox as evidence for unconscious psi, where the precognitive experience is felt and behaviour adjusted without conscious input. Because Pullman passengers had booked in advance, Cox believed that they were less dissuaded by feelings of impending doom, where the coach travellers were more able to be unconsciously saved from getting in a carriage that was very soon to be involved in an accident. To put it another way, it was nothing for a casual traveller to get a later train, but someone who had booked in advance may have had plans that they were unwilling to change, even if they experienced excessive feelings of impending disaster. Cox identified this disparity in passenger numbers as evidence of precognitive death-avoidance. Cox raised the question that, if this is what the data represent, why is the engineer in control of the train not the one to have the precognition and therefore prevent all of the deaths, to which he argued that it may be one of those times when the engineer's psi fails him.

With the tentative conclusions of his review of the train data, Cox (1956a) proposed an experiment which might provide an empirical test of this particular hypothesised ability (herein this will be called death-avoidance). Cox suggested that, assuming a psi ability in insects,

[O]ne quarter or some other fraction of an enclosed surface upon which a certain number of ants, e.g. freely roam, and could be designed to electrocute such as are not traversing elsewhere at that instant. The quadrant and instant would have to be randomly selected. (Cox, 1956a, p. 108)

This proposed experiment would be a test of animal psi, otherwise referred to as *anpsi* (not ant-psi). Obviously such an experiment would fit in well with the aims of this thesis. There are certainly a number of criticisms that can be levelled at Cox's review of the train crash data (Cox, 1956a). The small number of actual cases that he analyses prevents any real assessment of the crash data as a whole. Cox does not really acknowledge the possibility that his data set could in some way be artefactually distorted by any number of normal influences which could have affected passenger train travel. Certainly somebody conducting a similar review today for example on the London Underground, would have to take into account any perceived terrorist attack threat as a possibility for a deviation in passenger numbers. Perhaps a less contentious influence might be the weather. However, the hypothetical experiment Cox outlined certainly has a number of positive points which make it a candidate for a final test of both the paranormal hypothesis and the idea that psi, if real, might be a pre-sense. By removing the human element and replacing the participants with animal subjects, Cox removes all of the demand characteristics and potential confounding factors which accompany human research. It is as if Cox has been inspired to construct the perfect precognition experiment, whereby trains are randomly crashed, and to see whether the numbers of passengers in the cars that crash are different to those that do not (which is what Cox claimed to have found

in his review), except that the experiment can never use humans. Obviously, there is no way to empirically test death-avoidance in humans in an ethical way, however, it is certainly an experiment which is worthy of study in animals, if for no other reason than a positive result would support some of the ideas which have been put forward as to the purpose and mechanisms behind precognition.²

Some researchers have attempted to conduct a Cox (1956a) like study, although with mixed results and some of this evidence was covered in Chapter 1. Morris (1967) conducted a ground-breaking experiment in which the behaviour of an animal was monitored over a period of time and then used as an indicator as to whether or not the animal anticipated through psi, what was about to happen to it (which for half of them was to be killed). In the first study, 19 rats were inserted one at a time into an open field maze and a count was taken of the number of floor squares traversed by each rat, during a two minute interval. Following this, a second experimenter either killed the rat or let it live, according to whether an odd or an even digit had been assigned to the rat by a random number table in the possession of this second experimenter. Thus the procedure was double blind. The expectation was that rats about to die would show less activity than those that would live, because rats have been shown to 'freeze' in the presence of danger. Results were not significant.

On the face of it, Morris (1967) appears to be a methodologically sound test to investigate death-avoidance. However the research also raises a number of problems that may have influenced the results. First, the rationale for predicting what type of behaviour to expect before an animal is going to die is, if not tenuous, difficult to argue definitively. Rats may freeze in dangerous situations, but does that mean that precognitive detection of danger will cause the rat to freeze? Inaction seems to be the antithesis of intervention (something which appears to be important for precognitive experiences). So perhaps using it as an index of response to death in the future is not recommended. There is another (perhaps more important) methodological problem which might account for the negative result. At the end of the testing period, the rats that had 'survived' the psi task were also killed. Broad (1968) argues vehemently against the notion of precognition because, he reasons, that for a person to precognise the future, they must be able to change or avert the fate that they have precognised. It is thus argued that those events which cannot be avoided cannot be precognised. Does this explain the null result? Was there no difference between those animals who were about to die and those who were going to die some weeks later because of the fact that they all would die soon and there was nothing that the animals could do to change that outcome? It is not like the *Titanic*, for example, because people had the opportunity to not get on the *Titanic* and then go about their lives. This appears to be a much more damning criticism and one that is only rectified by allowing the surviving rats to live out their natural life-span (or even release into the "wild"), something that is difficult to do from both an ethical and practical standpoint.

Craig and Treurniet (1973) attempted to replicate and extend Morris (1967) work. They used 72 rats and claimed to have found that rats were significantly more active when in an open field environment, for those individuals who were randomly killed a day or two later, versus those killed a minimum of three-weeks later. The fact that this was in the opposite direction to that predicted on the basis of Morris's work is countered with three possible explanations.

²Good (1980, p. 12) wrote, "It is sometimes claimed that telepathy, clairvoyance, and precognition are apt to occur in situations of great emotional significance. It is difficult or expensive or unethical to bring such situations into a controlled experiment. It might be possible by offering potential percipients immense rewards such as reprieves to prisoners on Death Row. Or, more unethically, immense punishment could be given for failure, as by kings in fairy stories. Idi Amin could have organized such experiments if he had been interested in the scientific study of the paranormal."

Craig and Treurniet argued that as the rats in the two different experiments had come from different populations, their reactions might have been different. They also suggested that there might be an experimenter effect mediating the different results (and the experimenter's psychic powers somehow have influenced the results). A further possible explanation involved the phases of the moon. The authors concluded that (p. 77), '...it is now clear that time of death does have an effect on an animal's behaviour'. Given that conclusion, the authors argued that their findings had implications for other researchers using animals, in case a precognitive death-latency or arousal effect might influence results (that is, conventional animal experiments that terminate their subjects must themselves be influenced by a precognitive response to death). This last point is very interesting. One could well argue that the very fact that animal researchers and animal handlers regularly do not report a difference before an animal dies, is very telling (though we would only expect such a response in animals dying by accident or malice/predation rather than disease, childbirth or old-age, where intervention is less able to change the outcome or even impossible).

Bestall (1962) investigated a behavioural element in a similar methodology with mice. Rather than examine behaviour in an open field, Bestall used a two-choice maze, where a single choice was made per mouse. Mice were run in sessions and after all mice had been run through the maze (and made their choices), one of the choices was randomly chosen as the target (using weather report information and random numbers—using the weather was an attempt to ensure this was a test of true precognition as influencing the weather is generally considered outside of the ability of mouse PK). Those mice that made the incorrect choice were killed six hours after making their choice. Those who were successful were given access to a female mouse, also six hours after making their choice. Bestall reported a weak but significant effect, although again it is not reported what became of the successful mice (or their progeny) in the long term. (On a methodological note, it has been argued that animals in psi-tests where a choice is offered must be pre-trained to expect equality in the outcome of the choices so that biases or patterns do not unduly influence choice selection by the animal, e.g. Morris, 1977.)

Johnson (1982) conducted a very similar experiment to that described by Cox using the single-celled organism, *Paramecium aurelia*. The aim of Johnson's experiment was to select for a precognitive ability and thereby produce a psychic strain, in the hope that such a creature would represent unequivocal proof for psi functioning (a biological equivalent of a permanent paranormal object, something that could on the one hand be readily scrutinised by all and on the other hand defies conventional explanation). Johnson argued that such a line of enquiry presupposes that single-celled organisms possess psi functioning and that it has a genetic basis. Johnson had three groups of cultures of the organism: a precognitive group, a psychokinesis group and a control group. (In the psychokinesis group Johnson was aware of the result of the trial before testing and claimed that any deviation from chance could be the result of this knowledge and the action of his PK.) Every 24 hours for 14 days he would test the groups. Each precognition trial involved dividing each group into two, taking samples from each container, counting the number in each sample and then randomly destroying one or the other. Johnson found no difference between the number of paramecia in the sample that survived and the container that was destroyed and thus concluded that there was no indication that anything paranormal was at work.

Johnson (1982) appears to be a reasonable test of death-avoidance and the null results are disappointing. However, yet again there are a number of methodological elements which could somehow be eliminating a psi-effect. At the end of the study all samples were destroyed, so as

with Morris (1967), there is a concern that death-avoidance was eliminated because of a lack of continuation for the samples. Again, this particular concern would have been eliminated if the paramecia were released into the "wild". However there is also some concern regarding the evaluation of sample size. To determine the number of paramecia in each container a pipette was used to extract some from their solution and they were examined under a microscope. Johnson did not discuss the fact that to some degree this effectively constitutes a minor psi task: if the proposed death-avoidance behaviour exists then perhaps it would exert itself on those paramecia within the area of exposure to the pipette. Those that were collected were killed to make counting possible. As such, the fact that there was no difference in population sizes as determined by the sampling may not have represented an accurate figure for the actual populations within the containers. Had Johnson used a non-destructive method of assessing the total number of individuals in each population (perhaps with some kind of opacity index), this concern would not be important. Without an accurate assessment of the size of the populations it is difficult to assess the degree to which the populations differed and this significantly undermines the conclusions.

There is then some mixed evidence from experiments that have used death as the ultimate target in a psi-test. Other research has used less extreme measures. For example, Schmidt (1970) rigged a random number generator to a heat lamp and found that it stayed on significantly longer when his pet cat was in the area of the lamp. Schmidt also reported a significant effect with tanks of brine shrimp, that were subjected to fewer shocks from an RNG controlled electrical source than expected by chance (though a follow-up study was not significant, Morris, 1977). Morris (1967) found that goldfish who were randomly chosen to be caught in a net and raised out of the water for a brief period were more active (as rated by a blind observer) than those not selected (this selection was achieved through throwing dice directly after the observation period). A follow-up study was non-significant (Morris, 1977).

Another related experiment was conducted by Duval and Montredon (1968a). The experiment involved testing mice in a small cage divided in half by a small partition, over which the mice could easily jump. Every minute the animal's position was automatically recorded and immediately after this observation an RNG was used to select one side or the other to receive an electric shock. If the mouse was on the side chosen it was shocked through the floor. Although the overall results were positive but non-significant, when trials were dropped because of stereotyped behaviour the results were significant. Obviously there is a difficulty with this type of methodology in that the mice may begin to anticipate the shocks and act in a non-random way (this is what we mean by stereotyped behaviour in this paradigm). A follow-up study (Duval & Montredon, 1968b) also reported similar results. There have been several independent replications of this particular methodology. For example, Eysenck (1975) tested rats with a similar procedure and found significant under-scoring when the animals received 0.1 milliamps of electrical shock for some 3 seconds. In a second series this was increased to 0.2 milliamps and no significant effect was found. Eysenck argued that this could have been because the weaker shock represented an attractive stimulus, but that the larger shock of the second series resulted in the painfulness outweighing the attractiveness of the stimuli.

Finally, there are a number of even more novel experiments which, if not directly related to death-avoidance, are certainly related in a more tangential way. For example, Backster (1968) was interested in investigating a hypothesised form of communication between living cells. He monitored electrical resistance on the surface of one leaf of three plants to see if they would

“respond” to the death of brine shrimp (killed some distance away). The testing procedure was fully automated and an RNG was used to randomly select a time period (one out of six) to kill the brine shrimp by dumping them into boiling water. Three blind judges were used to analyse the electrical activity and the results were that there was a greater amount of electrical activity during the killing period than during the non-killing periods. Again, others have attempted to replicate this paradigm, with mixed results (e.g. Johnson, 1972; Horowitz, Lewis, & Gasteiger, 1975).

Overall then, there is some mixed empirical evidence for non-human animals to respond in a seemingly adaptive way to both death and other extreme stimuli. Although it may seem strange that parapsychologists have not made anpsi testing the main paradigm by which they have attempted to establish the reality of the underlying phenomena, there is a simple reason for that not being the case. Simply put, for any significant anpsi result it can always be argued that the animal is not the source of the psi influence, but is merely the focus of it. The human experimenters or handlers are often aware of what type of effect they wish to find given their methodology (and even in blind experiments such information could be acquired through psi) and thus, in some quarters it appears that keeping the phenomena wholly within the species keeps the explanation of the results “simpler”. If human beings have been shown to use psi in other regimes of testing (and most parapsychologists believe the case for human psi has already been made) then it is perhaps superfluous to argue that the results of anpsi tests are the consequence of animal psi. This is not the argument favoured by this author. Instead, as outlined in Chapter 1, there is a range of evidence that suggests that psi may have a biological basis, and since our biology has inherited and shared characteristics, it would be surprising for non-human animals to show *no* psi (given our conception of psi as a kind of death-avoidance behaviour).

As mentioned earlier in this chapter, a death-avoidance experiment was thus seen as a last, best test of the paranormal hypothesis of precognition. Although previous research following similar lines appears to be on the whole mixed, there is nevertheless some possibility that the mixed findings in the area are a consequence of other factors and not because the phenomenon is not real. The final series of experiments in this thesis was based on Cox’s thought experiment (Cox, 1956a), but would attempt to improve on previous methodologies used in the area. Specifically, a number of normal factors were identified which might be important to consider, when investigating death-avoidance.

Firstly, the previous rat research had measured the behaviour of the animal as an index of whether it was in some way aware of its future state. The difficulty in using this type of methodology is that there is little basis for distinguishing between a precognitive effect over normal behaviour, at the time of testing. Specifically, the rats were either killed soon after their behaviour was assessed or sometime after, and that afterwards, the behaviour of the two groups was compared. It is certainly possible that this particular method is not sensitive enough to distinguish between animals responding to a given presentiment, if any, and those just going about their regular routine. Also letting an animal wander around a given area might result in some stereotyped behaviour being exhibited, which somehow inhibits or overrides any precognitive effect. When using an insect species, it became apparent that using a square testing chamber might result in most of the animals exhibiting stereotyped behaviour, whereby the animals find a corner and stay there. It might then be argued that this over riding behaviour had eliminated the precognitive effect. Also, if there are any other stimuli in the testing environment, such as a light source, strong sounds or even electro-magnetic fields, they

could have a systematic effect on the behaviour of the animals, and thus influence the results. With these considerations in mind, it was hoped that by making the testing environment as uniform as possible, it would reduce the influence of any stereotyped behaviour and thus increase the chance of discovering a psi-effect.

Also, instead of developing an index of the animal's behaviour over a specific time-frame, this experiment was much closer to the thought experiment proposed by Cox (1956a), which employed a psi task with a behavioural aspect, where ants would be randomly electrocuted depending on their position in the testing environment. The difficulty of using Cox's proposed methodology is that it leaves open the possibility that the ants could be responding to electrical changes in the testing environment, and hence what might look like a paranormal effect, would have a normal basis. In this study, the animals' location within the uniform environment would be recorded and then their location would determine success in a subsequent precognitive test. It was hypothesised that given a uniform environment, the animals would show a preference for locating themselves in the testing apparatus on the side that is subsequently not chosen to be killed, and that such an effect would represent a precognitive death-avoidance behaviour. The null hypothesis was that there is no such ability and, in the absence of systematic bias, the overall number of animals surviving and killed will be equal (not significantly different from 0.5). So, where the previous research in this thesis had made only an indirect attempt to investigate death-avoidance with the use of fear-related stimuli, this last experiment was a direct attempt to simulate real life situations, like the sinking of the *Titanic*, but with non-human animals.

After a short period of testing (35 individual ants) it was found that the species itself was not suited to the conditions of laboratory testing. Approximately 100 ants, *Lasius flavus* (common name, yellow meadow ants) had been purchased for the first series in the experiment but this was subsequently abandoned due to the death of most of the insects (an environment that allowed the ants to thrive was much harder to provide than initial research had suggested). As such what had been intended to be a more extensive test of the paradigm was reduced to a pilot study until the problems could be rectified. Initial testing with the 35 ants resulted in more insects surviving than dying (21 compared to 14, which represents a 60 percent survival rate), although of course the results were probably due to the small number of trials and stopping before full testing was completed. As the experiment had been designed with a much greater number of trials in mind, this ant pilot has no real bearing on the later results, other than to represent an illustration of the difficulty of conducting non-human research, even with a species as simple as an ant.³ It should be noted that such unintended excessive mortality has been reported in the conventional insect testing literature (e.g. Park, Mertz, & Petruszewicz, 1961), who found an alarming and significant extrinsic mortality that was not a consequence of competition. The cause of the problem was a sporozoan parasite, which required the researchers to develop a technique to eradicate the "scourge". Thus excessive mortality is a problem which can affect even the most experienced researchers working with insect species. The problem also highlights the fact that it is possible that differences reported in the earlier research could be influenced by the state of an organism's health, about which, nothing is reported. It seems reasonable that unhealthy organisms will respond differently to healthy ones and that applies to conventional behaviour as well as possible psi responses. As it can be difficult to ensure populations are healthy over potentially extended periods, this is a

³Darwin (1871, p. 54), "... the brain of an ant is one of the most marvellous atoms of matter in the world, perhaps more so than the brain of man."

factor to be considered for future investigators of anpsi.

Due to the unforeseen difficulties in maintaining the ant species in captivity, a hardier species was sought for further testing, although most other aspects of the experiment would remain the same. The need to find a new species to test led to the identification of *Tribolium castaneum* (common name, the red flour beetle) as a better alternative to the ants. The red flour beetle is a hardy graminivorous insect. In the wild it is usually found under rotting wood, loose bark or in the nests of birds and bees. They are also a major agricultural pest of grain stores around the globe and a widely studied laboratory model organism. The beetles thrive at room temperature with little required for populations to flourish: small amounts of flour and some fruit peels. They have a short development period (~30 days), have high fecundity (up to 10 eggs per day per female) and are easy to maintain (Sokoloff, 1974). Not only were the flour beetles hardier, but they also allowed a slight modification in the methodology. Johnson (1982) had attempted to selectively breed for a precognitive paramecium, by only allowing those individuals who were successful in the psi task to live, and hence reproduce. The red flour beetle reproduces sexually and has a time from egg to egg of approximately 30 days, thus it was decided that the survivors of each round of testing would be allowed to reproduce, and that subsequent rounds of testing would occur with the offspring of those beetles that made the successful decision in the psi task (such selective breeding would not have been possible with the ants, where most individuals are sterile siblings and it is the queen who is the mother to all in the nest). The flour beetle made this particular methodological aspect very easy because the beetles have a larval form which is easy to separate from the fully formed adults. The flour beetle has been used in conventional laboratory tests of selective breeding, where they have shown themselves to be a useful species (e.g. Wade, 1976; McCauley & Wade, 1981; Wade, Shuster, & Stevens, 1996).

The hypothesis of this final round of testing was thus: that given a uniform environment where there is no particular stimulus and that given the consequences of being on one half of the testing chamber or the other, the beetles will show a preference for the side which allows them to continue living and that this success will improve over successive generations. It was hoped to be able to run this particular experiment as long as possible, although in actuality some five generations were run.⁴

6.1.2 Method

6.1.2.1 Design

A simple between-subjects design was used comparing the number of animals that passed the precognition task, and hence survived, with the number that failed the task and were consequently killed. Although implicit in the design, a secondary objective was to examine the survival rates across generations.

6.1.2.2 Subjects

The scientific testing of animals is covered by the Animals (Scientific Procedures) Act 1986, the guidance to which makes clear that the unlicensed testing of any invertebrate species

⁴Given the technical issues encountered with the ant species, it was thought prudent to not specify an endpoint, but to conduct as many trials as possible within a period of 5–6 months. Although optional stopping (where researchers end testing during a glut of hits) is a potential problem with an unspecified number of trials, the number of trials per generation was pre-set at large whole numbers (100 or 200) and thus although the number of generations was not set, the number of trials was to some degree.

(excepting the octopus) is legal (Department of Home Office, 2000). The relevant BPS ethical guidelines were consulted also. Because of this pragmatic restriction and the fact that Cox had proposed using ants, testing initially began using a purchased population of yellow meadow ants (*L. flavus*). It was planned that 100 ants were to be tested in the initial phase of the investigation. Due to the circumstances outlined above (the loss of the ants), 650 red flour beetles (*T. castaneum*) were tested over five generations.

6.1.2.3 Apparatus

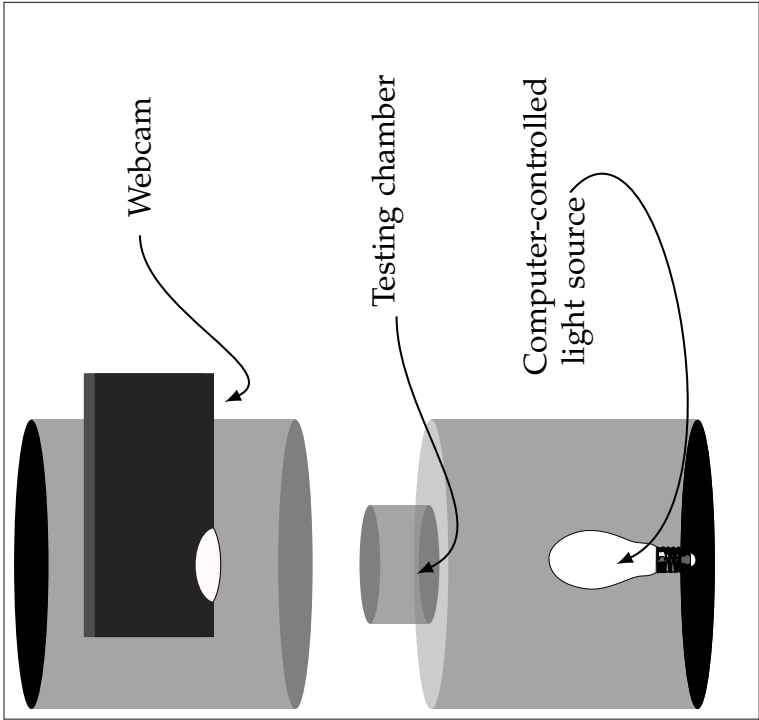
A special insect testing apparatus was designed for the purpose, and built by a departmental research technician (SY) (see Figure 6.1a). It was essentially a large digital microscope. A web-cam was fixed inside an opaque tube, and faced down over the testing chamber. The testing chamber consisted of a cylinder constructed from acrylic glass, 1 cm tall, 6 cm in diameter. When an insect was placed into the testing chamber, which locked with a pin also made from acrylic glass, the top half, with the web-cam, fixed onto the bottom section to ensure that the testing chamber was totally devoid of light.

A light-bulb was positioned underneath the testing chamber, which when illuminated, provided a view inside the testing chamber from the web-cam, which was attached to a desktop PC. An example of the type of image provided when the light was illuminated can be seen in Figure 6.1b. A piece of computer software was developed to control the apparatus and record the results (written in C++ by SY). The division of the testing chamber into two equal sections (the two semi-circles making up the testing chamber) was to be achieved via computer. However, the need to take apart the equipment during testing resulted in a slightly shifting field of vision, which made the computer division of the testing chamber a technologically difficult task. As such a physical demarcation line was drawn under the testing chamber, such that when illuminated it was apparent where the border between the two semi-circles was. As this demarcation line was not inside the testing chamber, the insects could not perceive it.

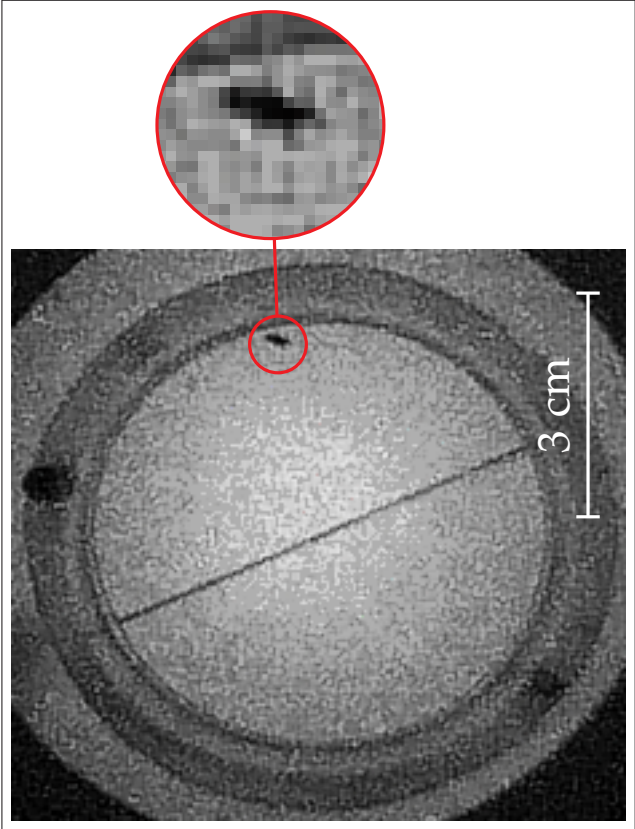
A number of well labelled animal tanks were used to house the different life-stages of the insects, around the lids of which was painted liquid Teflon to prevent the animals escaping. Beetles were housed in a flour-yeast medium (95% whole-wheat flour) and kept at a temperature of between 28–30°C (with a relative humidity of between 60–70%). Three sealed containers were marked, 'parent', 'survivor' and 'death' respectively for the tub containing the parent selection, the tub containing the survivors of the psi task and the death-jar used to kill the insects, which consisted of a small airtight jar into which cotton balls with three drops of acetone were placed. Although veterinary advice was sought to achieve the 'best' termination method for the insect species, the suggested method of liquid nitrogen was considered too technically challenging given the testing environment of a psychology laboratory. Sinnock (1970) used a preheated Petri dish taken from an oven maintained at 60°C and reports that the beetles were dead within 30 seconds. We settled on acetone as the preferred killing method as it is widely used by insect collectors to kill specimens and testing proved effective on the species used here (beetles were dead within a 30 second period).

6.1.2.4 Procedure

It is important to note that any insect selected to be tested was only going to be given one opportunity to pass the psi-test. Every insect after the first generation would be "descended" from an insect that had passed through the psi-test and survived. Insects in the population that were not subject to the psi-test were culled, as were individuals who were not successful



(a) Testing apparatus



(b) Image showing testing chamber

Figure 6.1: (a) Shows a diagram of the insect-testing apparatus. (b) An image showing inside the testing chamber with beetle highlighted (magnified).

in the psi-test. First an insect was taken from the container marked 'parent'. For the first round of testing this is essentially our first pool of subjects. The selection process was not based on any criteria other than being central to the container. As any insects not selected for testing during each stage were guaranteed to be culled, it can be argued that this selection process was an additional psi-test. For our purposes this is not a criticism. If psi has facilitated some insects in positioning themselves in favourable locations (to be selected for the psi-test) then we might expect that of those individuals who are selected for the psi task, there might be a propensity for the more sensitive individuals to be selected. Selected individuals were placed into the centre of the animal testing chamber (or as close to the centre as was possible). The lid was then closed and fixed with a pin made of acrylic glass. The camera section was then placed over the insect chamber. This process usually took seconds. Once the apparatus was ready the control software was initiated. The software was designed to turn on the bulb and take a picture of the insect's location in the chamber after a set period of time (in this case it was 1 minute). The software then had to wait a set period of time (another minute after the picture was taken) before randomly choosing a side (A or B), which corresponded to the two semi-circle areas. Once the computer had made the binary choice, the photograph of the insect's location in the apparatus (which had been taken before the side was chosen) was revealed to the experimenter and in the appropriate dialogue box the side that was the target for that trial was presented. A hit was counted if the insect was not on the side chosen by the computer and a miss was when the insect was on the side chosen.

Unlike other psi-tests however this was not the end of the process. A hit resulted in the insect being taken from the insect chamber and placed into a container clearly labelled as "survivors". A miss resulted in the insect being placed into the death-jar that contained the acetone. This would lead to the insect dying within approximately 30 seconds. To ensure that no biochemical markers would influence subsequent insect trials, the testing chamber was cleaned with an alcohol based wipe. This process of selecting individuals and either sorting them into the survivor tank or culling, is the basic procedure for each stage. At the end of a stage (a large whole number of testing—between 100 and 200) all the remaining insects in what had been the parent tank were culled (using the same method as was used during testing). At this between testing stage there is only one subset of individuals who remain alive: those who were chosen for and survived the psi-test. These survivors were housed in optimum conditions and thus they were intended to breed with one another. (Boake, 1986, reports that male flour beetles do not specifically court females and generally attempt to copulate with any beetle, including other males and pupae.)

As soon as a large number of larvae were visible in the survivors' container (that is, a number sufficient to supply enough adult insects for testing), the adults were separated from the juveniles. This is quite easy to do with the beetles as sieving the flour substrate easily separates the adults from the eggs and pupae. The removed adults were the survivors of the previous stage, who were at this point in the procedure now culled after separation via the methods described. The remaining lone juveniles were left to mature in their container. It is interesting to note that the adult beetles are apt to cannibalise both eggs and pupae (Wade, 1976) and thus any surviving pupae are not only descended from surviving individuals, but themselves have survived potential predation from their own kind.⁵ Thus it can be seen that the laboratory environment, rather than being a sedate environment punctuated by a death

⁵Park, Ziegler, Ziegler, and Mertz (1974) report that there are six predatory pathways (p. 37), "adult beetles eating eggs; adults eating pupae; larvae eating eggs; larvae eating pupae; larvae eating larvae, and adults eating larvae." Moreover, females are more voracious cannibals of eggs, whilst it is the males who are the greatest danger to pupae.

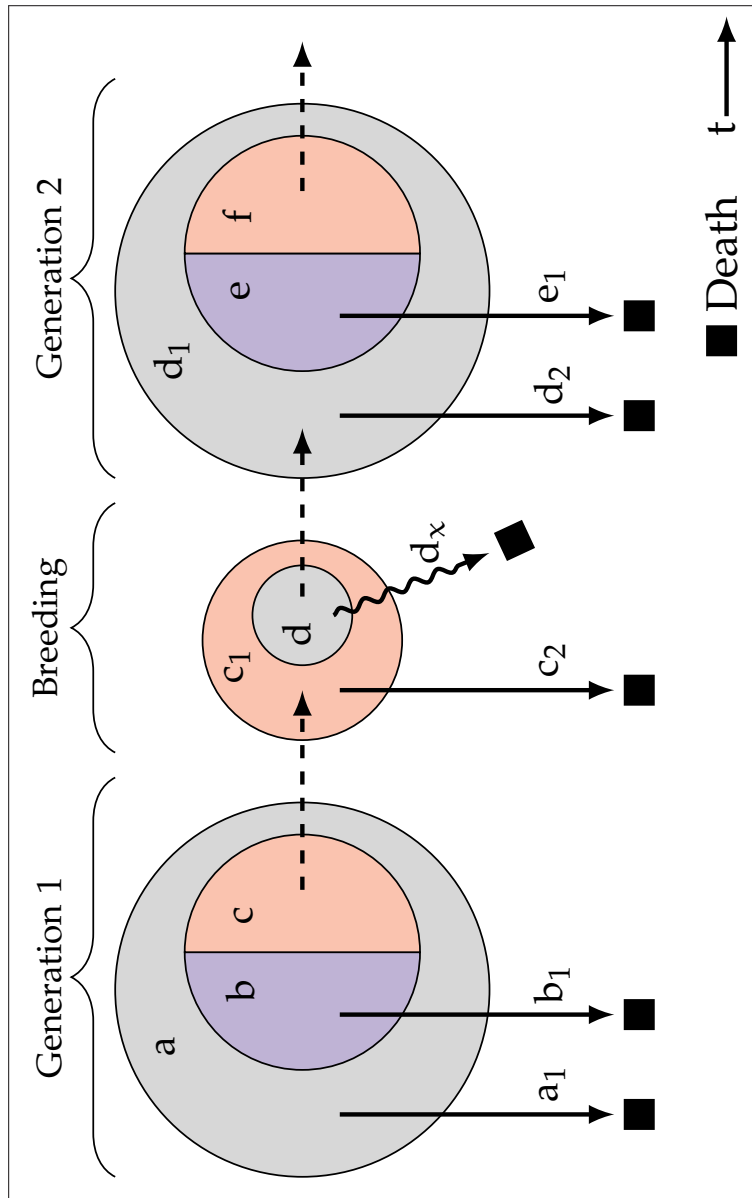


Figure 6.2: Graphical representation of insect testing procedure. First we begin the experiment with a population of several hundred insects, marked as (a). Of these insects a proportion of them are selected for individual testing within the apparatus, this is the whole circle (b+c) within (a). Of those that are tested, one proportion will pass the psi-test (c) and the other proportion will not pass (b). Any of (a) that are not selected for testing are culled (a₁). Any insects that do not pass the psi-test (b) are also culled (b₁). Those insects that do pass the psi-test (c) are housed together in a container. This population (c₁) is composed of adult insects and is left to breed together. After a period of time (~30 days) several hundred juvenile insects become visible. (The life-cycle of the beetle goes from egg → larvae (mobile immature) → pupae (immobile immature) → adult). Up to this point in the process there are a mixture of (c₁) adults (the “parent” population) and (d) eggs, larvae and pupae (the “offspring” population). Given the cannibalistic tendencies of the species there is a loss of (d) to (c₁) marked with the wavy line (d_x). This is the time when the adults (c₁) are sorted from the immature stages (larvae, eggs and pupae (d)). All of the adults are culled (c₂) and the eggs, larvae and pupae are allowed to mature into adults (d). (Early maturing adults in (d) can also predate on those beetles at an earlier stage of development and this also contributes to the loss of some beetles (d_x)). The matured adults of (a) then become the equivalent of (a) in the preceding circle (and is the source population from which insects will be drawn for the next stage). The experiment continues as before. A proportion of the insects (d₁) are drawn for testing in the apparatus (e+f). Those that do not pass the psi-test are culled (d₂). Those that pass the psi-test are culled (e₁). Those that pass the psi-test (f) are housed together and the experiment passes through another breeding stage (equivalent to c → c₁ → d → d₁). This procedure was repeated for 5 generations.

lottery, provides a fair analogue of 'nature, red in tooth and claw'.

As soon as a sufficient number of pupae have matured into adults, testing of the next generation began. This procedure was repeated over five generations of beetles. A graphical representation of the procedure (with further explanation) can be found in Figure 6.2.

6.1.3 Results

The results for all of the insect death-avoidance tests are reported in Table 6.1. When referring to the different stages of the experiment we will use the initial B which stands for beetle followed by the number of the testing stage (1–5) and the ant stage is provided only on the basis of completeness. The main test of significance is a binomial test, using the `binom.test` function of R. The binomial test gives us an exact p-value for the number of successes obtained, testing the null hypothesis of a Bernoulli experiment. Essentially it tells us if the number of survivors differs from chance (one-tailed). Only two results are in the correct direction (there are more surviving insects than those that were killed) although neither result is significantly different from chance. Thus there is no evidence for above-chance death-avoidance in any of the generations (B1–5) tested.

Table 6.1: Results for all insect trials. The P column refers to results of the exact binomial test (one-tailed) and is the probability of observing a result if the null is true. The empty cells denote a result that is in the opposite direction to that predicted and thus are not necessary to test.

Stage	N	Survivors	P
Ant	35	21	ns
B1	150	76	ns
B2	200	98	–
B3	100	50	–
B4	100	47	–
B5	100	45	–

Table 6.2: Two tables showing frequency of the side chosen and the side insect was on. The P column refers to the associated binomial test result which indicates whether the result is non-random.

Stage	A	B	P
Ant	17	18	ns
B1	69	81	ns
B2	102	98	ns
B3	53	47	ns
B4	50	50	ns
B5	48	52	ns
Total	339	346	ns

(a) Side Chosen

Stage	A	B	P
Ant	18	17	ns
B1	95	55	.001
B2	116	84	ns
B3	59	41	ns
B4	45	55	ns
B5	43	57	ns
Total	376	309	.01

(b) Side On

Two further tables are presented in Table 6.2. Table 6.2a shows the frequency of the sides chosen by the computer RNG and the associated p-value. These results are two-tailed (unlike the preceding table) and none are significant. This suggests that the computer-based RNG has not produced data that favours one side (A or B) over the other. Table 6.2b is the corresponding table for the side that the insect was pictured on. This table *does* show two significant p-values,

indicating that for the beetle stage 1 there was a significant propensity for the beetle to be found on side A than side B (95 against 55) and this trend is found in the totals (376 against 309). It is interesting that B1 is the only beetle testing stage that has a survival score in the correct direction (more survivors than those killed). It seems clear that there are three explanations for this significant trend for the insect to be found on side A over side B:

1. An environmental artefact is causing the insects to favour one side or the other. This could be as simple as biochemical markers or food particles, to more complex (and difficult to eliminate) factors such as an impinging electrical field, or infrasound source focus. Or it might indicate that the insects are showing stereotyped behaviour.
2. The trend to favour A over B is the result of a precognitive influence, reflecting the slight propensity for the RNG in beetle stage 1 to favour side B over side A. This would result in more survivors than those being killed (which was found, though non-significant).
3. Although the p-value is significant the confidence intervals for the same result range from .55 to .71, and thus the significant p-value may not indicate a truly non-random result, but may in fact be a chance finding (albeit an unlikely one).

Given the results, it is difficult to definitively conclude what the reason for the significant p-value actually is. However it is reassuring that the trend is not found throughout the stages of testing and so if it is a systematic bias like (1) it was minor and was eliminated or reduced. Thus, the hypothesis of insect death-avoidance has not been supported.

6.1.4 Discussion

From a parapsychological perspective, the results from testing the precognitive death-avoidance ability of five generations of flour beetles have not been successful and in fact provide demonstrable evidence of survival rates which do not differ significantly from that expected under the null hypothesis (that the probability of survival is at the chance level, 0.5 in all generations). Although no specific end point for the experiment had been specified a priori (and therefore we might be accused of ending the experiment during a dearth of hits—a kind of negative optional stopping), the results are not even suggestive of death-avoidance behaviour. Although much of the evidence reviewed throughout this thesis suggested a regime of testing such as this might be one of the better ways of establishing a precognitive influence, the failure to demonstrate extra-chance scoring is disappointing if not wholly definitive.

Given the mixed evidence for anpsi reviewed in the introduction, we can hypothesise a range of potential reasons for the failure to find even a very basic effect in this study. The most obvious conclusion must be that the null hypothesis is essentially correct and that precognitive death-avoidance is not a real behaviour, but a speculation in need of supporting data. Although this experiment included a selective breeding element, it is not true that we expected *no* above-chance survival in the first round of testing—the basic hypothesis is that nature has been selecting for this ability over the whole course of biological evolution. What we hoped to find was above-chance survival rates which *improved* over the course of successive generations. That no above-chance survival has been found is damning given the theoretical underpinnings of the experiment. However it is also true that no parapsychologist undertakes an experiment with an expectation of finding significant above-chance scores. Absence of evidence does not mean that there is evidence of absence and as stated there are any number of possibilities why

a significant result was not obtained (though such alternative explanations look weak in the face of the null hypothesis).

Firstly then, one of the stronger arguments must be that the red flour beetle just does not have the same abilities that humans do. Obviously the flour beetle has diminished cognitive and social abilities as compared to a human. However it is able to “choose” its location in space and at the very least is just as likely as other diverse species (such as the paramecia or rats) to exhibit a psi ability. An argument that animals do not have psi is often based on the idea that psi is something unique to humans, and whilst it is possible that psi may be a consequence of consciousness, much of the research cited in this thesis implies the very opposite (that psi is possibly shared throughout the animal kingdom). Even if psi is a universal trait due to its age (as a pre-sense say), that does not mean that all organisms will exhibit psi in the same manner and to the same degree. It is certainly possible that flour beetles have potentially traded the adaptive precognitive death-avoidance for a more adaptive ability—we might speculate that genes that may have been involved in increasing sensitivity to the psi-signal have been co-opted or replaced by genes that are now important for flight. Such quirks of evolutionary history may mean that response to the psi-signal is not found with equality throughout nature. On this argument it thus appears imperative that a researcher conducting such an insect based experiment must first demonstrate a psi-effect that appears to be focused on the species in question (rather than the experimenter being the source) before engaging in a more extended testing and breeding regime. The question as to whether *Tribolium* is actually capable of providing evidence for a psi-effect remains an open question for future anpsi research to answer.

Another related possible confound relates to the predation of the larvae, pupae and eggs by the adult form. This includes both the adults that pass successfully through the psi-test (marked on Figure 6.2 as (c₁)) but also the cannibalism of early maturing “offspring” on the less developed members of their cohort. It is possible (but difficult to estimate the importance of) that the selection process for the “psi-gene” has a developmental influence over the beetles. For example, those prodigious cannibals may perhaps be the less psychic individuals (or the more psychic individuals may be slower at developing). Thus although we may have a population at (d) which is composed of more or less psychic individuals, that the cannibalism is biased in favour of the less psychic (because they are the individuals who are slower at developing to an adult stage, and are thus in *more* danger of being predated on). This is a difficult problem to deal with (although it is not impossible to maintain the insects in individual habitats and thus reduce predation to an absolute minimum—the adults could be kept together for a brief period after successfully completing the psi task, sexed and then the females could be allowed to lay a certain number of eggs after which they are removed). Although technically feasible it appears to require an intensive methodology.

A further related confound relates to the health of the populations being tested and specifically the influence of parasites. It was noted in the introduction that maintaining the ant species in captivity was immensely more difficult than had been anticipated. It is widely recognised that *Tribolium castaneum* harbour a wide range of parasites, both endoparasites such as bacteria and sporozoa, to ectoparasites such as mites and parasitic wasps (Sokoloff, 1974). Although no significant extrinsic mortality was observed during the testing run and populations appeared healthy throughout the experiment, it is certainly possible that there was an unobserved parasitic pressure on the individual insects. As parasites bring with them their own set of genes it is difficult to say what kind of influence such a parasitic infestation

would manifest itself as. However, if we assume that parasites reduce the genetic fitness of the individual, then it is possible that the beetle could be weakened to a state that means that its response to the psi-signal is at chance levels. A highly speculative line of argument might run something akin to the opposite: that it is the “extended phenotype” (Dawkins, 1999) of a parasite that confers upon the host the psychic influence being sought. Just as there is some evidence that the protozoan *Toxoplasma gondii* inhibits rats’ fear of cats (thus facilitating the parasitic infection of the cat host), perhaps the psychic influence is not a consequence of the host biology, but the “extended phenotype” of a parasitic agent (the parasitic bacteria or virus confers the psi ability). Again, it is for future anpsi research to investigate the influence of parasites on animal subjects.

Moving on, if the species itself was then acceptable (health and predation concerns aside) the next major criticism must be that the length of the selective breeding element was too small to have had a large-scale influence on the population being tested. If we assume that the fact that generation B1 had an approximately 50/50 chance of survival (which might be argued is evidence of a lack of death-avoidance behaviour), then it is difficult to say what number of generations would be required before we could dismiss the underlying hypothesis altogether. However Wade (1976) reported that in his experimental test of group selection of the same species (*Tribolium*) selecting for adult population size, that:

The response in the group selection treatments occurred rapidly, within three or four generations, and was large in magnitude, at times differing from the controls by over 200%. (Wade, 1976, p. 4604)

It appears safe to conclude that the 5 generations tested here fall short of the number required to definitively answer the question (in the absence of significant results), but as a preliminary, exploratory investigation certainly had the potential to provide evidence which would have been important if forthcoming (this is a practical concern—time and resources—and one that could not be met in this particular empirical regime). Future researchers who undertake a similar experiment are urged that it is imperative that the number of generations be increased, perhaps substantially.

If the species was appropriate and the number of generations tested at least allows the possibility of discovering a significant effect, perhaps it was the testing environment which eliminated the effect. We noted in the results section that there is some evidence for the insects to prefer one side of the testing chamber over the other. Although it is not apparent what the cause of this preference is (if any) it is possible that if the influence is sensory, that there are strategies that might remove the confounding influence. For example, an alcohol wipe was used to remove any biochemical markers or particles of food that might influence the position of the insects within the chamber in a non-random way. Ryan, Park, and Mertz (1970) found that adults were attracted to the water extracts of macerated pupae and that removing adult antennae destroyed the preference. Given some of the evidence in the introductory chapter regarding the nature of sensory deprivation, perhaps removing the adult antennae would be a good strategy to removing any normal stimuli that might influence the beetle behaviour in a non-random way (however the antennae are not the only part of the beetle that contain sensory organs). If strong electrical fields or infrasound sources are biasing position in the apparatus there are also strategies available to the researcher (for example Faraday cages). Other ways of avoiding these potential confounds are described below when discussing possible future adaptations of the methodology.

A further example of a way in which the testing environment itself might influence the

results is suggested by the anecdotal finding that a great many of the web-cam shots showed the insects preferred the edges of the testing chamber to the middle. For example examination of the images shows that in B2 for only 34 trials out of 200 was the insect away from the edge (.17 trials). Did stereotyped behaviour (e.g. a flight urge) overwhelm the propensity to respond to the psi signal? As the whole chamber was deficient of light, it is easy to imagine that the beetles merely found an edge and proceeded to follow it, depending on which way they initially move from the centre point. As such, the beetles may have simply been travelling around and around the testing chamber, following their stereotyped innate behaviour, and did not or could not respond to any precognitive signal. This does, however, provide a good analogue to the real world experiences. Human beings are not constantly monitoring their internal states, attempting to discern any potential psi-signal that might potentially indicate a dangerous event in the future. Instead we go about our lives, often motivated by needs and intentions that may swamp any delicate and unusual responses indicating a psi-influence. The beetles obviously do not have psychological needs to fulfil, but the evolutionary history of their species has undoubtedly shaped the hierarchy of their responses. Alone and without any obvious source of food etc., it could be argued that the beetle is overwhelmed by the evolved predisposition to escape the container and find food and mates, that any precognitive influence is ignored or unnoticed.

If the methodology employed has encouraged stereotyped behaviour, are there any methods that might reduce this concern? One such suggestion might be that the testing of individual insects one after the other is not the most efficient method of testing for a precognitive ability in this species. Rather than take individual insects and test them for their ability to pass a psi-test, perhaps it would improve the results if large numbers of the insects were tested in batches. For example, we might imagine that the maturing tank is composed of two halves and the insects are free to move between them with no impediment. After a certain amount of time one half of the tank is culled and the remaining population allowed to mature and be the "parent" population to the next "offspring" generation. Counts of individuals could easily be taken by sorting the adults from the juveniles on both sides. This might be a preferable methodology to that employed in this experiment as it might amplify any psi-response by a kind of majority rule response. If only a certain percentage of the species are high-psychic individuals, perhaps they will remain on one side of the tank, and perhaps this propensity to remain on one side would allow less psychic individuals to themselves improve their genetic fitness and remain on the surviving side. (Stevenson, 1970a, describes the case of a villager in Japan who dreamt that his village was going to be washed away by a tidal wave. On waking (still the night) the villager woke the other villagers and they all escaped to safety. A tidal wave followed and would have killed them all.) Without further experimentation, this is of course highly speculative, but it remains an open question whether such a methodology is to be preferred over the procedure used in this experiment. A similar adaptation might allow selective routes of escape from the tank, such as a suspended thread from the level of the flour up into a collection chamber. The thread could be randomly placed in the tank for a set duration and the survivors taken only from individuals who leave the tank via the thread.

Another possible future adaptation could perhaps use not one, but two competing species. It is well established that different species of *Tribolium* compete readily against each other, both by exploiting the same food resource but also by mutual predation and cannibalism (e.g. Park, 1962; Goodnight, 1990). (The two species can also be bred in two different colours, which would facilitate distinguishing between the two and ensuring that equal numbers are selected

for testing.) If for example the initial populations consisted of *T. castaneum* and *T. confusum* we might expect an added level of pressure to be applied within the system. Whether this would facilitate precognition is again, an open question.

Finally, there is another *theoretical* possibility, that it was the fact that the entire group of animals was eventually killed, regardless of their success or failures individually, that eliminated the psi-effect. As with Morris's rats, it could be argued that since there is no way for the animals to avoid the fatal future event, it does not matter if they are successful on the precognitive test as it is only delaying the inevitable, that they will not be able to precognise the event. When the last generation of survivors in this experiment is culled, there are no more beetles left (from the original population obtained for testing). Thus it might be argued there is a corresponding cascading collapse of the precognitive influence, that would essentially prevent precognition at all stages of the experiment (like a rope bridge which has collapsed at the far side, travellers cannot even begin to cross the bridge). This is not an argument about the longevity of the individual animals. Were the final generation released into the "wild" (or allowed to live out their entire life-span), the argument itself would not apply. It is the fact that the generational testing itself forms a chain (B1 ← B2 ← B3 ← B4 ← B5) and that culling the successful individuals at B5 eliminates any precognitive influence throughout the system. It could be likened to running the whole experiment upon the *Titanic*: that the system as a whole allows no possibility of *actual* success, though there is the appearance of *possible* success. It appears to boil down as to whether the generational testing constitutes a psi-system through which there is information flow, or whether only individual success is important. It appears unlikely that successful precognition in humans is dependent on familial success across generations, although there is some minor evidence in this regard for example the Scottish tradition of second sight—often attributed to the seventh son of a seventh son, (Cohn, 1999; Campbell & Hall, 1968). This appears to be another question that must be left open for future researchers. Success at any of the stages alone would have helped provide some evidence that the system as a whole does not need to be successful along the chain, but again the absence of evidence does not support any theory either way. Lacking evidence to the extent to which that concern is the cause of the null results we might argue that in nature it seems that even a small time alive (extra) raises the likelihood of reproducing, which an animal cannot do when it's dead, thus in this current experiment, the fact that the survivors could reproduce, should to some extent have reduced the theoretical concerns (life will find a way). The solution, releasing the surviving individuals into the "wild" may allay this concern but releasing large numbers of beetles, ants or rats (or any other laboratory animal) is difficult to justify, and no laboratory is going to maintain animals for their natural life-span to investigate precognition.⁶ As such, any animal death-avoidance research must seemingly kill all of the testing subjects at some point or employ a different type of methodology.

Other parapsychologists might argue that there are other factors, such as the experimenter effect, which could explain the null findings (e.g. White, 1977; Smith, 2003). This is the idea that the experimenter's psi influences the outcome of the results. However, the range of authors that have conducted animal death-avoidance research have mostly been of a positive persuasion, which has not influenced the mixed nature of the overall results. Also the overall opinion at the start of this experiment was wholly positive, in spite of the lack of other empirical support for precognition. Perhaps this experimenter does lack the requisite psychic powers necessary

⁶Pearl, Park, and Miner (1941) report that the adult life-span of *Tribolium* is one of the longest of the "stored-product" insects and some individuals have been recorded to live over 3 years.

for above-chance scoring. As such it will be left for future researchers to show significant death-avoidance in a similar empirical regime (or not, as the case may be).

The final conclusion must be that in this study no death-avoidance behaviour was demonstrated and likewise none has been demonstrated reliably by other researchers. Without further experimentation with adaptation of the methodology, it will remain an open question as whether the lack of significant survival above the .05 level is evidence of absence of psi. Even testing the progeny of successful beetles, and in the last generation, progeny of several generations of successful beetles, revealed no significant trend to avoiding death. If there is precognitive death-avoidance, it does not seem unreasonable that it should have been demonstrated (though of course, there are seemingly numerous reasons why it might not be found even when such an ability exists). It is with minor regret that this experiment was not the first of this thesis and that it was not composed of a great deal more generations of testing, but as with evolution itself, quirks of history can impose restrictions upon us, and it will be left for future researchers to definitively show whether or not precognition equates to a kind of death-avoidance.

Conclusion

When we talk of tomorrow, the gods laugh.

—CHINESE PROVERB

7.1 What has been learnt?

The title of this thesis is a question: Is some of the evidence for ostensible precognition indicative of Darwinian adaptation to retrocausal influences? And at the very onset of the introduction we asked whether humans and other animals can perceive the future, not in relation to completely normal cognitive predictions about what the future *might* bring (a perfectly legitimate target for psychological investigation), but instead concentrating on whether the future is in some way knowable, and can provide information which is useful now in the present through the hypothesised parapsychological process of precognition. The majority of the research conducted during this empirical regime has attempted to directly glean evidence that allows us to answer these questions. Though, at the end of this thesis it appears clear that the answer to both sets of questions is “probably not”.¹

Parapsychology is a research area beset by numerous serious difficulties. Foremost is the indisputable fact that the underlying subject matter parapsychologists purport to investigate is denied by a majority of conventional scientists as having a veridical basis in reality. To put it bluntly, most scientists do not believe that psi exists (and this includes precognition) (Evans, 1973). The reasons for this disbelief are many and varied, and may be theoretical (belief that cause and effect are inviolate), to more practical matters (that the evidence purported to support the hypothesis has a mundane explanation, e.g. fraud). Given the work described in this thesis, it is clear that such scepticism is not without justification. We must have results that seriously challenge conventional theory before we can begin to accept functioning that falls far beyond what science currently accepts (it is *para* to psychology after all), and one of the main conclusions must be that not only is such evidence not easily forthcoming, but previous claims of such evidence appears to not be as compelling as it may have at first seemed. Furthermore, parapsychologists themselves are quick to acknowledge that there is no guarantee of paranormal results when one undertakes such research (Beloff, 1985; Schmeidler, 1994). Even replications of previously successful experimental techniques are not generally expected to produce results which indicate paranormal functioning on every occasion. Instead there is a

¹This is a somewhat subjective opinion, and it is certain that others evaluating much of the same evidence might come to a different (and arguably reasonable) conclusion. However despite the seeming simplicity of the conclusion, it belies a multifaceted assessment of the evidence for precognition (theoretical, methodological, philosophical etc.) which have informed that conclusion.

corpus of evidence that parapsychologists argue demand their paranormal hypotheses, even in the absence of replication.

Therefore, at the close of this investigation into precognition, it must be recognised that the majority of experimental methods employed during the empirical tests described herein were chosen because they appeared to be the most likely candidates at supplying evidence which would demand a paranormal explanation, that the chance of obtaining *no* such evidence was, of course, relatively high (that is, if psi does actually exist in the way parapsychologists envisage). Thus, the most general conclusion must be, that though we have not, on the whole, obtained statistical evidence that requires recourse to a paranormal hypothesis of precognition (or any other parapsychological hypothesis), we cannot categorically rule out the possibility that such a thing may be positively demonstrated in future empirical tests. Of course, given our investigation, we choose to make the slightly stronger conclusion that precognition (as useful knowledge acquired from the future) probably does not occur, but again, we must emphasise the possibility that future research may well yet yield definitive evidence of such phenomena. Although this conclusion may appear somewhat anti-climactic, it is certainly not true that this research has not yielded any interesting findings, many of which have been covered in the individual chapter conclusions.

A brief recapitulation of the thesis appears worthwhile. In Chapter 1 we examined a range of paranormal claims, looking at the historical, anecdotal and empirical evidence. Specifically, this concentrated on the best evidence for precognition, which suggested that premonitions seem strongest when connected with a future death or accident and thus that precognition, if real, might sometimes be useful. Moreover, there is some suggestion that precognition might have a biological basis (animal studies, evidence from altered states of consciousness and limited evidence of familial linkage). Given that particular conclusion, evolutionary perspectives for the study of parapsychology were considered with the summation that if precognition is real, it may represent an evolved ability to avoid a future fatal circumstance (and thus increase genetic fitness, increase time alive to reproduce, etc.). It was, however, recognised that the null hypothesis remains an attractive explanation for much of the evidence presented.

Given the conclusions of the introduction, Chapter 2 covered a review followed by two empirical tests of the *presentiment effect*, a proposed precognitive phenomenon whereby participants' physiology (namely skin conductivity) changes in the present, in "response" to arousing stimuli presented in the future. Although the results from neither experiment required recourse to a paranormal hypothesis, the research highlighted a number of hitherto unappreciated difficulties in conducting such parapsychological research. Firstly, using physiological measures as indicators of performance in a psi task proved itself to be a much more difficult process than previous research had suggested and that there was substantial variance associated with said measurements. In more conventional parapsychology tests evidence for psi is argued when hit rates are statistically significantly above what chance (and the null hypothesis) would expect. At least such above-chance scoring is an objective measure that all sides of the debate can agree is unexpected (even if the reasons for such scoring are less than agreed upon). But what is a paranormal response when we are measuring psychophysiological responses? As already stated, this is a much more difficult question (controlling the environment, individual differences in responding, etc.) that ultimately demands very careful experimentation and some expectation of a differential response. Such research is, in conclusion, much more demanding than simple card-guessing tasks, and is fraught with the possibility that some

kind of methodological error has crept into the procedure and eliminated the psi influence (or produced results which might be misconstrued as such). This is a most frustrating position to be in. Thus despite the initial enthusiasm for an unconscious physiological measure of psi that we had at the onset of this empirical regime, it was clear at the end of the presentiment research that psychophysiological responses such as those being measured offered great opportunities to mistake error for psi.

Secondly, given the effect sizes obtained, issues regarding the lack of power of the empirical tests were also deemed substantial impediments to a scientifically useful investigation. If what was found *was* precognition, it required a magnitude larger investigation than that undertaken (which was comparable to others in the area). It is certainly possible that the other significant findings in the literature are equally low-powered and with it, significant findings are false-positives (something that is a distinct possibility with low-powered studies). Finally, the use of fear-relevant stimuli (spider stimuli) and appropriate fear groups was a novel adaptation to the standard presentiment methodology. The reasoning behind using such stimuli was a) they were arguably less ethically challenging than the violent or pornographic images favoured in other similar research, b) they allowed a meaningful differential response to be investigated (between the fear and no-fear participants) and c) that spider-fear can be argued as being a form of death-avoidance (spider-fear has arisen as part of our evolutionary heritage whereby given an environment where death by spiders is possible, those individuals who are appropriately avoidant of spiders are less likely to be killed as a result of an interaction with one). The results obtained in this chapter suggest that this aspect was also unsuccessful. If such stimuli are to be shown to be empirically useful in investigations such as this, future research must attempt to recruit participants who will show an exaggerated differential response to them (by comparing a high-fear and no-fear group), a procedural element that may also increase the effect size and reduce power demands (e.g. number of participants tested). Given the substantial difficulties encountered with the presentiment effect (despite initial enthusiasm) it was deemed that a different paradigm might provide the elusive evidence for precognition being sought and with greater confidence that any unusual findings were the result of paranormal functioning and not of some error. As such, no further investigation of the presentiment effect was pursued.

In Chapter 3 we presented a review of, and three empirical investigations into, what was termed *time-reversed interference*, a proposed retrocausal influence said to have been discovered during conventional cognitive tests based on the Stroop paradigm. Some 134 participants were tested in three different experiments, the first using an oral test, the second a replication using a manual task and the last a novel adaptation of the procedure using an emotional Stroop (and this adaptation continued the testing of spider-based stimuli and fear groups begun in Chapter 2). This chapter provides substantial investigation of the TRI effect and argues quite definitively that the influence almost certainly is not a precognitive effect but what evidence there is, is most likely the result of a conventional correlation that has been misconstrued as paranormal (we are not the only researchers to have made this conclusion). The methodology of Study II provides a unique difference from previous research (the manual task used required participants to respond through colour-coded keys) and the results of this study suggested uniquely that it is the task itself (and the ease with which responses are made) which has resulted in the misconstruing of a normal correlation between two participant responses being misidentified as a precognitive influence from the future to the present. (This finding is based on an interaction between ease with which the required response is made, i.e. orally calling one of four colours versus pressing a colour-coded button on a computer keyboard, and the

likelihood that participants will respond in a conventional Stroop-based manner). Although it was disappointing that the TRI effect did not yield evidence demanding a paranormal explanation, Chapter 3 demonstrates yet again the difficulty in conducting parapsychological research if the hope is to counter conventional scepticism and prove that seemingly impossible phenomena do occur. Moreover, as well as methodological issues, all three studies of Chapter 3 highlight the wide scope that researchers have in analysing their results. Although the variety of methods employed are justified to some degree by the previous researchers, it is clear that establishing unequivocal evidence for precognition is difficult when many potential statistical analyses may have been employed (that is multiple testing of the same data) and with it the potential for unreported fishing expeditions, looking for any potential deviations from chance that might exist in the data.

Study III also continued the test of spider stimuli as evolutionary based targets, although yet again, the main conclusion in that regard must be that if such stimuli are to be employed in parapsychological tests some effort must be made to recruit high-fear participants and no-fear participants, lest the difference between moderately fearful participants and low-fear participants force power requirements beyond a reasonable limit (i.e. the number of participants needing to be recruited is excessive). Finally, it was observed that TRI has been substantially over-simplified by some authors and that current presentations of the paradigm substantially overstate the case for a precognitive effect, a situation that is inherently irresponsible especially when such presentations are targeted at the lay person. Given the conclusions from this research, the TRI paradigm was (like the presentiment research before it) deemed as being unlikely to provide evidence requiring a parapsychological explanation and no further experimentation along these lines was pursued.

Given that the two investigated empirical regimes proved themselves to be much less simple than they had at first seemed (and this added complexity made ensuring results were truly paranormal extremely difficult), the research described in Chapter 4 was an attempt to again investigate a previously successful paradigm, this time using a very simple methodology that required little conscious input from participants, other than deciding when to press a button. As such, the psi-timing paradigm sat between the two empirical regimes already tested (it was neither entirely unconscious, nor required complex cognitive processing). Some 130 participants were tested in both precognitive and contemporaneous ESP tasks over three experiments. (Because the first precognitive experiment yielded no significant findings it was deemed useful to replicate the earlier published research that had claimed statistical significance, and this had been a test of clairvoyance, not precognition.) Interestingly, Study II produced results that could be construed as providing support for the initial parapsychological claims (though if this result is paranormal, it is not precognition).

Study I again attempted to test the value of employing fear-based spider stimuli as a target in a parapsychological setting, and again the lack of statistical significance highlights the importance for future researchers to attempt to tease apart the level of fear reported by participants and thus employ a more extreme contrast than was employed in this research. If spider-fear and spider stimuli are useful in a parapsychological setting, it will be in comparing high-fear and no-fear participants. Finally, the psi-timing experiment was the first experiment in this thesis that employed a more traditional parapsychological scoring method (akin to the Zener cards employed in some of the earliest parapsychological experiments). As such it allowed an investigation of two further potential locations of a psi influence, in terms of unanticipated patterning in participant scores (displacement and U-shaped scoring). No

striking evidence of psi-functioning was demonstrated beyond the result that we believe is a false-positive. Again this highlights the importance of power considerations and effect sizes when conducting parapsychological investigations. Although such issues have not been traditionally considered in the literature, it is clear that *all* parapsychological research must attempt to be as powerful as possible lest false-positives and low-power lead to studies being written off as scientifically useless and with it, much effort wasted. Given our conclusions regarding the psi-timing paradigm, no research along these lines was pursued further.

The research outlined in Chapter 5 demonstrated a significant deviation from the research of earlier chapters, and detailed the testing of a person who claimed an exceptional precognitive ability. DM claimed to have experienced over 200 precognitive dreams and often produced a pictorial record of his dream, which he claimed matched a later occurring disaster. Unusually for cases such as this, DM often attempted to get a photographic record of the picture with a date, taken at a reputable institution (e.g. a bank) to provide corroboration of the picture's prior existence to the disaster. At first examination DM's overall claim looked very impressive and difficult to explain as anything other than direct evidence for precognition. However, when testing whether independent judges would show preference for DM's events over another event chosen because of a subjective similarity to the images that he produced, we found that for the majority of pictures no such preference was found. This suggested that DM's "matches" were psychological and the result of subjective validation (once an event had occurred that matched, this was seen as the event that had caused the dream) and thus was not premonitory. Of the small number of DM's events that were preferred statistically significantly over the alternative, this was most likely due to the sometimes encountered problem that some of the events were more unique than others (and thus it was hard to provide good alternatives). Although this may appear to be a "head I win, tails you lose" scenario, it was established that one of DM's best matches (for the *Concorde* crash picture) was not rated by the participants as significantly better than the alternative event (evidently because the alternative event was just as good a match to the picture) and this suggests that given enough time, eventually events will occur that reduce the seeming uniqueness of even the most impressive matches. It is however true, that a secondary analysis found a small, but highly statistically significant difference between the ratings for DM's event and the alternative (favouring DM's event). Although some may attribute this result as supporting DM's initial claims, we identify a number of reasons why such a small preference may be found in the absence of such paranormal explanations.

Although not strictly an experimental test of precognition, the research outlined in Chapter 5 is important as we identified in the introduction that the spontaneous reports of individuals such as DM represent the very bedrock of parapsychology. Without such naturalistic accounts there would have been no impetus to engage in the experimental tests in the first place and if a seemingly impressive case such as DM has a more prosaic explanation the conclusion must be that perhaps many other such experiences have an equally mundane basis (psychological not parapsychological). Research into dream ESP continues (e.g. Sherwood & Roe, 2003, 2013), but most such research is more experimental in nature (generally comparing ratings of participant dreams with randomly chosen targets) and does not directly pertain to the type of study that we were restricted to (because of the nature of DM's claimed ability).

Finally, Chapter 6 detailed a large-scale anpsi (animal psi) study. Five generations of beetles were subjected to a behavioural test of precognition, where the consequence of failing was subsequent death. (Using personal death in this manner was a direct attempt at testing the hypothesis that precognition might be a kind of death-avoidance behaviour that had, up until

this point, only been alluded to by the use of the spider stimuli/spider-fear group element used in some of the previous research.) Yet again, the actual precognition test was extremely simple (spatial location within the testing environment) in an attempt to strip away some of the confounding variables that have complicated such previous tests. It was also argued that this particular regime of testing was a last, best test of the secondary hypothesis (within the confines of the research of this thesis), that precognition is a kind of death-avoidance behaviour. A novel aspect of this regime was the inclusion of a selective breeding element (the progeny of those subjects who passed the precognition test were the subjects of the next round of testing). Along with a specific testing occasion, the environment within which the insects matured (from egg to adult) provided potential predation pathways (entirely natural in origin) and these also had the potential for improving the genetic fitness of the underlying population (all insects tested could be described as survivors in this regard).

A total of 650 adult beetles were tested over five generations and the results obtained did not suggest any precognitive influence in the data. From a general perspective data collection went almost entirely as expected, except for the loss of the initial ant subjects due to difficulty in maintaining a suitable habitat for the species. Switching to the flour beetles seemingly rectified this problem (no extrinsic mortality was observed during the experiment) but the problem did highlight the possibility that population health might be a confound in this type of anpsi experiment (and a problem seemingly rarely discussed in the relevant literature). The lack of precognitive results in this test could not definitively be used as evidence for a lack of precognition per se and a number of possible reasons for the results (or lack thereof) were suggested. These included the possibility that the species itself might not possess the requisite precognitive ability, that perhaps the number of generations was too small, to more theoretical concerns that the destruction of the final generation of “survivors” may itself disrupt any potential precognitive influence. Again this extended discussion highlighted the numerous problems that can confound even a very simple precognition experiment, which lacks some of the demand characteristics that can influence human-based research. Finally although the conclusion must be that nothing paranormal was discovered with the insect trials, it must be left for future researchers to continue such research before we can confidently dismiss it as an entirely fruitless scientific pathway.

Summary

It is clear from this recapitulation that there are a number of common themes that have been made throughout this thesis:

1. Some conventional scientists readily dismiss the veridical nature of parapsychological phenomena (including precognition).
2. Previously published evidence purporting to indicate paranormal functioning has generally proved itself far more suspect than it appeared at first examination.
3. Replications of research whose results are taken as indicative of paranormal functioning are not expected to produce similar results every time. There are numerous reasons for a failure to replicate.
4. It is evidently common for an experimental technique to be hailed as a successful method of eliciting paranormal functioning, only for independent replications to fail and the area to fall into desuetude.

5. The dismissal of parapsychological results as being entirely normal in origin is theoretically difficult to do (given the nature of the hypothesis).
6. Despite (5) the null hypothesis remains constantly attractive given the many difficulties identified.
7. Although precognitive research can appear simple (and thus difficult to confound) in practice even the simplest paradigms seem to entail significant difficulties for the researcher. These might include:
 - (a) Small effect sizes mean small studies have low power and significant results may actually be false-positives. Such small effects in reality require large numbers of participants (or some other strategy for improving the study power).
 - (b) Difficulty in distinguishing conventional and parapsychological correlations. Normal influences may be misconstrued as paranormal.
8. Death-as-a-target appears to have been no better than other target type in eliciting a paranormal effect, however this may not be because the underlying theory is wrong.
9. Using physiological measures as indicators of psi functioning opens up a whole range of difficulties, which makes arguing that results are paranormal difficult, and thus ultimately they are poor candidates for such research unless the research is of an extremely high quality.
10. Our testing of spider stimuli and spider-fear groups as an attempt to both use ethically less problematic stimuli and attempt to force a differential effect has not been successful (although it should be recognised that this cannot be dismissed as a potentially useful methodological device until more extreme spider-fear participants have been compared with their no-fear counterparts).
11. Adaptation of previously successful methods can provide unique insight into the reason why normal influences might be misconstrued as paranormal.
12. There are a wide variety of statistical analyses open to researchers (including manipulations and adjustments) that may be used consciously or unconsciously to fish or mislead, including investigation of so-called secondary effects such as displacement or U-shaped curves.
13. Even the best cases of precognition may have mundane psychological explanations.
14. Those claiming paranormal abilities may not technically be fraudulent but may have little insight into the psychological motivations driving them.

7.2 Future Directions

Despite our conclusions regarding the veridical nature of precognition, we have little doubt that parapsychological research, and research into precognition in particular, will continue. Suggestions of the death of parapsychology have been made for almost a century, but the state of the field appears little different. However, in spite of our negative opinion of the underlying phenomena there are any number of directions researchers may wish to follow in the future.

1. Presentiment

The presentiment research described in this thesis was conducted over a decade ago. It should therefore not be surprising that presentiment research has continued (though there has not yet been any great breakthrough). There is one clear requirement of any presentiment research: It must be conducted carefully, by experienced psychophysiological researchers who understand the demands of such research and have the laboratory environment to meet them. Research that cannot meet this imperative is doomed to become bogged down in many of the issues already discussed.

2. Time-Reversed Interference

There is only one aspect of the TRI research that appears to demand further work and it is the procedure for Klintman's Studies II-V (Klintman, 1983, 1984). To date, despite the small number of replications of his Study I procedure, no experiment has attempted to replicate his secondary procedure, and thus it is imperative, if we are to conclude (once and for all) that Klintman did make a mistake when claiming to have discovered a paranormal effect in Stroop-based tasks, that his secondary procedure be evaluated. We await future replication to determine the final status of Klintman's claims.

3. Psi-Timing

Although the psi-timing research has fallen into desuetude, it is clear that it may be worth some effort in repeating the original procedure, ensuring the design is powerful enough to discover an effect. (The simple methodology is well suited to data collection over the web and such an investigation may allow many participants to be recruited.) There is a reason why this area is no longer pursued by other parapsychologists, but without a definitive investigation, it is difficult to conclude that psi-timing is a complete dead end.

4. DM

We have not remained in contact with DM and given his advancing age, it is doubtful whether DM will participate in any further parapsychological research. However, were such an individual to make similar claims in the future, it is clear that there are a number of aspects that would improve upon our own procedure. Firstly, although not strictly demanded by the empirical aspect of the procedure, it is regretful that we did not subject DM to a more extensive psychological and psychiatric evaluation. Was DM suffering from any particular mental health issues? (It is not DM's claims that prompt this question, but observational evidence regarding his living conditions and personal neglect.) This is not a question we are in a position to answer. But such evaluations would have resulted in a more well rounded examination of DM and his claims. Secondly, it would have been interesting to bear first-hand witness to DM's precognitive experience: if not monitoring him in a sleep-lab (which given the infrequent nature of his premonitions would be prohibitively resource intensive), then at the least being contacted by him before or during his picture production stage and thus attesting to the existence of the picture before production (and with it, ensuring all elements of the picture were there at the time of production). However we do not expect that a full psychiatric evaluation nor witnessing picture production would change our understanding of the underlying processes involved in DM's experience. But it would, as stated, provide a more comprehensive account of this very interesting case.

5. Animal Death-Avoidance

We suggested a number of adaptations to the insect study described in Chapter 6. It is clear that any future insect death-avoidance study must invest a great deal of time and effort, to improve over the previous research (and with no guarantee of success it is a brave researcher who undertakes that challenge). First, it must be established that the species itself is capable of responding in a way which indicates a psi influence. This is challenging in itself. If the researcher is happy that there is at least the potential for psi in the species used, then it is imperative that the number of generations tested be increased to a much larger number than that tested in our research. Such research would have demanding logistical requirements and be a significant undertaking. We do not envy the researcher who takes up such a challenge, though there is the possibility that subsequent significant results may provide the best evidence yet attained for the existence of precognition.

6. General Comments

The use of spider-fear is the other main aspect of this thesis and it has been noted several times during this conclusion that it is imperative that any future research that attempts to investigate spider-fear and spider-stimuli-as-targets, must endeavour to force a much larger differential effect and compare high-fear and low-fear participants. This may inevitably mean seeking out those suffering from some kind of clinical fear of spiders. Such high-fear participants were not generally sought out because: a) spider-fear was found to split evenly during testing of normal participants and thus was an economically simple comparison to investigate, along with the other reasons already stated for investigating this aspect, and b) that a clinical fear of spiders is not a guarantee of a more significant precognitive response. We would expect that if psi does exist, such comparisons of those with a clinical high-fear and no-fear participants has the potential for providing evidence that at least equals that already published.²

The success of using death-as-a-target has also not been supported by the research of this thesis. That being said, it appears too early to completely write off this aspect as being a complete dead end, unworthy of further examination. Although such a procedure is obviously off-limits with human beings, animal precognition tests where death is the ultimate target appear to continue to be one of the more interesting procedures open to parapsychological researchers. Significant results might still make it difficult to argue where the psi has originated from—experimenter or subject—though this problem might be lessened by a selective breeding element, as generational increase in survival rate appears to be more parsimoniously explained as a subject effect rather than coming from the experimenter.

7.3 Final Comments

After an extensive investigation of the problem of precognition, it appears fair to conclude that evidence has not been found that incontrovertibly requires the hypothesis of precognition and thus it does not appear likely that there exists an ability to “see into the future”, the question asked at the very onset of this investigation. Also, the evidence that stimulated the undertaking of this research, appears (with hindsight) to be most easily explained away by a myriad of conventional reasons, rather than rely on the extremely controversial hypothesis that is precognition. Why do some believe in precognition? A final conclusion is that precognition

²I have been involved in testing another precognitive paradigm, namely Daryl Bem’s *Precognitive Habituation* paradigm, again continuing the use of spider stimuli and spider-fear groups (Savva et al., 2004, 2005, 2006). This research produced generally negative findings, though the continuing criticism remains that it is imperative that high-fear participants must be recruited to provide a more extreme contrast.

is essentially a hypothesis whose existence is the byproduct of wishful thinking: we are a species with an enormous capacity for understanding and yet we are in a way trapped between an ever-unchangeable past and a future that represents an ever-present unknown. From this stark perspective there is little wonder that certain humans throughout history have believed that this position is not as fixed as it appears, and that the future might be knowable. Although this conclusion may appear to some as uncharitably bleak, it is true that the question regarding the reality of precognition will only be settled by careful experimentation, something that parapsychologists will continue to pursue. However, the parapsychological research conducted over the last century has ultimately not delivered on its initial promise. Despite that conclusion, there remains much about the universe which is unknown and mysterious.

References

- Akers, C. (1987). Parapsychology is science, but its findings are inconclusive. *Behavioral and Brain Sciences*, 10(4), 566–568.
- Alcock, J. (2011a). Back from the future: Parapsychology and the Bem affair. *Skeptical Inquirer*, 35(2), 31–39.
- Alcock, J. (2011b, 6th January). *Response to Bem's comments*. Retrieved June 24, 2013, from CSICOP website: http://www.csicop.org/specialarticles/show/response_to_bems_comments.
- Alvarez, F. (2010). Anticipatory alarm behaviour in Bengalese finches. *Journal of Scientific Exploration*, 24(4), 599–610.
- Alvarez, L. W. (1965). A pseudo experience in parapsychology. *Science*, 148(3677), 1541.
- Anderson, M. (1959). A precognition experiment comparing time intervals of a few days and one year. *Journal of Parapsychology*, 23(2), 81–89.
- Arnholt, A. T. (2010). *BSDA: Basic statistics and data analysis* [Computer software]. Available from <http://CRAN.R-project.org/package=BSDA> (R package version 1.0)
- Augustine, S. (2008). *Saint Augustine's confessions* (H. Chadwick, Trans.). Oxford: Oxford University Press.
- Averill, R. L. & Rhine, J. B. (1945). Effect of alcohol upon performance in PK tests. *Journal of Parapsychology*, 9, 32–41.
- Backster, C. (1968). Evidence of a primary perception in plant life. *International Journal of Parapsychology*, 10(4), 320–348.
- Barker, J. C. (1967a). Disaster early warning system. *Journal of the Society for Psychical Research*, 44(732), 107–108.
- Barker, J. C. (1967b). Premonitions of the Aberfan disaster. *Journal of the Society for Psychical Research*, 44(734), 169–181.
- Barker, J. C. (1968). *Scared to death*. London: Frederick Muller Ltd.
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275(5304), 1293–1295.
- Beck, A. T., Ward, C. H., Mendelson, M. L., Mock, J. E., & Erbaugh, J. K. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4(6), 561–571.
- Beloff, J. (1985). Research strategies for dealing with unstable phenomena. In B. Shapin & L. Coly (Eds.), *The repeatability problem in parapsychology* (pp.1–21). New York, NY: The Parapsychology Foundation.
- Bem, D. (2011, 6th January). *Response to Alcock's "Back from the Future: Comments on Bem"*. Retrieved June 24, 2013, from CSICOP website: http://www.csicop.org/specialarticles/show/response_to_alcocks_back_from_the_future_comments_on_bem/.
- Bem, D. J. (2003). Precognitive habituation: Replicable evidence for a process of anomalous cognition. In S. Wilson (Ed.), *Proceedings of the 46th annual convention of the Parapsychological Association* (pp.6–20). Vancouver, Canada.

- Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology, 100*(3), 407.
- Bem, D. J. & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin, 115*(1), 4–18.
- Bem, D. J., Palmer, J., & Broughton, R. S. (2001). Updating the ganzfeld database: A victim of its own success? *Journal of Parapsychology, 65*(3), 207–218.
- Bem, D. J., Tressoldi, P. E., Rabeyron, T., & Duggan, M. (2014). Feeling the future: A meta-analysis of 90 experiments on the anomalous anticipation of random future events. Available at SSRN. Retrieved July 17, 2014, from <http://ssrn.com/abstract=2423692>
- Bem, D. J., Utts, J., & Johnson, W. O. (2011). Must psychologists change the way they analyze their data? *Journal of Personality and Social Psychology, 101*(4), 716–719.
- Bender, H. (1966). The Gotenhafen case of correspondence between dreams and future events: A study of motivation. *International Journal of Neuropsychiatry, 2*(5), 398–407.
- Bestall, C. M. (1962). An experiment in precognition in the laboratory mouse. *Journal of Parapsychology, 26*, 269.
- Besterman, T. (1932). The psychology of testimony in relation to parapsychical phenomena. *Proceedings of the Society for Psychical Research, 40*, 363–387.
- Bierman, D. J. (1997). Emotion and intuition I, II, III, IV and V. Unravelling variables contributing to the presentiment effect. In R. Wiseman (Ed.), *Proceedings of the 40th annual convention of the Parapsychological Association held in conjunction with the Society for Psychical Research* (pp.49–62). Brighton, UK.
- Bierman, D. J. (2000). Anomalous baseline effects in mainstream emotion research using psychophysiological variables. *Journal of Parapsychology, 64*(3), 239–239.
- Bierman, D. J. & Radin, D. I. (1997). Anomalous anticipatory response on randomized future conditions. *Perceptual and Motor Skills, 84*(2), 689–690.
- Bierman, D. J. & Radin, D. I. (1998). *Anomalous unconscious emotional responses: Evidence for a reversal of the arrow of time*. Available from http://www.parapsy.nl/uploads/w1/retrocausal_tucson3.pdf
- Bishop, G. M. (1941). Foreknowledge in dreams. *Journal of the Society for Psychical Research, 32*, 50–59.
- Blackmore, S. J. (1995). What's in the box? An ESP test with Chris Robinson. *Journal of the Society for Psychical Research, 60*(840), 322–324.
- Blackmore, S. J. & Hart-Davis, A. (1997). *Test your psychic powers*. New York, NY: Sterling Pub.
- Blackmore, S. J. & Troscianko, T. (1985). Belief in the paranormal: Probability judgements, illusory control, and the “chance baseline shift”. *British Journal of Psychology, 76*(4), 459–468.
- Blagrove, M., French, C. C., & Jones, G. (2006). Probabilistic reasoning, affirmative bias and belief in precognitive dreams. *Applied Cognitive Psychology, 20*(1), 65–83.
- Boake, C. R. B. (1986). A method for testing adaptive hypotheses of mate choice. *American Naturalist, 127*(5), 654–666.
- Braud, L. W. & Braud, W. G. (1974). Further studies of relaxation as a psi-conducive state. *Journal of the American Society for Psychical Research, 68*, 229–245.
- Braud, W. G. (1990). Meditation and psychokinesis. *Parapsychology Review, 20*(1), 9–11.
- Braud, W. G. (2000). Wellness implications of retroactive influence: Exploring an outrageous hypothesis. *Alternative Therapies in Health & Medicine, 6*(1), 37–48.

- Braud, W. G. (2002). Psi favorable conditions. In V. G. Rammohan (Ed.), *New frontiers of human science: A festschrift for K. Ramakrishna Rao* (pp.95–118). Jefferson, NC: McFarland.
- Braud, W. G. & Braud, L. W. (1973). Preliminary explorations of psi-conducive states: Progressive muscular relaxation. *Journal of the American Society for Psychical Research*, 67, 26–46.
- Braud, W. G. & Shafer, D. (1989). Successful performance of a complex psi-mediated timing task by unselected subjects. *Journal of the American Society for Psychical Research*, 83(3), 205–216.
- Broad, C. D. (1935). Mr. Dunne's theory of time in "An experiment with time". *Philosophy*, 10(38), 168–185.
- Broad, C. D. (1953). *Religion, philosophy, and psychical research*. New York, NY: Harcourt, Brace.
- Broad, C. D. (1968). The notion of precognition. *International Journal of Parapsychology*, 10(2), 165–195.
- Broughton, R. S. (1988). If you want to know how it works, first find out what it's for. In D. H. Weiner & R. L. Morris (Eds.), *Research in parapsychology* (pp.187–202). Metuchen, NJ: The Scarecrow Press, Inc.
- Broughton, R. S. (2004). Exploring the reliability of the presentiment effect. In S. Schmidt (Ed.), *Proceedings of the 47th annual convention of the Parapsychological Association* (pp.15–24). Vienna, Austria.
- Broughton, R. S. (2010, April). *An evolutionary approach to anomalous intuition*. Paper presented at the 8th Symposium of the Bial Foundation. Porto, Portugal.
- Cadoret, R. J. (1953). The effect of amytal and dexedrine on ESP performance. *Journal of Parapsychology*, 17, 259–274.
- Camfferman, G. (1987). Time reversed interference: A replication study. *European Journal of Parapsychology*, 7(1), 13–31.
- Campbell, J. L. & Hall, T. H. (1968). *Strange things: The story of Fr. Allan McDonald, Ada Goodrich Freer, and the Society for Psychical Research's enquiry into Highland second sight*. London: Routledge & K. Paul.
- Cardeña, E., Iribas, A. E., & Reijman, S. (2012). Art and psi. *Journal of Parapsychology*, 76(1), 3.
- Carpenter, J. C. (2004). First sight: Part one, a model of psi and the mind. *Journal of Parapsychology*, 68(2), 217.
- Carpenter, J. C. (2005). First sight: Part two, elaboration of a model of psi and the mind. *Journal of Parapsychology*, 69(1), 63.
- Casler, L. (1962). The improvement of clairvoyance scores by means of hypnotic suggestion. *Journal of Parapsychology*, 26, 77–87.
- Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). *Graphical methods for data analysis*. Belmont: Wadsworth.
- Champely, S. (2012). pwr: Basic functions for power analysis [Computer software manual]. Available from <http://CRAN.R-project.org/package=pwr> (R package version 1.1.1)
- Child, I. L. (1985). Psychology and anomalous observations: The question of ESP in dreams. *American Psychologist*, 40, 1219–1230.
- Cicero, M. T. (1887). *The treatises of Cicero* (C. D. Yonge, Trans.). London: G. Bell.
- Cohn, S. A. (1999). Second sight and family history: Pedigree and segregation analyses. *Journal of Scientific Exploration*, 13(3), 351–372.
- Cox, W. E. (1951). The effect of PK on the placement of falling objects. *Journal of Parapsychology*,

- 18, 234–239.
- Cox, W. E. (1956a). Precognition: An analysis. I. *Journal of the American Society for Psychical Research*, 50, 47–58.
- Cox, W. E. (1956b). Precognition: An analysis. II. *Journal of the American Society for Psychical Research*, 50, 99–109.
- Cox, W. E. (1982). An auditory precognition within a radio newscast. *Journal of the Society for Psychical Research*, 51(792), 378–381.
- Craig, J. G. & Treurniet, W. (1973). Precognition in rats as a function of shock and death. In Morris, J. D., Roll, W. G., & R. L. Morris (Eds.), *Research in parapsychology* (pp.75–78). Metuchen, NJ: Scarecrow Press.
- Dalkvist, J. & Westerlund, J. (2006). A bias caused by inappropriate averaging in experiments with randomized stimuli. *Journal of Parapsychology*, 70(2), 233–254.
- Darwin, C. (1871). *The descent of man*. London: John Murray.
- Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.
- Dawkins, R. (1999). *The extended phenotype* (Reissue ed.). Oxford: Oxford Paperbacks.
- Department of Home Office. (2000). *Guidance on the operation of the Animals (Scientific Procedures) Act 1986*. London: The Stationary Office.
- Diaconis, P. & Mosteller, F. (1989). Methods for studying coincidences. *Journal of the American Statistical Association*, 84(408), 853–861.
- Dodds, E. R. (1971). Supernormal phenomena in classical antiquity. *Proceedings of the Society for Psychical Research*, 55(202), 189–237.
- Dosa, D. M. (2007). A day in the life of Oscar the cat. *New England Journal of Medicine*, 357(4), 328–329.
- Dosa, D. M. (2010). *Making rounds with Oscar: The extraordinary gift of an ordinary cat*. New York, NY: Hyperion.
- Dossey, L. (2008). Premonitions. *Explore: The Journal of Science and Healing*, 4(2), 83–90.
- Dukhan, H. & Rao, K. R. (1973). Meditation and ESP scoring. In W. G. Roll, R. L. Morris, & J. D. Morris (Eds.), *Research in parapsychology 1972* (pp.148–151). Metuchen, N.J.: Scarecrow Press.
- Dunne, J. W. (1927). *An experiment with time*. New York, NY: Macmillan.
- Dutton, D. & Williams, C. (2009). Clever beasts and faithful pets: A critical review of animal psi research. *Journal of Parapsychology*, 73(1), 43–68.
- Duval, P. & Montredon, E. (1968a). ESP experiments with mice. *Journal of Parapsychology*, 32, 153–166.
- Duval, P. & Montredon, E. (1968b). Further psi experiments with mice. *Journal of Parapsychology*, 32, 260.
- Edeal, G. (1950, 27th March). Why the choir was late. *Life*, pp.19–23.
- Eisenbud, J. (1976). Evolution and psi. *Journal of the American Society for Psychical Research*, 70(3), 35–53.
- Eisenbud, J. (1977). Paranormal photography. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp.414–432). London: MFarland.
- Evans, C. (1973). Parapsychology—what the questionnaire revealed. *New Scientist*, 57(830), 209.
- Evans, C. (1983). *Landscapes of the night: How and why we dream*. London: Victor Gollancz Ltd.
- Eysenck, H. J. (1975). Precognition in rats. *Journal of Parapsychology*, 39, 222–227.

- Fahler, J. (1957). ESP card tests with and without hypnosis. *Journal of Parapsychology*, 21, 179–185.
- Feather, S. R. & Schmicker, M. (2005). *The gift: ESP, the extraordinary experiences of ordinary people*. New York, NY: St. Martin's Press.
- Festinger, L. (1957). *A theory of cognitive dissonance*. London: Stanford University Press.
- Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1(3), 288–299.
- Flammarion, C. (1922). *Death and its mystery* (Vol. 2). New York: The Century Company.
- Ford, K. W. (1963). Magnetic monopoles. *Scientific American*, 209, 122–31.
- Forwald, H. (1952). A further study of the PK placement effect. *Journal of Parapsychology*, 19, 273–283.
- Frazier, K. (1980). Prediction, yes. Psychic, no! *The Skeptical Inquirer*, IV(3), 13–14.
- Freeman, J. & Nielsen, W. (1964). Precognition score deviations as related to anxiety levels. *Journal of Parapsychology*, 28(4), 239–249.
- French, C. C. (2003). Fantastic memories: The relevance of research into eyewitness testimony and false memories for reports of anomalous experiences. *Journal of Consciousness Studies*, 6(7), 153–174.
- French, C. C. & Wilson, K. (2006). Incredible memories: How accurate are reports of anomalous events? *European Journal of Parapsychology*, 21(2), 166–181.
- Gardner, M. (1981). *Science: Good, bad, and bogus*. New York, NY: Avon Books.
- Gardner, M. (1998). *The wreck of the Titanic foretold?* Buffalo, NY: Prometheus Books.
- Geer, J. H. (1966). Fear and autonomic arousal. *Journal of Abnormal Psychology*, 71(4), 253–255.
- Godley, J. (1950). *Tell me the next one*. London: Gollancz.
- Good, I. J. (1980). Scientific speculations on the paranormal and the parasciences. *Zetetic Scholar*, 7, 9–29.
- Goodnight, C. J. (1990). Experimental studies of community evolution I: The response to selection at the community level. *Evolution*, 44(6), 1614–1624.
- Green, C. (1960). Analysis of spontaneous cases. In *Proceedings of the Society for Psychical Research* (Vol. 53, pp.97–161).
- Gurney, E., Myers, F. W. H., & Podmore, F. (1918). *Phantasms of the living*. London: Trubner & co.
- Haidinger, W. (1844). Über das direkte erkennen des polarisierten lichts und der lage der polarisationsebene. *Annalen der Physik und Chemie*, 63, 29–39.
- Hansel, C. E. M. (1966). *ESP: A scientific evaluation*. New York, NY: Scribner.
- Hansen, G. P. (1990). Deception by subjects in psi research. *Journal of the American Society for Psychical Research*, 84(1), 25–80.
- Haraldsson, E. (1975). Reported dream recall, precognitive dreams, and ESP. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in parapsychology 1974* (pp.47–48). Metuchen, NJ: Scarecrow Press.
- Hardy, A. (1965). *The living stream: A restatement of evolution theory and its relation to the spirit of Man*. London: Collins.
- Hearne, K. M. T. (1982a). An ostensible precognition of the 1974 Flixborough disaster. *Journal of the Society for Psychical Research*, 51(790), 210–213.
- Hearne, K. M. T. (1982b). An ostensible precognition of the accidental sinking of H.M. submarine Artemis in 1971. *Journal of the Society for Psychical Research*, 51(791), 283–287.

- Hearne, K. M. T. (1982c). Three cases of ostensible precognition from a single percipient: 1. Sadat assassination 2. Reagan assassination attempt 3. S. S. Achille Laura incident. *Journal of the Society for Psychical Research*, 51(791), 288–291.
- Hearne, K. M. T. (1984). A survey of reported premonitions and of those who have them. *Journal of the Society for Psychical Research*, 52(796), 261–270.
- Hearne, K. M. T. (1986). An analysis of premonitions, deposited over one year from an apparently gifted subject. *Journal of the Society for Psychical Research*, 53(804), 376–382.
- Hearne, K. M. T. (1989). *Visions of the future*. Wellingborough: The Aquarian Press.
- Hintze, J. (2013). Pass 12 [Computer software manual]. Available from <http://www.ncss.com> (NCSS, LLC. Kaysville, Utah, USA)
- Hitchman, G. A., Roe, C. A., & Sherwood, S. J. (2012). A re-examination of nonintentional precognition with openness to experience, creativity, psi beliefs, and luck beliefs as predictors of success. *Journal of Parapsychology*, 76(1), 109–145.
- Hoebens, P. H. (1986). Sense and nonsense in parapsychology. In P. Kurtz (Ed.), *Science confronts the paranormal* (pp.28–42). Buffalo, NY: Prometheus Books.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.
- Honorton, C. (1970). Psi and internal attention states. In D. Shapin & L. Coly (Eds.), (pp.79–90). New York, NY: Parapsychology Foundation.
- Honorton, C. (1972). Reported frequency of dream recall and ESP. *Journal of the American Society for Psychical Research*, 66, 369–374.
- Honorton, C. (1977). Psi and internal attention states. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp.435–472). London: McFarland.
- Honorton, C. & Ferrari, D. C. (1989). Future telling: A meta-analysis of forced-choice precognition experiments, 1935–1987. *Journal of Parapsychology*, 53(28), 1–308.
- Honorton, C. & Stump, J. (1969). A preliminary study of hypnotically-induced clairvoyant dreams. *Journal of the American Society for Psychical Research*, 63, 175–184.
- Hoppe, R. B. (1987). A note on statistics for psientists. *Journal of the American Society for Psychical Research*, 81(3), 261–276.
- Horowitz, K. A., Lewis, D. C., & Gasteiger, E. L. (1975). Plant “primary perception”: Electro-physiological unresponsiveness to brine shrimp killing. *Science*, 189(4201), 478–480.
- Humphrey, N. K. (1996). *Leaps of faith*. New York: Basic Books.
- Humphrey, N. K. & Keeble, G. R. (1977). Do monkeys’ subjective clocks run faster in red light than in blue? *Perception*, 6(1), 7–14.
- Humphrey, N. K. & Keeble, G. R. (1978). Effects of red light and loud noise on the rate at which monkeys sample the sensory environment. *Perception*, 7(3), 343–348.
- Inglis, B. (1977). *Natural and supernatural: A history of the paranormal from earliest times to 1914*. London: Hodder and Stoughton.
- Inglis, B. (1990). *Coincidence*. London: Hutchinson.
- Jensen, A. R. & Rohwer Jr, W. D. (1966). The Stroop color-word test: A review. *Acta Psychologica*, 25(1), 36–93.
- Johnson, M. (1968). Relationship between dream recall and scoring direction. *Journal of Parapsychology*, 32, 56–57.
- Johnson, M. (1972). Letter to the editor. *Journal of Parapsychology*, 36, 71–72.
- Johnson, M. (1982). An attempt to select for psi ability in *Paramecium aurelia*. *Journal of the Society for Psychical Research*, 51(791), 272–282.

- Kammann, R. (1983). Uncanny prophecies in New Zealand: An unexplained scientific anomaly. *Zetetic Scholar*, 11, 15–21.
- Kammann, R. (1987). New Zealand prophecies exposed as a hoax. *Zetetic Scholar*, 12/13, 34–38.
- Katkin, E. S. & Hoffman, L. S. (1976). Sex differences and self-report of fear: A psychophysiological assessment. *Journal of Abnormal Psychology*, 85(6), 607–610.
- Kelly, E. & Kanthamani, K. (1972). A subject's efforts towards voluntary control. *Journal of Parapsychology*, 36, 185–197.
- Kennedy, J. E. (2004). What is the purpose of psi? *Journal of the American Society for Psychical Research*, 98, 1–127.
- Klintman, H. (1983). Is there a paranormal (precognitive) influence in certain types of perceptual sequences? I. *European Journal of Parapsychology*, 5(1), 19–49.
- Klintman, H. (1984). Is there a paranormal (precognitive) influence in certain types of perceptual sequences? II. *European Journal of Parapsychology*, 5(2), 125–140.
- Krippner, S., Honorton, C., & Ullman, M. (1971). A precognitive dream study with a single subject. *Journal of the American Society for Psychical Research*, 65, 192–203.
- Krippner, S., Honorton, C., & Ullman, M. (1972). A second precognitive dream study with Malcolm Bessent. *Journal of the American Society for Psychical Research*, 66, 269–279.
- Kumar, V. K., Pekala, R. J., & Gallagher, C. (1994). *The Anomalous Experiences Inventory*. Unpublished psychological test, Department of Psychology, West Chester University, West Chester, PA.
- Kurtz, P. (Ed.). (1986). *Science confronts the paranormal*. Buffalo, NY: Prometheus Books.
- Lachapelle, S. & Healey, J. (2010). On Hans, Zou and the others: Wonder animals and the question of animal intelligence in early twentieth-century France. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 41(1), 12–20.
- Lang, P. J., Bradley, M. M., Cuthbert, B. N., et al. (1999). International affective picture system (iaps): Instruction manual and affective ratings. *The center for research in psychophysiology, University of Florida*.
- Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English: Based on the British National Corpus*. London: Longman.
- Lépes, I. T. (1992). A psi effect with the *Drosophila melanogaster* fly. *Journal of the Society for Psychical Research*, 58(826), 181–188.
- Lépes, I. T. & Argibay, J. C. (1994). Possible psi effect with *Drosophila melanogaster*. *Journal of the Society for Psychical Research*, 60(837), 78–85.
- Levin, M. (1996). On the lack of evidence for the evolution of psi as an argument against the reality of the paranormal. *Journal of the American Society for Psychical Research*, 90(3), 221–230.
- Levy, W. J., Mayo, L. A., Andre, E., & McRae, A. (1971). Repetition of the French precognition experiments with mice. *Journal of Parapsychology*, 35, 1–17.
- Levy, W. J., Terry, J. C., & Davis, J. W. (1973). A precognition test with hamsters. *Journal of Parapsychology*, 37, 97–104.
- Lodge, O. (1909). *Survival of Man*. London: Methuen.
- Loftus, E. F. (1979). *Eyewitness testimony*. London: Harvard University Press.
- Lowe, W. L. (1980, March). Bad dreams in the future tense. *Playboy*, pp.102–108.
- Luke, D. P. (2011). Anomalous phenomena, psi and altered consciousness. In E. Cardeña &

- M. Winkelman (Eds.), *Altering consciousness: A multidisciplinary perspective* (Vol. 2, pp.355–374). Westport, CT: Praeger.
- Luke, D. P. (2012). Psychoactive substances and paranormal phenomena: A comprehensive review. *International Journal of Transpersonal Studies*, 31(1), 97–156.
- Luke, D. P., Delanoy, D., & Sherwood, S. J. (2008). Psi may look like luck: Perceived luckiness and beliefs about luck in relation to precognition. *Journal of the Society for Psychical Research*, 72, 193–207.
- Luke, D. P. & Kittenis, M. (2005). A preliminary survey of paranormal experiences with psychoactive drugs. *Journal of Parapsychology*, 69(2), 163–185.
- Luke, D. P. & Morin, S. (2009, September). *Luck beliefs, pmir, psi and the sheep-goat effect: A replication*. Paper presented at the 33rd International Conference of the Society for Psychical Research. University of Nottingham.
- Luke, D. P., Roe, C. A., & Davison, J. (2008). Testing for forced-choice precognition using a hidden task: Two replications. *Journal of Parapsychology*, 72, 133–154.
- MacKenzie, A. (1974). *Riddle of the future*. London: Arthur Baker Ltd.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109(2), 163–203.
- Mangan, G. L. (1955). Evidence of displacement in a precognition test. *Journal of Parapsychology*, 19, 35–44.
- Marks, D. F. & Kammann, R. (1980). *The psychology of the psychic*. Buffalo, NY: Prometheus Books.
- Markwick, B. (1978). The Soal-Goldney experiments with Basil Shackleton: New evidence of data manipulation. *Proceedings of the Society for Psychical Research*, 56, 250–277.
- McCauley, D. E. & Wade, M. J. (1981). The populational effects of inbreeding in *Tribolium*. *Heredity*, 46, 59–67.
- McCurdy, H. G. (1950). Consciousness and the galvanometer. *Psychological Review*, 57(6), 322–327.
- McMahan, E. A. & Bates Jr, E. K. (1954). Report of further Marchesi experiments. *Journal of Parapsychology*, 18, 82–92.
- McMahan, E. A. & Rhine, J. B. (1947). A second Zagreb-Durham ESP experiment. *Journal of Parapsychology*, 11, 244–253.
- Milton, J. & Wiseman, R. (1999). Does psi exist? Lack of replication of an anomalous process of information transfer. *Psychological Bulletin*, 125(4), 387–391.
- Milton, J. & Wiseman, R. (2001). Does psi exist? Reply to Storm and Ertel (2001). *Psychological Bulletin*, 127(3), 434–438.
- Mineka, S. (1992). Evolutionary memories, emotional processing, and the emotional disorders. In D. L. Medin (Ed.), *The psychology of learning and motivation* (pp.161–206). London: Academic Press Limited.
- Morris, R. L. (1967). Some new techniques in animal psi research. *Journal of Parapsychology*, 31, 316–317.
- Morris, R. L. (1977). Parapsychology, biology, and anpsi. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp.687–715). London: McFarland.
- Morris, R. L. (1982). Assessing experimental support for true precognition. *Journal of Parapsychology*, 46(4), 321–336.
- Mossbridge, J., Tressoldi, P. E., & Utts, J. (2012). Predictive physiological anticipation preceding seemingly unpredictable stimuli: a meta-analysis. *Frontiers in psychology*, 3(390).

- Mundle, C. W. K. (1964). Does the concept of precognition make sense? *International Journal of Parapsychology*, 6, 179–198.
- Murphy, G. (1960). Future research in precognition. *International Journal of Parapsychology*, 2(3), 5–20.
- Nash, C. B. (1960). Can precognition occur diametrically? *Journal of Parapsychology*, 24, 26–32.
- Nash, C. B. (1982). ESP of present and future targets. *Journal of the Society for Psychical Research*, 51(792), 374–377.
- Nickell, J. (2013). Oscar, the death-predicting cat. *The Skeptical Inquirer*, 34(4). Available from http://www.csicop.org/si/show/oscar_the_death-predicting_cat/?ref=nf
- Nicol, J. F. (1961). Apparent spontaneous precognition: A historical review. *International Journal of Parapsychology*, 3(2), 26–45.
- Öhman, A. & Soares, J. J. (1993). On the automaticity of phobic fear: Conditioned skin conductance responses to masked phobic stimuli. *Journal of Abnormal Psychology*, 102, 121–132.
- Orme, J. E. (1974). Precognition and time. *Journal of the Society for Psychical Research*, 47(760), 351–365.
- Orme, J. E. (1979). A note on the Nostradamus prophecies. *Journal of the Society for Psychical Research*, 50(781), 163–166.
- Orne, M. T. & Scheibe, K. E. (1964). The contribution of nondeprivation factors in the production of sensory deprivation effects: The psychology of the “panic button”. *The Journal of Abnormal and Social Psychology*, 68(1), 3–12.
- Osis, K. (1953). A test of the relationship between ESP and PK. *Journal of Parapsychology*, 17, 298–309.
- Osis, K. (1955). Precognition over time intervals of one to thirty-three days. *Journal of Parapsychology*, 19, 82–91.
- Pablos, F. de. (2004, September). *Rivastigmine induces precognitive dreams in normal subjects*. Paper presented at the 28th International Conference of the Society for Psychical Research. West Downs Conference Centre, Winchester.
- Palmer, J. (1977). Attitudes and personality traits in experimental ESP research. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp.175–201). London: McFarland.
- Park, T. (1962). Beetles, competition, and populations: An intricate ecological phenomenon is brought into the laboratory and studied as an experimental model. *Science*, 138(3548), 1369–1375.
- Park, T., Mertz, D. B., & Petruszewicz, K. (1961). Genetic strains of *Tribolium*: Their primary characteristics. *Physiological Zoology*, 34(1), 62–80.
- Park, T., Ziegler, J. R., Ziegler, D. L., & Mertz, D. B. (1974). The cannibalism of eggs by *Tribolium* larvae. *Physiological zoology*, 47(1), 37–58.
- Parker, A. (2005). Psi and altered states. In L. Storm & M. Thalbourne (Eds.), *Parapsychology in the 21st Century* (pp.65–89). Jefferson, NC: McFarland.
- Pearl, R., Park, T., & Miner, J. R. (1941). Experimental studies on the duration of life. XVI. Life tables for the flour beetle *Tribolium confusum*. *American Naturalist*, LXXV(756), 5–19.
- Peoc'h, R. (1994). Psychokinetic action of young chicks on the path of an illuminated source. *Journal of Scientific Exploration*, 9(2), 223–229.
- Pfungst, O. (1911). *Clever Hans (the horse of Mr. von Osten): A contribution to experimental animal and human psychology*. New York, NY: Holt, Rinehart and Winston.

- Pigliucci, M. (2010). *Nonsense on stilts: How to tell science from bunk*. London: University of Chicago Press.
- Prasad, J. & Stevenson, I. (1968). A survey of spontaneous psychical experiences in school children of Uttar Pradesh, India. *International Journal of Parapsychology*, 10, 241–261.
- Pratt, J. G. (1973). A decade of research with a selected ESP subject: An overview and reappraisal of the work with Pavel Stepanek. In J. Beloff (Ed.), *New directions in parapsychology*. London: Paul Elek.
- Prigatano, G. P. & Johnson, H. J. (1974). Autonomic nervous system changes associated with a spider phobic reaction. *Journal of Abnormal Psychology*, 83(2), 169–177.
- Pütz, P., Braeunig, M., & Wackermann, J. (2006). Eeg correlates of multimodal ganzfeld induced hallucinatory imagery. *International journal of psychophysiology*, 61(2), 167–178.
- Radin, D. I. (1997a). *The conscious universe*. New York, NY: HarperEdge.
- Radin, D. I. (1997b). Unconscious perception of future emotions: An experiment in presentiment. *Journal of Scientific Exploration*, 11(2), 163–180.
- Radin, D. I. (2002). A dog that seems to know when his owner is coming home: Effect of environmental variables. *Journal of Scientific Exploration*, 16(4), 579–592.
- Radin, D. I. (2003). Electrodermal presentiments of future emotions. In S. Wilson (Ed.), *Proceedings of the 46th annual convention of the Parapsychological Association*. Vancouver, Canada.
- Radin, D. I. & Bosworth, J. L. (1985). Response distributions in a computer-based perceptual task: Test of four models. *Journal of the American Society for Psychical Research*, 79(4), 453–483.
- Radin, D. I. & Bosworth, J. L. (1987). On statistics for “psientists” and critics. *Journal of the American Society for Psychical Research*, 81(3), 277–291.
- Radin, D. I. & May, E. (2000). Evidence for a retrocausal effect in the human nervous system. In F. Steinkamp (Ed.), *Proceedings of the 43rd annual convention of the Parapsychological Association* (pp.249–246). Raleigh, NC.
- Randall, J. L. (1976). *Parapsychology and the nature of life*. New York, NY: Harper & Row.
- Randi, J. (1982). *The truth about Uri Geller*. New York, NY: Prometheus Books.
- Randi, J. (1983a). The Project Alpha experiment: Part 1. The first two years. *The Skeptical Inquirer*, 7(4), 24–33.
- Randi, J. (1983b). The Project Alpha experiment: Part 2. Beyond the laboratory. *The Skeptical Inquirer*, 8(1), 36–45.
- Randi, J. (1986). The Project Alpha experiment. In P. Kurtz (Ed.), *Science confronts the paranormal* (pp.158–165). Buffalo, NY: Prometheus Books.
- Rattet, S. L. & Bursik, K. (2001). Investigating the personality correlates of paranormal belief and precognitive experience. *Personality and Individual Differences*, 31(3), 433–444.
- Rhine, J. B. (1934). *Extra-sensory perception*. Boston, MA: Bruce-Humphries.
- Rhine, J. B. (1938). Experiments bearing on the precognition hypothesis. *Journal of Parapsychology*, 2, 38–54.
- Rhine, J. B. (1941). Experiments bearing upon the precognition hypothesis. III. Mechanically selected cards. *Journal of Parapsychology*, 5, 1–57.
- Rhine, J. B. (1942). Evidence of precognition in the covariation of salience ratios. *Journal of Parapsychology*, 6, 111–143.
- Rhine, J. B. (1969). Position effects in psi test results. *Journal of Parapsychology*, 33, 136–157.

- Rhine, J. B. (1977). History of experimental studies. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp.25–47). London: McFarland.
- Rhine, J. B. & Humphrey, B. M. (1942). A confirmatory study of salience in precognition tests. *Journal of Parapsychology*, 6, 190–219.
- Rhine, J. B. & Pratt, J. G. (1957). *Parapsychology, frontier science of the mind: A survey of the field, the methods, and the facts of ESP and PK research*. Springfield, IL: Charles C. Thomas.
- Rhine, J. B. & Rhine, L. E. (1929a). An investigation of a “mind-reading” horse. *The Journal of Abnormal and Social Psychology*, 23, 449–466.
- Rhine, J. B. & Rhine, L. E. (1929b). Second report on Lady, the “mind-reading” horse. *The Journal of Abnormal and Social Psychology*, 24, 287–292.
- Rhine, L. E. (1954). Frequency of types of experience in spontaneous precognition. *Journal of Parapsychology*, 18, 93–122.
- Rhine, L. E. (1955). Precognition and intervention. *Journal of Parapsychology*, 19, 1–34.
- Rhine, L. E. (1962). Psychological processes in ESP experiences. I. Waking experiences. *Journal of Parapsychology*, 26, 88–111.
- Richards, A., French, C. C., Johnson, W., Naparstek, J., & Williams, J. (1992). Effects of mood manipulation and anxiety on performance of an emotional Stroop task. *British Journal of Psychology*, 83(4), 479–491.
- Ritchie, S. J., Wiseman, R., & French, C. C. (2012). Failing the future: Three unsuccessful attempts to replicate Bem’s ‘retroactive facilitation of recall’ effect. *PloS one*, 7(3), e33423.
- Robinson, D. (1980). A preliminary study of heredity and ESP scoring patterns in mice. , 44(1), 77.
- Roe, C. A., Sherwood, S. J., Farrell, L., Savva, L., & Baker, I. S. (2007). Assessing the roles of sender and experimenter in dream ESP research. *European Journal of Parapsychology*.
- Roe, C. A., Sherwood, S. J., Savva, L., & Baker, I. S. (2005). Assessing the role of sender and experimenter in dream ESP research. In L. Savva (Ed.), *Proceedings of the 48th annual convention of the Parapsychological Association* (pp.242–245). Petaluma, CA.
- Rogo, D. S. (1985). JB Rhine and the Levy scandal. In P. Kurtz (Ed.), *A skeptic’s handbook of parapsychology*. Buffalo, NY: Prometheus Books.
- Roll, W. G. (1961). The problem of precognition. *Journal of the Society for Psychical Research*, 41(709), 115–128.
- Ryan, M. F., Park, T., & Mertz, D. B. (1970). Flour beetles: Responses to extracts of their own pupae. *Science*, 170(3954), 178–180.
- Ryback, D. & Sweitzer, L. (1988). *Dreams that come true: Their psychic and transforming powers*. London: Doubleday.
- Saltmarsh, H. F. (1934). Report on cases of apparent precognition. *Proceedings of the Society for Psychical Research*, 42, 49–103.
- Savva, L. (Ed.). (2005). *Proceedings of the 48th annual convention of the Parapsychological Association*. Petaluma, CA.
- Savva, L., Child, R., & Smith, M. D. (2004). The precognitive habituation effect: An adaptation using spider stimuli. In S. Schmidt (Ed.), *Proceedings of the 47th annual convention of the Parapsychological Association* (pp.223–229). Vienna, Austria.
- Savva, L., Roe, C. A., & Smith, M. D. (2005). Further testing of the precognitive habituation effect. In L. Savva (Ed.), *Proceedings of the 48th annual convention of the Parapsychological Association* (pp.163–170). Petaluma, CA.
- Savva, L., Roe, C. A., & Smith, M. D. (2006). Further testing of the precognitive habituation

- effect using spider stimuli. *Journal of the Society for Psychical Research*, 70(885), 225–234.
- Schmeidler, G. R. (1970). High ESP scores after a swami's brief instruction in meditation and breathing. *Journal of the American Society for Psychical Research*, 64, 100–103.
- Schmeidler, G. R. (1994). Experiments 1978-1992: The glass is half full. In S. Krippner (Ed.), *Advances in parapsychological research 7* (pp.104–197). Jefferson, NC: McFarland.
- Schmeidler, G. R. & McConnell, R. A. (1957). *ESP and personality patterns*. Binghamton, NY: Yale University Press.
- Schmidt, H. (1969). Precognition of a quantum process. *Journal of Parapsychology*, 33, 99–108.
- Schmidt, H. (1970). PK experiments with animals as subjects. *Journal of Parapsychology*, 34, 255–261.
- Schmidt, S. (Ed.). (2004). *Proceedings of the 47th annual convention of the Parapsychological Association*. Vienna, Austria.
- Schmidt, S. & Walach, H. (2000). Electrodermal activity: State-of-the-art measurement and techniques for parapsychological purposes. *Journal of Parapsychology*, 64(2), 139–164.
- Schriever, F. (1987). A 30-year "experiment with time": Evaluation of an individual case study of precognitive dreams. *European Journal of Parapsychology*, 7(1), 49–72.
- Scott, C. & Haskell, P. (1973). General: "Normal" explanation of the Soal-Goldney experiments in extrasensory perception. *Nature*, 245(5419), 52–54.
- Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior Therapy*, 2(3), 307–320.
- Sheldrake, R. (1981). *A new science of life: The hypothesis of formative causation*. Los Angeles: JP Tarcher.
- Sheldrake, R. (2011). *Dogs that know when their owners are coming home and other unexplained powers of animals*. London: Arrow Books.
- Sheldrake, R. & Morgana, A. (2003). Testing a language-using parrot for telepathy. *Journal of Scientific Exploration*, 17(4), 601–616.
- Sheldrake, R. & Smart, P. (2000a). A dog that seems to know when his owner is coming home: Videotaped experiments and observations. *Journal of Scientific Exploration*, 14, 233–255.
- Sheldrake, R. & Smart, P. (2000b). Testing a return-anticipating dog, kane. *Anthrozoös*, 13, 203–212.
- Sherwood, S. J. (2002). Relationship between hypnagogic/hypnopompic states and reports of anomalous experiences. *Journal of Parapsychology*, 66, 127–150.
- Sherwood, S. J. (2012). A web survey of the content, sensory modalities, and interpretation of hypnagogic and hypnopompic experiences. *Journal of Parapsychology*, 76(1).
- Sherwood, S. J. & Roe, C. A. (2003). A review of dream ESP studies conducted since the Maimonides dream ESP programme. *Journal of Consciousness Studies*, 10, 85–109.
- Sherwood, S. J. & Roe, C. A. (2013). An updated review of dream ESP studies conducted since the Maimonides dream ESP programme. In S. Krippner, A. J. Rock, J. Beischel, H. L. Friedman, & C. L. Fracasso (Eds.), *Advances in parapsychological research 9*. Jefferson, NC: McFarland.
- Shimizu, H. (2002). Measuring keyboard response delays by comparing keyboard and joystick inputs. *Behavior Research Methods, Instruments, and Computers*, 34(2), 250–256.
- Sidgwick, E. M. (1888). On the evidence for premonitions. *Proceedings of the Society for Psychical Research*, 5, 288–354.
- Sinnock, P. (1970). Frequency dependence and mating behavior in *Tribolium castaneum*. *American Naturalist*, 104(939), 469–476.
- Skeat, W. W. (1882). *An etymological dictionary of the English language*. London: Macmillan.

- Smith, M. D. (2003). The role of the experimenter in parapsychological research. *Journal of Consciousness Studies*, 10, 6(7), 69–84.
- Smith, M. D. & Savva, L. (2004a). Experimenter effects and psi performance using a digital autoganzfeld system. In S. Schmidt (Ed.), *Proceedings of the 47th annual convention of the Parapsychological Association* (pp.461–463). Vienna, Austria.
- Smith, M. D. & Savva, L. (2004b, September). *Experimenter effects and psi performance using a digital autoganzfeld system*. Paper presented at the 28th International Conference of the Society for Psychical Research. West Downs Conference Centre, Winchester.
- Smith, W. R., Dagle, E. F., Hill, M. D., & Mott-Smith, J. (1963). *Testing for extrasensory perception with a machine* (Project 4610 No. AFCRL-63-141). Hanscom AFB, MA: Air Force Cambridge Research Labs.
- Soal, S. G. & Goldney, K. M. (1943). Experiments in precognitive telepathy. *Proceedings of the Society for Psychical Research*, 47, 21–150.
- Sokoloff, A. (1974). *The biology of Tribolium* (Vol. 2). Oxford: Oxford University Press.
- Sondow, N. (1988). The decline of precognized events with the passage of time: Evidence from spontaneous dreams. *Journal of the American Society for Psychical Research*, 82(1), 33–51.
- Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Spottiswoode, S. J. P. & May, E. (2003). Skin conductance prestimulus response: Analyses, artifacts and a pilot study. *Journal of Scientific Exploration*, 17(4), 617–641.
- Stanford, R. G. (1974a). An experimentally testable model for spontaneous psi events: I. Extrasensory events. *Journal of the American Society for Psychical Research*, 68(1), 34–57.
- Stanford, R. G. (1974b). An experimentally testable model for spontaneous psi events: II. Psychokinetic events. *Journal of the American Society for Psychical Research*, 68(4), 321–356.
- Stanford, R. G. (1977). Conceptual frameworks of contemporary psi research. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp.823–858). London: McFarland.
- Stanford, R. G. (1978). Toward reinterpreting psi events. *Journal of the American Society for Psychical Research*, 72(3), 197–214.
- Stanford, R. G. (1982). An experimentally testable model for spontaneous extrasensory events. In I. Grattan-Guinness (Ed.), *Psychical research: a guide to its history, principles and practices* (pp.195–205). Wellingborough: Aquarian Press.
- Stanford, R. G. (1990). An experimentally testable model for spontaneous psi events: A review of related evidence and concepts from parapsychology and other sciences. In S. Krippner (Ed.), *Advances in parapsychological research* 6 (pp.54–167). Jefferson, NC: McFarland.
- Stanford, R. G. & Stein, A. G. (1994). A meta-analysis of ESP studies contrasting hypnosis and a comparison condition. *Journal of Parapsychology*, 58, 235–269.
- Steinkamp, F., Milton, J., & Morris, R. L. (1998). A meta-analysis of forced-choice experiments comparing clairvoyance and precognition. *Journal of Parapsychology*, 62(3), 193–195.
- Stevenson, I. (1960). A review and analysis of paranormal experiences connected with the sinking of the Titanic. *Journal of the American Society for Psychical Research*, 54, 153–171.
- Stevenson, I. (1965). Seven more paranormal experiences associated with the sinking of the Titanic. *Journal of the American Society for Psychical Research*, 59, 211–225.
- Stevenson, I. (1970a). Precognition of disasters. *Journal of the American Society for Psychical Research*, 64(2), 187–210.
- Stevenson, I. (1970b). *Telepathic impressions*. Charlottesville: University Press of Virginia.
- Storm, L. & Ertel, S. (2001). Does psi exist? Comments on Milton and Wiseman's (1999)

- meta-analysis of ganzfeld research. *Psychological Bulletin*, 127(3), 424–433.
- Storm, L., Tressoldi, P. E., & Di Risio, L. (2010). Meta-analysis of free-response studies, 1992–2008: Assessing the noise reduction model in parapsychology. *Psychological Bulletin*, 136(4), 471.
- Stroop, J. R. (1935). Studies of interference in verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662.
- Suddendorf, T. & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, 30(03), 299–313.
- Szymanski, J. & O' Donohue, W. (1995). Fear of spiders questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 26(1), 31–34.
- Taleb, N. N. (2011). *The black swan: The impact of the highly improbable*. London: Random House.
- Targ, R., Cole, P., & Puthoff, H. (1974). *Development of techniques to enhance Man/machine communication* (Tech. Rep. No. NASA-CR-157886). Stanford, CA: Stanford Research Institute.
- Tart, C. T. (1977). Drug-induced states of consciousness. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp.500–525). London: McFarland.
- Tart, C. T. (1979). An emergent-interactionist understanding of human consciousness. In B. Shapin & L. Coly (Eds.), *Brain/mind and parapsychology* (pp.177–200). New York, NY: Parapsychology Foundation.
- Tart, C. T. (1983). Information acquisition rates in forced-choice ESP experiments: Precognition does not work as well as present-time ESP. *Journal of the American Society for Psychical Research*, 77(4), 293–310.
- Taylor, J. A. (1953). A personality scale of manifest anxiety. *Journal of Abnormal and Social Psychology*, 48(2), 285–290.
- Taylor, R. (2003). Evolutionary theory and psi: Reviewing and revising some need-serving models in psychic functioning. *Journal of the Society for Psychical Research*, 67, 1–17.
- Thalbourne, M. A. (2010). The Australian sheep-goat scale: Development and empirical findings. *Australian Journal of Parapsychology*, 10(1), 5.
- Thalbourne, M. A. & Delin, P. S. (1993). A new instrument for measuring sheep-goat variable: Its psychometric properties and factor structure. *Journal of the Society for Psychical Research*, 59(832), 172–186.
- Thouless, R. H. & Wiesner, B. P. (1946). On the nature of psi phenomena. *Journal of Parapsychology*, 10, 107–119.
- Thouless, R. H. & Wiesner, B. P. (1947). The psi process in normal and “paranormal” psychology. *Proceedings of the Society for Psychical Research*, 48, 117–196.
- Truzzi, M. (1987). Reflections on “Project Alpha”: Scientific experiment or conjuror’s illusion? *Zetetic Scholar*, 12(13), 73–98.
- Ullman, M. & Krippner, S. (1970). *Dream studies and telepathy*. New York, NY: Parapsychology Foundation.
- Ullman, M. & Krippner, S. (1973). *Dream telepathy*. London: Turnstone Books.
- van de Castle, R. L. (1977). Sleep and dreams. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp.473–499). London: McFarland.
- Vassy, Z. (1986). Experimental study of complexity dependence in precognition. *Journal of Parapsychology*, 50(3), 235–270.
- Wackermann, J., Pütz, P., & Allefeld, C. (2008). Ganzfeld-induced hallucinatory experience, its

- phenomenology and cerebral electrophysiology. *Cortex*, 44(10), 1364–1378.
- Wackermann, J., Pütz, P., Buchi, S., Strauch, I., & Lehmann, D. (2000). A comparison of ganzfeld and hypnagogic state in terms of electrophysiological measures and subjective experience. *Journal of Parapsychology*, 64(3), 254–255.
- Wade, M. J. (1976). Group selections among laboratory populations of *Tribolium*. *Proceedings of the National Academy of Sciences*, 73(12), 4604–4607.
- Wade, M. J., Shuster, S. M., & Stevens, L. (1996). Inbreeding: Its effect on response to selection for pupal weight and the heritable variance in fitness in the flour beetle, *Tribolium castaneum*. *Evolution*, 50(2), 723–733.
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., Kievit, R., & van der Maas, H. L. J. (2011, 18th February). *Yes, psychologists must change the way they analyse their data: Clarifications for Bem, Utts, and Johnson (2011)*. Retrieved June 24, 2013, from: <http://www.ruudwetzels.com/articles/ClarificationsForBemUttsJohnson.pdf>.
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, 100(3), 426–432.
- Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). Colour naming of phobia-related words. *British Journal of Psychology*, 77(1), 97–108.
- Weaver, W. (1963). *Lady Luck: The theory of probability*. London: Dover Publications.
- Webb, J. (1974). *The occult underground*. Chicago, IL: Open Court.
- Westerlund, J., Parker, A., Dalkvist, J., & Hadlaczky, G. (2006). Remarkable correspondences between ganzfeld mentation and target content—A psychical or psychological effect? *Journal of Parapsychology*, 70(1), 23–48.
- White, R. A. (1977). The influence of experimenter motivation, attitudes, and methods of handling subjects on psi test results. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp.273–301). London: McFarland.
- Wilson, G. D. (1967). GSR responses to fear-related stimuli. *Perceptual and Motor Skills*, 24(2), 401–402.
- Wilson, S. (Ed.). (2003). *Proceedings of the 46th annual convention of the Parapsychological Association*. Vancouver, Canada.
- Wiseman, R., Smith, M. D., & Milton, J. (2000). The “psychic pet” phenomenon: A reply to Rupert Sheldrake. *Journal of the Society for Psychical Research*, 64(858), 46–49.
- Wolman, B. B. (Ed.). (1977). *Handbook of parapsychology*. London: McFarland.
- Wood, G. H. & Cadoret, R. J. (1958). Tests of clairvoyance in a man-dog relationship. *Journal of Parapsychology*, 22, 29–39.
- Wood-Trost, L. (1981). Possible precognitions of the Teton Dam disaster. *Journal of the Society for Psychical Research*, 51, 65–74.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9, 1–27.
- Zarakhovich, Y. (2000, August). The Concordski. *Time*, 156(6). Available from <http://www.time.com/time/world/article/0,8599,2056192,00.html>
- Zohar, D. (1983). *Through the time barrier: A study in precognition and modern physics*. London: Paladin Books.
- Zusne, L. & Jones, W. H. (1989). *Anomalistic psychology: A study of magical thinking* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Author Index

- Akers, C., 29, 198
Alcock, J., 30, 198
Allefeld, C., 45, 211
Alvarez, F., 41, 198
Alvarez, L. W., 166, 198
Anderson, M., 33, 198
Andre, E., 42, 204
Argibay, J. C., 41, 204
Arnholt, A. T., 125, 126, 131, 198
Augustine, S., 10, 198
Averill, R. L., 45, 198
- Backster, C., 173, 198
Baker, I. S., 43, 208
Barker, J. C., 24, 25, 34, 35, 141, 146, 198
Bates Jr, E. K., 33, 205
Bechara, A., 62, 198
Beck, A. T., 112, 198
Beloff, J., 188, 198
Bem, D., 30, 198
Bem, D. J., 30, 31, 45, 58, 198, 199
Bender, H., 21, 199
Bestall, C. M., 172, 199
Besterman, T., 140, 199
Bierman, D. J., 61, 62, 64, 69, 71, 75, 199
Bishop, G. M., 140, 199
Blackmore, S. J., 34, 144, 147, 151, 199
Blagrove, M., 34, 199
Boake, C. R. B., 179, 199
Borsboom, D., 30, 212
Bosworth, J. L., 120–122, 130, 207
Bradley, M. M., 66, 204
Braeunig, M., 45, 207
Braud, L. W., 46, 199, 200
Braud, W. G., 44, 46, 118, 121, 122, 130–132, 199, 200
Broad, C. D., 13, 15, 19, 31, 44, 171, 200
Broughton, R. S., 45, 52, 53, 56, 59, 77, 199, 200
Buchi, S., 45, 212
Bursik, K., 34, 35, 207
- Cadore, R. J., 40, 46, 200, 212
Camfferman, G., 84–88, 92–96, 98, 99, 101, 103–107, 109, 116, 117, 200
Campbell, J. L., 145, 186, 200
Cardeña, E., 168, 200
Carpenter, J. C., 54, 200
Casler, L., 44, 200
Chambers, J. M., 125, 200
Champely, S., 69, 90, 132, 135, 200
Child, I. L., 43, 200
Child, R., 30, 208
Cicero, M. T., 13, 200
Cleveland, W. S., 125, 200
- Cohn, S. A., 35, 145, 186, 200
Cole, P., 120, 211
Corballis, M. C., 10, 53, 211
Cox, W. E., 21, 23, 27, 143, 169–172, 174, 175, 200, 201
Craig, J. G., 40, 171, 201
Cuthbert, B. N., 66, 204
- Dagle, E. F., 119, 210
Dalkvist, J., 76, 162, 201, 212
Damasio, A. R., 62, 198
Damasio, H., 62, 198
Darwin, C., 175, 201
Davis, J. W., 42, 204
Davison, J., 52, 205
Dawkins, R., 49, 184, 201
Delanoy, D., 52, 205
Delin, P. S., 153, 154, 156, 211
Diaconis, P., 150, 201
Di Risio, L., 45, 211
Dodds, E. R., 13, 201
Dosa, D. M., 39, 201
Dossey, L., 26, 59, 120, 201
Duggan, M., 31, 199
Dukhan, H., 44, 201
Dunne, J. W., 18, 19, 139–141, 201
Dutton, D., 41, 201
Duval, P., 173, 201
- Edeal, G., 25, 201
Eisenbud, J., 48, 49, 201
Erbaugh, J. K., 112, 198
Ertel, S., 45, 210
Evans, C., 142, 188, 201
Eysenck, H. J., 173, 201
- Fahler, J., 44, 202
Farrell, L., 43, 208
Feather, S. R., 148, 202
Ferrari, D. C., 29, 203
Festinger, L., 140, 202
Fischhoff, B., 162, 202
Flammarion, C., 166, 202
Ford, K. W., 165, 202
Forwald, H., 27, 202
Frazier, K., 148, 202
Freeman, J., 64, 202
French, C. C., 23, 31, 34, 110, 199, 202, 208
- Gallagher, C., 64, 204
Gardner, M., 119, 138, 169, 202
Gasteiger, E. L., 174, 203
Geer, J. H., 63, 202
Godley, J., 18, 202

- Goldney, K. M., 28, 126, 210
 Good, I. J., 171, 202
 Goodnight, C. J., 185, 202
 Gorsuch, R. L., 112, 210
 Green, C., 36, 202
 Gurney, E., 13, 202

 Hadlaczky, G., 162, 212
 Haidinger, W., 11, 202
 Hall, T. H., 145, 186, 200
 Hansel, C. E. M., 127, 128, 132, 202
 Hansen, G. P., 166, 202
 Haraldsson, E., 42, 202
 Hardy, A., 47, 48, 56, 202
 Hart-Davis, A., 144, 147, 199
 Haskell, P., 28, 209
 Healey, J., 38, 113, 204
 Hearne, K. M. T., 32–37, 54, 55, 142–146, 152, 202, 203
 Hill, M. D., 120, 210
 Hintze, J., 126, 203
 Hitchman, G. A., 52, 203
 Hoebens, P. H., 118, 203
 Hoffman, L. S., 63, 204
 Holm, S., 90, 137, 156, 157, 203
 Honorton, C., 29, 42–45, 139, 199, 203, 204
 Hoppe, R. B., 120, 121, 123, 138, 203
 Horowitz, K. A., 174, 203
 Humphrey, B. M., 27, 33, 208
 Humphrey, N. K., 55, 56, 117, 203

 Inglis, B., 11, 13, 42, 44, 142, 203
 Iribas, A. E., 168, 200

 Jacobs, G. A., 112, 210
 Jensen, A. R., 79, 203
 Johnson, H. J., 63, 207
 Johnson, M., 42, 172–174, 203
 Johnson, W., 110, 208
 Johnson, W. O., 30, 199
 Jones, G., 34, 199
 Jones, W. H., 166, 212

 Kammann, R., 84, 151, 152, 162, 166, 167, 204, 205
 Kanthamani, K., 139, 204
 Katkin, E. S., 63, 204
 Keeble, G. R., 117, 203
 Kelly, E., 139, 204
 Kennedy, J. E., 48, 204
 Kievit, R., 30, 212
 Kittenis, M., 46, 205
 Kleiner, B., 125, 200
 Klintman, H., 58, 59, 79, 80, 82–87, 90–94, 96–99, 101–105, 107, 109–111, 115–119, 195, 204
 Krippner, S., 42, 43, 204, 211
 Kumar, V. K., 64, 204
 Kurtz, P., 204

 Lachapelle, S., 38, 113, 204
 Lang, P. J., 66, 204
 Leech, G., 10, 204
 Lehmann, D., 45, 212
 Lépes, I. T., 41, 204

 Levin, M., 55, 56, 204
 Levy, W. J., 42, 204
 Lewis, D. C., 174, 203
 Lodge, O., 11, 204
 Loftus, E. F., 23, 204
 Lowe, W. L., 141, 147, 204
 Luke, D. P., 45, 46, 52, 204, 205
 Lushene, R. E., 112, 210

 MacKenzie, A., 146, 147, 205
 MacLeod, C. M., 80, 99, 113, 205
 Mangan, G. L., 27, 29, 205
 Marks, D. F., 84, 151, 152, 162, 166, 205
 Markwick, B., 28, 205
 May, E., 77, 81, 85, 86, 88, 96, 97, 99, 100, 107–109, 116, 207, 210
 Mayo, L. A., 42, 204
 McCauley, D. E., 176, 205
 McConnell, R. A., 82, 209
 McCurdy, H. G., 64, 205
 McKenna, F. P., 110, 212
 McMahan, E. A., 33, 205
 McRae, A., 42, 204
 Mendelson, M. L., 112, 198
 Mertz, D. B., 175, 179, 184, 206, 208
 Milton, J., 29, 41, 45, 205, 210, 212
 Mineka, S., 64, 205
 Miner, J. R., 186, 206
 Mock, J. E., 112, 198
 Montredon, E., 173, 201
 Morgana, A., 41, 209
 Morin, S., 52, 205
 Morris, R. L., 29, 40, 171–173, 205, 210
 Mossbridge, J., 77, 205
 Mosteller, F., 150, 201
 Mott-Smith, J., 120, 210
 Mundle, C. W. K., 21, 206
 Murphy, G., 21, 206
 Myers, F. W. H., 13, 202

 Naparstek, J., 110, 208
 Nash, C. B., 37, 206
 Nickell, J., 40, 206
 Nicol, J. F., 21, 206
 Nielsen, W., 64, 202

 O' Donohue, W., 66, 112, 124, 211
 Öhman, A., 63, 64, 206
 Orme, J. E., 33, 35, 57, 206
 Orne, M. T., 46, 206
 Osis, K., 33, 37, 206

 Pablos, F. de, 44, 206
 Palmer, J., 45, 82, 199, 206
 Park, T., 175, 179, 184–186, 206, 208
 Parker, A., 44, 45, 162, 206, 212
 Pearl, R., 186, 206
 Pekala, R. J., 64, 204
 Peoc'h, R., 40, 206
 Petruszewicz, K., 175, 206
 Pfungst, O., 38, 206
 Pigliucci, M., 39, 207
 Podmore, F., 13, 202

- Prasad, J., 21, 207
 Pratt, J. G., 34, 139, 207, 208
 Prigatano, G. P., 63, 207
 Puthoff, H., 120, 211
 Pütz, P., 45, 207, 211, 212
- Rabeyron, T., 31, 199
 Radin, D. I., 41, 58–64, 66, 69, 73, 75, 81, 85, 86, 88, 96, 97, 99, 100, 107–109, 116–118, 120–122, 130, 199, 207
 Randall, J. L., 48, 207
 Randi, J., 139, 151, 207
 Rao, K. R., 44, 201
 Rattet, S. L., 34, 35, 207
 Rayson, P., 10, 204
 Reijman, S., 168, 200
 Rhine, J. B., 13, 26, 27, 33, 34, 39, 45, 126, 127, 198, 205, 207, 208
 Rhine, L. E., 19–21, 32, 34, 36, 39, 44, 208
 Richards, A., 110, 208
 Ritchie, S. J., 30, 208
 Robinson, D., 41, 208
 Roe, C. A., 30, 43, 44, 52, 130, 192, 203, 205, 208, 209
 Rogo, D. S., 42, 208
 Rohwer Jr, W. D., 79, 203
 Roll, W. G., 27, 28, 208
 Ryan, M. F., 184, 208
 Ryback, D., 34, 208
- Saltmarsh, H. F., 16, 17, 19, 21, 36, 165, 208
 Savva, L., 30, 43, 130, 162, 196, 208, 210
 Scheibe, K. E., 46, 206
 Schmeidler, G. R., 36, 44, 82, 162, 188, 209
 Schmicker, M., 148, 202
 Schmidt, H., 27, 40, 119, 173, 209
 Schmidt, S., 71, 72, 209
 Schriever, F., 140, 209
 Scott, C., 28, 209
 Seligman, M. E. P., 63, 209
 Shafer, D., 121, 122, 130–132, 200
 Sharrock, R., 110, 212
 Sheldrake, R., 39, 41, 48, 209
 Sherwood, S. J., 42–44, 52, 192, 203, 205, 208, 209
 Shimizu, H., 102, 209
 Shuster, S. M., 176, 212
 Sidgwick, E. M., 14–16, 19, 21, 36, 209
 Sinnock, P., 177, 209
 Skeat, W. W., 21, 209
 Smart, P., 41, 209
 Smith, M. D., 30, 41, 130, 162, 186, 208, 210, 212
 Smith, W. R., 119, 210
 Soal, S. G., 28, 126, 210
 Soares, J. J., 63, 64, 206
 Sokoloff, A., 176, 183, 210
 Sondow, N., 140, 210
 Spielberger, C. D., 112, 210
 Spottiswoode, S. J. P., 77, 210
 Stanford, R. G., 44, 50–52, 54, 210
 Stein, A. G., 44, 210
 Steinkamp, F., 29, 37, 210
 Stevens, L., 176, 212
- Stevenson, I., 21–24, 31, 32, 35, 36, 44, 145, 146, 160, 185, 207, 210
 Storm, L., 45, 210, 211
 Strauch, I., 45, 212
 Stroop, J. R., 79, 211
 Stump, J., 44, 203
 Suddendorf, T., 10, 53, 211
 Sweitzer, L., 34, 208
 Szymanski, J., 66, 112, 124, 211
- Taleb, N. N., 24, 211
 Targ, R., 120, 211
 Tart, C. T., 37, 46, 120, 211
 Taylor, J. A., 64, 211
 Taylor, R., 53–55, 211
 Terry, J. C., 42, 204
 Thalbourne, M. A., 153, 154, 156, 211
 Thouless, R. H., 12, 50, 211
 Tranel, D., 62, 198
 Tressoldi, P. E., 31, 45, 77, 199, 205, 211
 Treurniet, W., 40, 171, 201
 Trezise, L., 110, 212
 Troscianko, T., 34, 199
 Truzzi, M., 151, 211
 Tukey, P. A., 125, 200
- Ullman, M., 42, 43, 204, 211
 Utts, J., 30, 77, 199, 205
- Vagg, P. R., 112, 210
 van de Castle, R. L., 43, 211
 van der Maas, H. L. J., 30, 212
 Vassy, Z., 123, 211
- Wackermann, J., 45, 207, 211, 212
 Wade, M. J., 176, 179, 184, 205, 212
 Wagenmakers, E. J., 30, 212
 Walach, H., 71, 72, 209
 Ward, C. H., 112, 198
 Watts, F. N., 110, 112, 113, 212
 Weaver, W., 10, 26, 212
 Webb, J., 11, 212
 Westerlund, J., 76, 162, 201, 212
 Wetzels, R., 30, 212
 White, R. A., 186, 212
 Wiesner, B. P., 12, 50, 211
 Williams, C., 41, 201
 Williams, J., 110, 208
 Wilson, A., 10, 204
 Wilson, G. D., 63, 212
 Wilson, K., 23, 202
 Wilson, S., 212
 Wiseman, R., 30, 41, 45, 205, 208, 212
 Wolman, B. B., 40, 212
 Wood, G. H., 40, 212
 Wood-Trost, L., 146, 212
- Zajonc, R. B., 58, 212
 Zarakhovitch, Y., 164, 212
 Ziegler, D. L., 179, 206
 Ziegler, J. R., 179, 206
 Zohar, D., 39, 146, 212
 Zusne, L., 166, 212

TRI Appendix

A.1 Summary of Questionnaire Data

This section contains summary results for the questionnaires completed as part of TRI III as outlined in Section 3.3.2.3. This data was not used in the analysis, mainly because of the lack of significant paranormal results, and is included here for completeness.

Table A.1: Summary results for the Spielberger State-Trait Anxiety Questionnaire, administered before participation

Min	1st Qu.	Median	Mean	3rd Qu.	Max
21	30.2	34.5	36.6	40.8	77

Table A.2: Summary results for the Spielberger State-Trait Anxiety Questionnaire, administered after participation

Min	1st Qu.	Median	Mean	3rd Qu.	Max
21	30	35	36.6	41	69

Table A.3: Summary results for the Spielberger Trait Scale, administered after participation

Min	1st Qu.	Median	Mean	3rd Qu.	Max
20	35	40	42	48	66

Table A.4: Summary results for the Beck Depression Inventory, administered after participation

Min	1st Qu.	Median	Mean	3rd Qu.	Max
0	3.0	6.5	8.3	11	31

DM Appendix

B.1 Booklet Instructions

These are the instructions as presented to the participants of the DM study:

For many decades, David Mandell has been having vivid dreams which he feels offer him glimpses into the future, often relating to disasters, terrorist attacks or gruesome crimes. Throughout history, such precognitive dreams have been recorded. Because he is an artist, David took to making paintings or drawings of the scenes represented in his dreams and he took steps to obtain independent proof of the time and date when the pictures had been made.

Is it possible that David really does have some kind of psychic gift that allows his sleeping mind to witness events that have yet to take place? Perhaps. It is also possible that, given the large number of drawings and paintings that he has made, some of the pictures are bound to match some future event just by chance alone. It is the latter possibility that we are investigating in this current study.

You will be presented with 40 dream-pictures depicting events that David feels have already come to pass. You will be presented with two brief descriptions of events that possibly match the pictures. One will be the actual event that David feels he predicted; the other will be an event selected from news archives to also be a reasonable match to the picture. You will not always know which is which. Your task will be simply to rate the degree to which you think the picture matches the event.

The copies of the pictures have not been edited in any way other than to remove the date on which they were produced. During the task, you may remove the picture from its plastic folder to examine it but please ensure that you put it back in the correct folder (all folders and pictures are numbered).

As you will see, some of the pictures are paintings and some are drawings. Some are quite realistic depictions of events, others are more abstract. Some pictures have David's notes written on them, some do not. Some of the interpretations are quite literal, others are symbolic. Your task is to simply make a global judgement of the degree of match between the events and the pictures in terms of how convincing you find the pictures to be. There are no right or wrong answers. We simply want *your* opinion—so please do not discuss the pictures with others.

Take as long as you like over each picture. You may feel that one event is clearly a better match to the picture than the other or that they both match to a similar extent. In the latter case, you may feel that there is a high degree of match or very little correspondence. Whatever you decide, please ensure that ratings awarded reflect your judgement accurately. Finally, it is possible that David Mandell is simply a clever con-artist. We have no direct evidence to suggest that this is the case. For this purposes of the present exercise, please complete the task on the assumption that David is not a hoaxer. Thank you for agreeing to take part in this study. We hope you find it interesting. The experiment will probably be included in a Channel 5 documentary on David Mandell.

Please work your way through the set of pictures at your own pace. Do not discuss the pictures with anyone else. For each picture, judge the degree to which you think it matches the events described by circling one number only on the rating scale.

B.2 Rating Task

This section provides the information regarding DM's specific claims and provides the written details regarding DM's events and the alternative events (as presented within the testing booklet). Abbreviations used are: DoP for Date of Prediction; BP indicates if there is a bank photograph and when it was taken; C refers to cutting and indicates the newspaper issue where a corresponding cutting can be found. Where the difference between the rating for DM's event and the ALT event was significant, a copy of the picture is reproduced in Section B.3. A digital appendix to this section can be found at <https://sites.google.com/site/dmdigitalarchive> where the reader can find copies of DM's pictures, along with examples of the bank-photographs and other supplementary material.

Picture 1: "Marchioness Disaster"

DoP: 01/02/88 | BP: 14:24, 01/02/88 | C: *The Times*: 11/06/96
Not used because original painting could not be found.¹

Picture 2: "Hammer attack"

DoP: 11/03/89 | BP: No. | C: *The Times*: 11/06/96

- DM: 10th July 1996: The three female members of the Russell family were attacked with a claw hammer in a country lane near Chillenden, near Canterbury, Kent. Only Josie, aged 9, survived. Both Josie and her mother were blindfolded and suffered severed blows to the head. Meagan, aged 6, was hit at least seven times. The words *hammer* and *hammering* appear in the text.
- ALT: 27th September 1986: Donna Jester (aged 37), her blind cousin Dalpha (aged 64), and Laura Lee Owens (aged 20) were discovered dead in Lancaster, Texas. All three died as a result of numerous chopping wounds to their heads and faces, which were inflicted with a hatchet. The text refers to *three women*.

Picture 3: "San Francisco Storm"

DoP: 13/05/89 | BP: No. | C: *The Guardian*: 19/10/89

- DM: 17th October 1989: San Francisco Oakland earthquake. The suspension bridge is similar in its "4 sections" construction details to the Golden Gate Bridge. More than 270 people died and hundred were injured in the earthquake. David Mandell reported in a letter to a newspaper that the picture refers to an earthquake on the West Coast of America "later this year".
- ALT: 3rd December 1983: High winds closed the Golden Gate Bridge, San Francisco, for the longest period in its history: 3 hours and 27 minutes. Wind gusts reached 75 miles per hour, but the bridge suffered no structural damage. Text on the picture mentions the bridge shuddering with a storm effect.

Picture 4: "Soldiers"

DoP: 24/06/90 | BP: No. | C: *The Times*: 13/08/90

- DM: 7th August 1990: Symbolic representation of "Operation Desert Shield", the American defensive military operation in Saudi Arabia and the Gulf. As a symbol, it should obviously not be matched to a news photograph of the event.

¹Included for consistency with numbering.

- ALT 1969 to date: Symbolic representation of British soldiers in Northern Ireland, first sent in response to an upsurge in sectarian violence in 1969. As a symbol, it should obviously not be matched to a news photograph of the event.

Picture 5: “Floods in Europe”

DoP: 16/04/91 | BP: no | C: *The Times*: 27/12/93

- DM: 21–31 December 1993: In Germany, France, Belgium and the Netherlands, the worst flooding in 60–100 years took place. River water poured into many riverside towns.
- ALT: 27–30 December 1998: Over a period of four days, rivers in Oregon, Washington, USA, were severely flooded. The number of known dead was undetermined, but 70 people were displaced.

Picture 6: “Kosovo: refugee camp forced evacuation”

DoP: 05/07/91 | BP: no | C: *Daily Mail*: 08/04/99

- DM: 8th April 1999: Forced evacuation of a refugee camp in Kosovo. The red crosses are consistent with news reports. The picture is said to correspond closely with the attached newspaper photograph.
- ALT: Autumn 2001: One hundred and sixty seven foreign nationals were evacuated from the northern border town of Peshawar, in the aftermath of the US attack on Afghanistan. All of them were taken by road to Islamabad.

Picture 7: “Heathrow mortar attack”

DoP: 24/07/91 | BP: 13:41, 27/07/91 | C: *The Independent*: 10/03/94 | Figure B.2

- DM: 9th March 1994: Heathrow mortar attack. The IRA threw air services over London into chaos with mortar attacks on Heathrow airport. The burnt cars in the drawing are similar to those features in the attached news photograph. The words *possibly Heathrow* also appear in the text.
- ALT: August 27th 2001: A car bomb exploded in a parking garage at Madrid’s Barajas Airport, 10 miles east of down-town Madrid, causing damage but no injuries, in an attack apparently carried out by the armed Basque separatist group ETA, police said.

Picture 8: “Dream of Gigantic Seas—Langland Bay”

DoP: 13/09/91 | BP: no | C: *The Times*: 30/09/91

- DM: September 1991: Severe storms in Langland Bay, Wales. The words “Dream of gigantic seas in Langland Bay” appear in the picture. In the top left-hand corner is a ship that was blown aground.
- ALT: October 2000: Horrific storms in the West-Country and South Wales—the worst in a decade—left many parts of these areas underwater.

Picture 9: “Dead fish in River Irfon, Wales”

DoP: 27/10/91 | BP: no | C: Information from Wye foundation | Figure B.2

- DM: November 1995: A chemical spill occurred in the river Elan, Wales. The spill killed all the fish along a 20-mile stretch. The words *River Irfon* appear on the picture. The river *Irfon*² is connected to the River Elan.
- ALT: Summer 1994: A pollution disaster occurred in the Hereford area of the west Midlands, causing a catastrophic loss of fish.³

Picture 10: “Earthquake, Turkey”

DoP: 23/11/91 | BP: no | C: *The Times*: 02/10/95 & 03/10/95

²Although this is the phrasing in the testing booklet, closer examination shows that the river name is either Irfun or Irfon.

³Although not noted in the testing booklet, this river is the Wye and the river Elan, mentioned in DM’s event account, is a tributary of the Wye.

- DM: 1st October 1995: A strong earthquake, measuring 6.1 on the Richter scale, hit the city of Dinar, Turkey, causing the deaths of between 90 and 100 people, and wounding between 230 and 270. A newspaper picture is said to correspond to the painting.
- ALT: 17th August 1999: A devastating earthquake (magnitude 6.8 Mj) hit Izmit in Turkey's heavily populated north-west. It left some 17,000 dead and thousands more homeless.

Picture 11: "Baltic Exchange Bomb"

DoP: 29/02/92 | BP: 12:33, 31/03/92 | C: *Daily Mail*: 13/04/92

NB: Accounts refer only to the picture on right half of page.

- DM: 13th April 1992: Three killed when IRA bomb exploded outside the Baltic Exchange, near to the spot shown in the picture.
- ALT: April 1993: An IRA truck bomb devastated the Bishopsgate area of London's financial centre, killing one and injuring 44, as well as causing damage running into hundreds of millions of pounds, near to the spot shown in the picture.

Picture 12: "Neftegorsk Earthquake"

DoP: 10/04/92 | BP: no | C: *Daily Telegraph*: 31/05/95

- DM: 28th May 1995: Neftegorsk earthquake, Russia. The aerial perspective of the picture shows three shapes, resembling the perspective and the shapes formed by the rubble in the attached news photograph.
- ALT: 4th October 1994: A major earthquake occurred near Shikotan Island, in the South Kuril Island Group of Russia's Far East, causing several landslides. The picture represents the desolate landscape post-quake.

Picture 13: "Los Angeles Earthquake"

DoP: 16/05/92 | BP: no | C: *Daily Telegraph* & *The Times*: 18/01/94

- DM: 17th January 1994: A violent earthquake killed at least 24 people when it struck the San Fernando valley, north west of Los Angeles. It was the worst quake in California for forty years. Highways were turned to rubble, thousands of homes were wrecked, and hundred of fires were sparked. It is claimed that there is correspondence between the picture and the attached news photograph.
- ALT: 27th March 1964: One of the largest earthquakes (and ensuing tsunami) in the USA occurred at Prince William Sound, Alaska, taking 125 lives and causing millions of pounds in property loss. Anchorage, about 120 km north-west of the epicentre, sustained the most severe damage to property.

Picture 14: "Tank convoy"

DoP: 06/06/92 | BP: no | C: *The Times*: 01/07/92

- DM: 2nd July 1992: The UN peacekeeping effort in the Bosnian capital, Sarajevo, was to be bolstered with the arrival of a battalion of Canadian troops, but the convoy was split into two before being reunited.
- ALT: April 2002: A tank brigade made up of reservists rolled into Nablus, West Bank, Israel, from two directions, sealing off the city and its satellite refugee camps. Dozens of bombs were detonated in the paths of the tanks.

Picture 15: "Indian earthquake"

DoP: 08/06/92 | BP: 16:11, 01/10/93 | C: *The Guardian*: 01/10/93 | Figure B.3

- DM: 30th September 1993: Indian earthquake. 10,000 people died in India's worst earthquake for 50 years. The word *India* is at the top of the picture. The shock measured 6.5 on the Richter scale, shattering 50 towns and villages in Maharashtra state, east of Bombay. A column in the enter of the picture is almost identical to a column in the news photograph. It is claimed that there is a correspondence between the picture and the attached news photograph.

- ALT: 26th January 2001: Indian earthquake: More than 1000 people were killed in a powerful earthquake that struck Western India and parts of Pakistan. The word *India* is at the top of the picture. The shock measured 6.9 to 7.9 on the Richter scale. Nearly all the victims were from the Indian town of Bhuj, in the state of Gujarat, and the nearby city of Ahmedabad.

Picture 16: "Power failure on the London Underground"

DoP: 27/09/1992 | BP: no | C: *The Times*: 25/11/93

- DM: 24th November 1993: Power failure on the Central Line, London Underground. Reference is made to High Holborn, which is a station on the Central Line.
- ALT: 28th November 1996: Power failure occurred whilst engineers were testing a compressor at Lots Road power station. 30,000 people were reported to have been trapped underground for three hours.

Picture 17: "Braer tanker"

DoP: 12/12/92 | BP: 14:48, 15/12/92 | C: *The Daily Express*: 07/01/93

- DM: 1st June 1993: The Braer tanker sank off the Shetland Islands. 85,000 tonnes of Norwegian light crude spilled into the sea. The funnel, white masts and the triangular shape of the bow bear some resemblance to those seen in photographs of the Braer. It is claimed that there is correspondence between the picture and the attached news photograph.
- ALT: 16th February 1996: The Sea Empress sank off the coast of South West Wales, causing the largest oil spill in the region, involving the release of approximately 70,000 tonnes of North Sea light crude.

Picture 18: "Airliner crash on boulders"

DoP: 24/05/93 | BP: 15:04, 24/05/93 | C: *The Times*: 28/05/93 & 06/06/93

- DM: 27th May 1993: Represents the crash of a Lockheed RAF C-130 Hercules Transport into a Perthshire hillside, Scotland. The upturned tail of the plane looks like that of a Hercules plane. All nine people on board were killed.
- ALT: 21st December 1988. Pan Am flight 103 exploded and pieces of the plane fell onto the Scottish town of Lockerbie, killing 259 people on the plane and 11 people on the ground.

Picture 19: "Heavy rain and floods—Dave Armitage"

DoP: 06/93 | BP: no | C: *The Times*: 14/10/93

- DM: August/September 1993: The south and west of Britain experienced some of the worst weather since the great storm of October 1987. Hundreds of people left their homes as they were flooded.
- ALT: October/November 2000: Wales was hit by major storms. The Meteorological office indicated that the rainfall recorded during the autumn across England and Wales was the highest in over 200 years. Over 1900 separate incidents of properties flooded were identified.

Picture 20: "Bubonic plague scare, Heathrow"

DoP: 03/10/93 | BP: no | C: *The Times*: 27 & 28/09 & 03/10 1994

- DM: September/October 1994: During the bubonic plague epidemic in India, airports across Europe imposed immigration and travel restrictions from India. On 3rd of October, ambulances were sent to Heathrow to meet a plane carrying plague sufferers.
- ALT: May 8th 2002: Twelve French men, wounded when their bus was blown up in a deadly suicide bomb attack in Karachi, Pakistan, were taken to the airport where a flying hospital was to take them to Paris.

Picture 21: "Tokyo Gas Attack"

DoP: 17/11/93 | BP: no | C: no

- DM: 20th March 1995: The Tokyo subway sarin gas attack occurred, killing nearly a dozen people and incapacitating or injuring approximately 5,000 others. Note the reference to information packs and leaflets. The poison gas was left in container wrapped up in newspapers and magazines. The attack was carried out by the Aum Shinrykio cult. The words *secret worship by rebel terrorist group* and *secret worship group or terrorist group* appear in the picture.
- ALT: 18th December 1997: At least 50 passengers on a packed commuter train in Tokyo were taken to hospital after a gang of pickpockets sprayed tear gas to escape undercover police. The group of about seven thieves, some armed with knives, fled at Ikebukuro station. Passengers were treated for eye and throat complaints.

Picture 22: "St. Malo ferry"

DoP: 05/01/94 | BP: 05:01:94 | C: *The Guardian*: 18/04/95

- DM: 17th April 1995: The St. Malo Ferry disaster. The boat, with 307 people on board, apparently hit a rock off the Jersey coast. Sixty-five people were injured. There were orange squares on the side of the ship and funnel.
- ALT: 6th March 1987: The Herald of Free Enterprise capsized about 90 seconds after leaving the harbour, ending on her side half-submerged in shallow water. At least 150 passengers and 38 members of the crew lost their lives, most inside the ship.

Picture 23: "Boy in river"

DoP: 31/01/94 | BP: 13/4/94 | C: *Daily Mail*: 11/08/95

- DM: 11th August 1995: The event involved a mother drowning after successfully rescuing her four-year old son who had fallen into a river at Curzon Park near Calne, Wiltshire.
- ALT: 18th March 2002: Four-year old boy drowns on a boating holiday in Upton upon Severn Marina in Worcestershire while his family slept.

Picture 24: "Fred West suicide"

DoP: 11/03/94 | BP: 13/4/94 | C: *Daily Mail*: 11/08/95

- DM: 1st January 1995: Fred West, the man accused of the Gloucester "House of Horrors" killings, hangs himself in his cell with strips of his prison bed-sheets.
- ALT: 6th February 1999: Simon Barker, 25, hanged himself in his prison cell in Doncaster jail. This controversial prison, awarded a charter mark for the care of its inmates, is facing a fresh investigation after this fifth death at the prison in one year alone.

Picture 25: "Dublin Bomb"

DoP: 19/04/94 | BP: no | C: *The Times & Daily Telegraph*: 23/05/94

- DM: 23rd May 1994: Dublin bomb. Two explosions relate to the gunfire and the bomb which flashed but failed to detonate properly. One man died and another was wounded.
- ALT: 17th May 1974: Dublin bomb. Two explosions relate to the two locations of the bombs—Dublin and Monaghan. Thirty people died that day.

Picture 26: "Crash Plane USA Whitehouse"

DoP: 15/07/94 | BP: no | C: *The Times & Washington Post*: 13/09/94

- DM: 12th September 1994: Plane crashing into the White House. A student pilot stole the single-engine Cessna from the airfield north of Baltimore and died in the crash. The words *Crash plane USA White House* appear at the top of the picture.

- ALT: 11th September 2001: Plane crashing into the Pentagon. Top government officials suggested that the American airliner was originally headed for the White House, although the recorded flight path shows otherwise.

Picture 27: "Oxford St. riots"

DoP: 23/09/94 | BP: no | C: *The Times*: 10/10/94

- DM: 9th October 1994: Oxford Street riots. The picture shows baton-wielding police with shields, broken glass, and a rioter with an eye patch.
- ALT: 1st May 2001: May Day Riots. Rocks, bottles and paving slabs were thrown at the police lines. Many police officers wore riot gear.

Picture 28: "Rackhams knife attack"

DoP: 23/10/94 | BP: 25/12/94 | C: *Evening Standard*: 08/12/94 | Figure B.4

- DM: 8th December 1994: Fourteen people were stabbed by a man wielding two knives in Rackhams department store in Birmingham. The man randomly attacked staff and shoppers, slashing at their throats and faces. The words *Edgebaston* and *throwing knives* appear in the picture.
- ALT: October 2000. A police and Travel West Midlands scheme to cut crime on buses in Birmingham has scooped a national community safety award. Operation Hackney found bus users with guns, knives and drugs, which led to around 250 arrests.

Picture 29: "Huge tanker in storm, Wales"

DoP: 22/10/95 | BP: no | C: *The Guardian*: 17/02/96

- DM: 16th February 1996: The Sea Empress, a super-tanker, ran aground in Wales in high winds.
- ALT: December 1999: The coastal tanker Blackfriars, in ballast, but carrying 180 tonnes of diesel oil, ran aground on sand off the West Wales coast in heavy weather.

Picture 30: "Kobe earthquake"

DoP: late/94 | BP: no | C: *The Independent*: 18/1/95

- DM: 17th January 1995: An earthquake of magnitude 7.2 Mj struck the region of Kobe and Osaka in south-central Japan. The elevated railroad structures were particularly hard hit. Nearly 5,500 deaths were confirmed, and roughly 35,000 people were injured. Strong ground shaking lasted for about 20 seconds and caused severe damage over a large area.
- ALT: 24th March 2001: The Geiyo earthquake (magnitude 6.9 Mj) struck Hiroshima, Japan, ca. 430 miles south-west of Tokyo. Ground shaking was intense in certain localised areas. Two people were reported dead, and nearly 200 injured.

Picture 31: "British soldier beaten to death in Bosnia"

DoP: 14/05/1996 | BP: no | C: Information from army press office

- July 1997: A British soldier was beaten to death after a charity race, in Bosnia, by local youths. The text reads "dream of mob kicking and killing British soldiers in Bosnia".
- September 2001: A Protestant was viciously beaten by a mob of Nationalists in north Belfast. They kicked and punched the man, George Courtney, before dragging him along the ground.

Picture 32: "Watford train crash"

DoP: 07/08/96 | BP:16:35, 08/08/96 | C: *Mirror & The Times*: 09/08/1996

- DM: 8th August 1996: Watford train crash. One woman died in the crash; one woman is shown in the picture. Sixty-seven people were injured. The left-hand side of the painting could represent the underside of the upturned carriage, thrown up in the air after the collision.

- ALT: 28th February 2001: Selby Train Crash. The picture shows a car in the middle-top part of the picture. Thirteen people were confirmed dead, and 70 people were injured when a passenger train hit a car, then collided with a freight train on the East Coast main line in North Yorkshire. The picture depicts events before the train hit the car and subsequently hit the other train in the blood-red distance.

Picture 33: "Twin Towers sketch"

DoP: 11/09/96 | BP:16:19, 11/09/96 | C: Various 11/09/01 | Figure B.5

- 11th September 2001: Collapse of the Twin Towers, New York. The Statue of Liberty is represented on the right, with its prominent pronged head-dress. The smaller skyscraper with its distinctive pointed top, shown in news photographs, is on the left. Note that the left tower collapses into the right.
- 5th September 1944: Earthquake in Massena Center, New York State. The effects of the quake were felt as far north as James Bay, Canada, and as far south as the state of Virginia. Numerous buildings were damaged, mostly in Massena Center.

Picture 34: "Diego Garcia Bomber Crash"

DoP: 16/09/96 | BP:16:46, 16/09/96 | C: *The Times*: 13/12/01

- DM: 12 December 2001: An Air Force B1 bomber crashed in the Indian Ocean shortly after taking off from its base on the island of Diego Garcia. The crew was rescued by helicopters. The sketch of King George VI refers to the island being given to the Americans. David Mandell is quoted in a German Newspaper before the crash as saying the the dream refereed to Diego Garcia.
- ALT: 17th May 2002: A Tornado GR4 aircraft from RAF Marham crashed into the River Humber, East Yorkshire. RAF rescue helicopters were called to the scene. None of the crew were seriously injured. The sketch of King George VI refers to the scene of the crash being in England.

Picture 35: "Twin Towers colour painting"

DoP: 27/03/97 | BP:16:18, 27/03/97 | C: Various 11/09/01 | Figure B.6

- DM: 11th September 2001: Collapse of the Twin Towers, New York. The statue of liberty is represented on the right, with its prominent pronged head-dress. The smaller sky scraper with its distinctive pointed top, shown in news photographs, is on the left. Note that the left tower collapses into the right.
- ALT: 18th April 2002: A small plane slammed into a landmark skyscraper in Milan, killing at least four people and injuring 60. Smoke poured out of the building, but fire-fighters quickly put out the flames.

Picture 36: "Death of Princess Diana"

DoP: 20/05/97 | BP:16:48, 27/05/97 | C: *The Times*: 31/08/97

- DM: 30th August 1997: The image relates to the death of Princess Diana in a car crash in a Parisian tunnel. There is fire in the painting, which symbolises great danger. The words *Di in tunnel* appear in the text.
- ALT: November 2000: 155 people die after a faulty radiator sparks a blaze in a tunnel in the Austrian Alps.

Picture 37: "Two miners killed"

DoP: 29/05/97 | BP:16:56, 03/06/97 | C: *Daily Mail* 05/06/98 | Figure B.7

- DM: 4th June 1998: The picture shows two coalminers killed in a mine cave-in accident. The roof collapsed without warning. The words *collapsed coal face* and *pit prop broken* appear in the picture.
- ALT: September 1995: The picture depicts the deaths of two miners killed in a rockfall at AngloGold's Nologwa Mine (formerly Vaal Reefs) in Klerksdorp, South Africa.

Picture 38: "Concorde Crash"

DoP: 04/06/97 | BP:11:30, 12/06/97 | C: BBC news 25/07/00 | Figure 5.1

- DM: 25th July 2000: Concorde crashed in France, on Air France flight 4590. Just before impact, the control tower told the pilot, "you have flames behind you." Note the French Flag. The words *Concorde crash* appear in the picture.
- ALT: 3rd June 1973: Russia's supersonic rival to Concorde, the TU-144, crashed in France. The French eventually acknowledged that they had sent up a Mirage III jet to photograph the TU-144 in flight, without telling the Russians. This is the French jet shown in the picture.

Picture 39: "Italian mud slide"

DoP: 03/08/97 | BP:16:12, 04/08/97 | C: *The Times*: 07/05/98

- DM: 6th May 1998: Italian mudslides in May 1997. At least 33 people were confirmed dead, whilst 70 others were listed as missing or feared dead. The car in the foreground is similar to an upturned police car. The cylindrical object on the roof of the car could relate to the police beacon. It is claimed that there is correspondence between the picture and the attached news photograph.
- ALT: 11th September 2000: Italian mudslide in September 2000. The picture depicts the sea of mud with overturned cars and other debris. Twelve people died, and at least five people were reported missing or feared dead.

Picture 40: "Space shuttle fuel problems"

DoP: 27/07/98 | BP:14:01, 27/07/98 | C: *The Times*: 24/07/99

- DM: 22nd July 1999: The space shuttle mission, commanded by a woman (the blonde haired person to the left), was dogged with fuel problems before and after take-off. Leaking fuel is shown in brown around the sides of the spacecraft. Take-off was delayed and the mission had to be shortened.
- ALT: 28th January 1986: The picture depicts the moments before the space shuttle Challenger blows up. Depicted are five of the seven astronauts. Their space helmets convey an iconic/religious feel. The day is bright.

Picture 41: "Dunblane school shooting"

DoP: 11/04/94 | BP: 13/04/94 | C: *The Daily Mail*: 14/03/96

- DM: 13th March 1996: 17 people (16 pupils and one teacher) were killed at Dunblane Primary School. The figure in the top right-hand corner of the painting represents the murderer, Thomas Hamilton. In the drawing the bag contains 4 tools; the killer used 4 guns. The composition is similar to a class photograph that appeared in many national newspapers at the time.
- ALT: 21st October 1966: At 9.15am, a waste tip slid down a mountainside into the mining village of Aberfan in South Wales. Pantglas Junior School, just below, was engulfed as were about 20 houses in the village. 144 people died in the disaster. 116 of them were school children. About half of the children at Pantglas Junior School, and five of their teachers, were killed. The composition is similar to a class photograph. The man in the top right-hand corner is a miner. Notice the words "black background", signifying the coal waste.

B.3 Images

The following are the seven pictures produced by DM that were rated significantly better than the alternative events as outlined in Section 5.1.3.

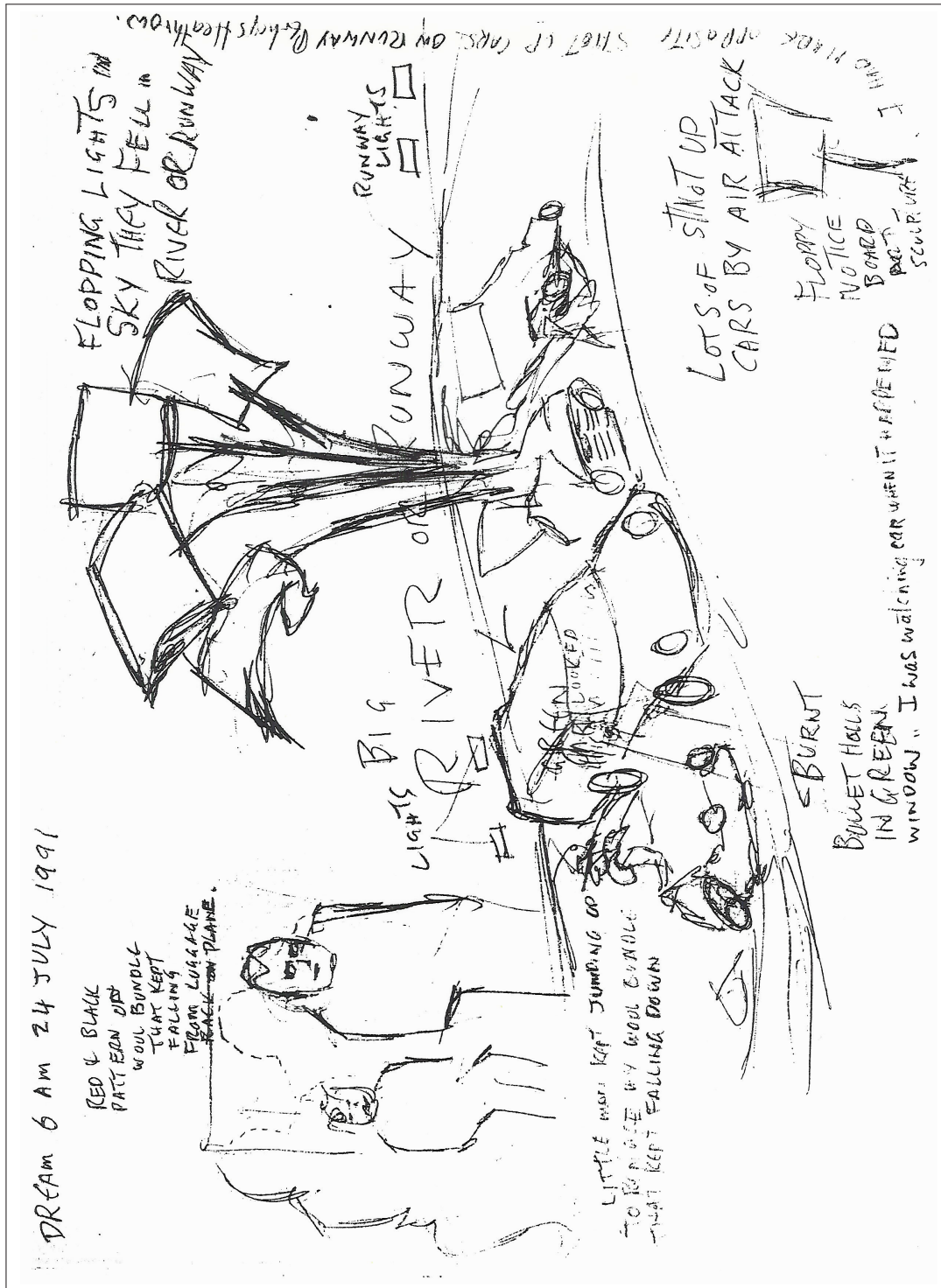


Figure B.1: Picture 7: "Heathrow mortar attack"



Figure B.2: Picture 9: "Dead fish in River Irfon, Wales"



Figure B.3: Picture 15: "Indian earthquake"



Figure B.4: Picture 28: "Rackhams knife attack"



Figure B.5: Picture 33: "Twin Towers sketch"

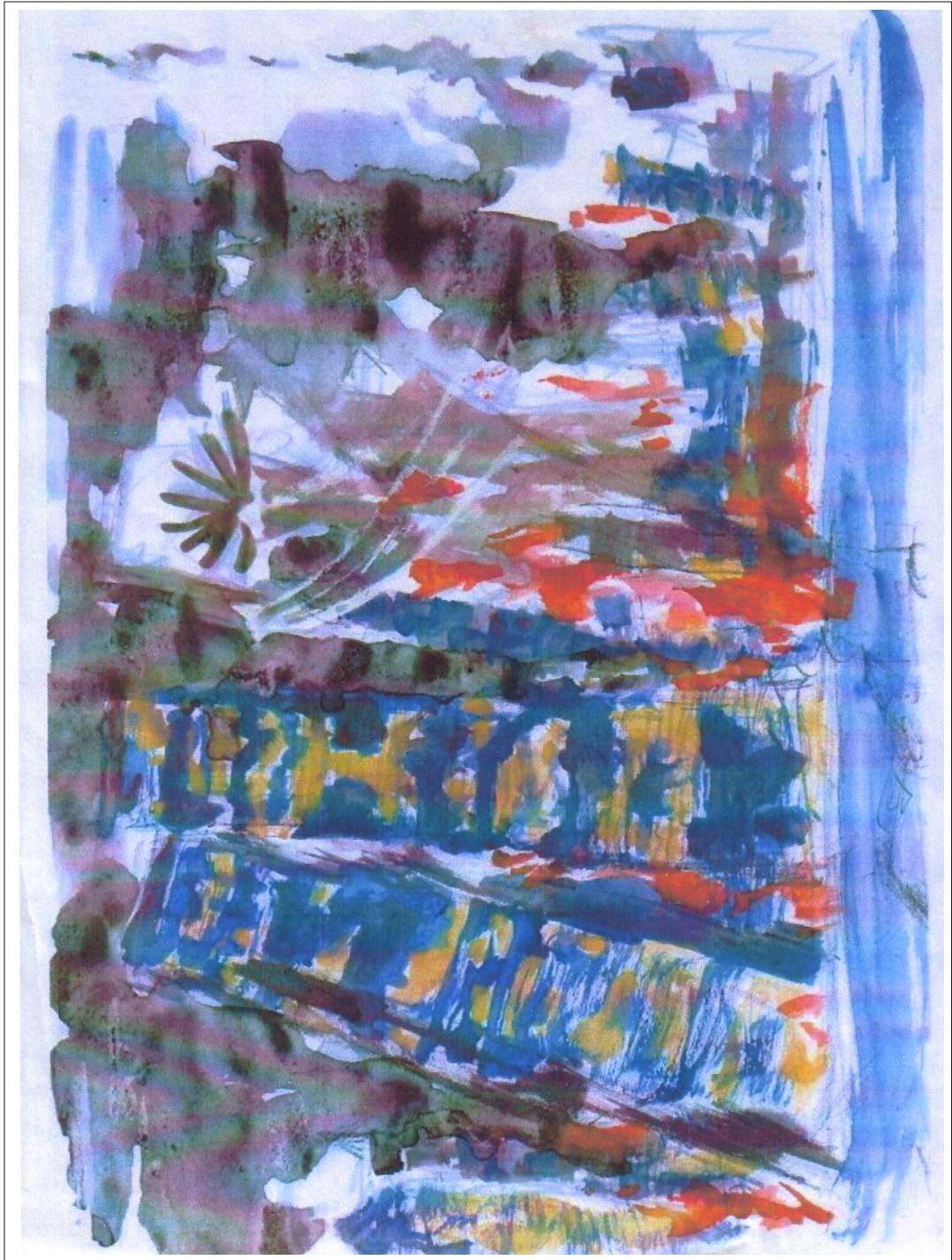


Figure B.6: Picture 35: "Twin Towers colour painting"



Figure B.7: Picture 37: "Two miners killed"

B.4 Supplementary Evidence

This section reproduces two examples of the bank photographs, taken to 'prove' that the dream occurred before the event it has precognised, and two handwritten notes by DM explaining some details, after the event has occurred. These details are provided to show the apparent sincerity of DM's belief in the veridical nature of his experiences, and to give some insight into his character. Digital versions of these and other similar images can be found in the DM digital archive at <https://sites.google.com/site/dmdigitalarchive/>



(a) DM bank photo for picture 18



(b) DM bank photo for picture 22

Figure B.8: Two example bank photographs. Photograph (a) pertains to picture 18 (see page 220) and photograph (b) pertains to picture 22 (see page 221).

DREAM OF RUSSIAN EARTHQUAKE AT NEFTEGORSK

I HAD DREAM ON 10th APRIL 92
EVENT OCCURRED ON 28 MAY 95.

This premonition dream of the earthquake in Neftegorsk on the remote Pacific Island of Sakhalin was very short in duration and frighteningly accurate in its details of the shapes of buildings and the rubble and debris left afterwards. This accuracy is clearly shown in the amazing similarity of shapes and masses to the newspaper photographs in the Evening Standard Telegraph not only showing the 3 main lines of rubble from the collapsed buildings in the centre of the picture but the diagonal angle they form across the picture from the top left to bottom right corners of the scene as it was photographed. Also more amazingly the shape of the building still partly standing in the background behind these 3 lines of rubble shown by 3 cylindrical shapes in my sketch of this dream done immediately on awaking from the dream.

I had this drawing photographed with me holding it at Barclay Bank, Sudbury Hill on the 5th MAY 1992 but owing to the drawing being in pencil it didn't come out - only a faint trace of the shapes coming out on a darkened print although the lettering of Dream 10 APRIL 92 has shown on the top of the photograph. I also had my original pencil drawing signed by 2 independent witnesses on the back of the drawing on 10th MAY 1992.

The dream itself was very short and showed the tremor and buildings collapsing and the scene disappeared in dust which when it cleared revealed the 3 lines of rubble shown in Telegraph photo and my dream drawing. During the dream I was made aware or given a knowledge that it was in Russia that this was happening.

Figure B.9: Note written after the Neftegorsk earthquake, pertaining to picture 12 (see page 219).

NOTES MADE AFTER THE EVENT

TOKIO TERRORIST GAS ATTACK ON UNDERGROUND DREAM

This dream came in 2 different parts or scenes. The first showed subversive figures hiding in a large cave, and the second part showed the interior view of an underground tube train.

The impression I received very strongly in the first part of this dream was of "secrecy and being hidden." I saw seven or eight subversive and silent figures moving stealthily into and from a secret hiding place within a rock cave with a tiny low entrance underground or below ground level so that this entrance would not be easily detected. This meant that the members of this secret society had to squeeze through a tiny underground crack in the rock to get in and out of the cave.

These secret figures were engaged in some kind of what appeared to be terrorist activities 'but underground' as everything in this premonition dream seemed to be happening underground.

The dream then switched to the inside of an underground tube train carriage that had suddenly been evacuated in panic by people who seemed to be Japanese. I as an observer became aware of a very strong nasty smell that affected my nose, throat and lungs and I felt myself gasping for breath in the dream which was so bad that it woke me up from the dream, rather like people suddenly wake up frightened and distressed from a nightmare.

As a result of waking up so suddenly while still in the dream I managed to remember the details of the dream quite well.

Although I don't know Japan at all having never been there I assumed that Tokio would be the only city with an underground tube or Metro system so I guessed that this event when it happened would be in Tokio.

I had a pencil drawing of this dream photographed with me holding it by the date display in Philip & Co estate agents and Building Society in Sudbury Hill on 13 April 1994

DATE OF DREAM 17 NOVEMBER 93
EVENT OCCURRED 20 MARCH 95

Figure B.10: Note written after the Tokyo terrorist attacks, pertaining to picture 21 (see page 221).