CHAPTER 1
INTRODUCTION TO PARAPSYCHOLOGY

Instead of knowledge of what we are, we have belief.
J.B. Rhine (1948/1954, p. 7)

Whence has it [the mind] all the materials of reason and knowledge? To this I answer, in one word, from EXPERIENCE.
John Locke (1679, cited in Drever, 1968, p. 12)

1.1 Overview of the Thesis

In this first chapter, a socio-empirical[1] approach is taken to explore the current state of parapsychology[2] and its subject matter: alleged paranormal[3] phenomena, sometimes referred to as psi.[4] Skepticism from the point of view of polemicists who have "already decided the issue" (Hansen, 1991, p. 202) about these phenomena is the main focus of this chapter, and three forms of inquiry will be made that ultimately challenge this extremely skeptical position on psi. Of course, only a limited number of issues can be covered in a single chapter. First, the singular yet paradoxical nature of psi is scrutinized in terms of its specifiability as an 'acceptable' scientific phenomenon by comparison with other allegedly specified scientific phenomena. Second, the impact on parapsychology of the philosophical and sociological factors of 'paradigm incommensurability' and 'socio-cognitive discontinuity' is examined. Third, three socio-empirical effects on

[1] The term 'socio-empirical' is used here to describe a fusion of, or inter-relatedness between, sociological and scientific ideas and principles both of which underlie the so-called 'objective' findings and statements of scientists.

[2] Parapsychology is the "scientific study of [alleged] paranormal phenomena" (Thalbourne, 1982, p. 51). See n3 for a definition of paranormal.

[3] Both paranormal and psi phenomena "[exceed] the limits of what is deemed physically possible on current scientific assumptions" (Thalbourne, 1982, p. 50). The general distinction in parapsychological circles is that psi "simply denotes a communications anomaly" (Hyman & Honorton, 1986, p. 353), whereas paranormal phenomena "must be adequately explained scientifically . . . by a theory or model the core assumption(s) of which violate one or more of C.D. Broad's so-called 'Basic Limiting Principles' " (Palmer, 1988, p 154; see also, Palmer, 1987).

[4] Thouless and Wiesner, 1947, p. 179, proposed a unitary process underlying ESP and PK. They called it 'psi' from the Greek letter ψ.

parapsychology—experience, belief, and the scientific method—are considered. These three factors are dynamic forces that drive the field of parapsychology onwards. On the basis of the above considerations, it is argued that scepticism, particularly of a form that denies rather than doubts the possibility of psi, has not produced cogent arguments that undermine the psi hypothesis.

Having thus set a context for this thesis, it will be seen that subsequent chapters, although concerned with more experimental and theoretical issues, will nevertheless give us cause to refer again and again to the issues raised in this first chapter.

Chapters 2 and 3 are reviews of the meta-analytic literature in parapsychology, which represent a period of more than 70 years of theoretical and experimental research in the field. The main purpose of these two chapters is to establish justifiable reasons to pursue serious investigations of psi, but the meta-analytic findings presented in these two chapters will also be drawn upon throughout the thesis for theoretical and experimental purposes (see especially Chapter 4 and Chapter 10).

In Chapter 4, a new theory—Thalbourne's (2000a) theory of psychopraxia—will be introduced as a paradigmatic 'shift' in parapsychological thinking that has emerged as a result of the socio-empirical findings of more than a century of research into the paranormal. It will be seen that the meta-analytic findings give some support to the claims made for the theory of psychopraxia.

Seven subsequent chapters (Chapters 5-11) are reports on experimental investigations into the theory of psychopraxia, and other research which has bearing on the theory. Finally, Chapter 12 will be a discussion of the findings of these experiments and research, with some general conclusions about the theory of psychopraxia.

1.2 Prologue

In any scientific field of research that involves experimental work and theoretical speculation there is a widely held assumption that such activities are conducted in a dispassionate way by investigators and theorists, as if the socio-cultural and behavioural influences on that field had been isolated and removed as confounding variables. Far from being removed, they are actively present. But these influences are more than just variables since they not only underpin the whole enterprise of science, but also constitute its very core. A number of critics of science have already made these observations. For example, Latour and Woolgar (1979) state the case:

> Whereas we now have fairly detailed knowledge of the myths and circumcision rituals of exotic tribes, we remain relatively ignorant of the details of equivalent activity among tribes of scientists whose work is commonly heralded as having startling or at least extremely significant effects on our civilization. (p. 11)

Collins and Pinch (1982) also recognize this state of affairs as problematic for science and society, and the following statement implies a solution:

> On the whole, science is done in 'controlled conditions'. It is done in laboratories, conferences, journals, books, universities, and, of course, it is done here and now. Thus while it is certainly not true that every influence upon scientific development is transparent, most of what happens in science today is associated with a visible and investigable 'outcrop'. (p. 3)

Collins and Pinch recommend an ongoing sociological investigation into the scientific enterprise. This approach may lead to an understanding of how science works, but without realization, or even a tacit acknowledgement from scientists that transparent influences (i.e., the social forces that act upon science) exist, they remain invisible and therefore uncontrollable. More to the point, even if these social forces are identified, scientists may still fail to see the relativity and the subjectivity of their claims and findings if they do not adopt a reflexive (i.e., introspective) attitude to those forces. Consequently, they will pursue the same time-honoured course they have always followed. This state of affairs is analogous to a blinkered horse that can have no concept at all of the broader nature of its environment—not unless it bothers to turn its head and look.

Thus we may lack a keen insight into, or at least remain ignorant of, how these influences have had considerable effects on (a) the way experimental and theoretical problems are resolved, (b) the criteria for interpreting results, (c) the future path and attitude of the investigator, and (d) the direction of a discipline. Metaphorically speaking, these influences work in the same way as the invisible 'forces' that shape an environment.

To extend the environmental metaphor, if we want fully to comprehend and appreciate a new territory, then not only must we circumnavigate its borders and coastline, scale its peaks, and enter into its heartland, we must also consider the forces and actions that have shaped and are still shaping that territory. In like manner, parapsychologists must not only seek to discover empirically verifiable phenomena through direct observation, whereupon they posit theories, but they must also consider the field's current status (specifically how it came to be that way) in terms of those crucial but sometimes invisible influences that help give shape and form to those theories. Thus, a careful consideration of

the socio-empirical forces that have shaped parapsychology as a discipline, and continue to shape that discipline, is a necessary undertaking in terms of the importance to the field of maintaining its consistent identity and purpose in a constantly changing social and intellectual environment.

1.3 A Socio-Empirical Perspective on the Paranormal

1.3.1 The Problems of Specification and Replication in Parapsychology

The 'landscape' metaphor used above is reasonably well applied in the case of parapsychology: Just as any territory is dominated and shaped by, or held within the framework of, specific environmental forces, the constantly changing 'landscape' of parapsychology does find itself 'conditioned' by a long-running socio-cultural and historical process of paradigm construction and re-construction. Often the process involves the so-called paradigm 'shift', where new theories and ways of thinking are necessarily put forward that challenge the received view of psi (for examples, see May, Utts, & Spottiswoode, 1995a; Rhine, 1937/1950, pp. 99-101; Stanford, 1977a, 1978; Storm & Thalbourne, 2000a; and Thalbourne, 1981, 1982, 2000a. See also the reviews of theories compiled by Stokes, 1987, pp. 77-189, and more recently, Irwin, 1999, pp. 159-174.)

Though manifold and disparate in principle, these constantly changing constructions and re-constructions, sometimes incorporating philosophical ideas that may or may not be empirically testable, have nevertheless, helped 'integrate' parapsychology into a unified field of inquiry, thus rendering it a 'landscape' or unique *topos* (Greek: 'place') in its own right. In fact, ironically, parapsychology has maintained its 'integrity' strictly on account of the particular nature of its subject matter: alleged paranormal phenomena (i.e., extrasensory perception [ESP], psychokinesis [PK], and life after death).[5]

Skeptics have disputed the above claim (see Alcock, 1985; Blackmore, 1985a; Bunge, 1991; Hansel, 1966; Hyman, 1989; Kurtz, 1985), but parapsychology, having identified its subject matter more than 100 years ago—"all types of phenomena called parapsychological or paranormal"[6]—has achieved this unity by consistently refusing to be

[5] ESP is "paranormal cognition" and PK is "paranormal action" (Thalbourne, 1982, pp. 27, 61). Life after death refers to the survival of human consciousness (or conservation of identity) in one form or another after the demise of the corporeal body.

[6] From the frontispiece of the *Journal of the American Society for Psychical Research*.

intimidated by the specious claim that its subject is unspecifiable (for example, see Boring, 1955, 1966).

To clarify this point about the specification of the "parapsychological or paranormal" from a socio-empirical perspective, it is important to recognize that the related term 'psi' is often used interchangeably as both a 'construct' and a 'paranormal phenomenon' (see Palmer, 1988, pp. 155-156). Palmer warns that the "traditional" approach of positing paranormal theories to explain psi is sometimes confused with the goal of confirming the psi construct. For example, the oft-used existentialist question "Does psi exist?" suggests a need for "construct-validation" of psi in its descriptive capacity as an 'anomaly' (Palmer, 1987, p. 156), yet psi was validated in this way as early as 1886—i.e., by the time of the founding of the American Society for Psychical Research (Palmer, 1987, p. 111). Palmer argues that the investigator might be more gainfully employed in seeking an "explanation of (ostensible) psi phenomena" as possibly being a paranormal process—clearly a different goal to that of the verification of the psi construct. Palmer's focus, then, is on reaching a greater clarity of understanding about psi than is currently circulating inside and outside the parapsychological community.

Before discussing paranormal phenomena any further, we must ask how it is that any phenomenon comes to be specified. Traditionally, specification has been regarded as a process of recognizing how different one phenomenon is from another phenomenon by means of a categorization procedure—thus a phenomenon can be regarded as specified if it has a unique identity. But, here we fall victim to taken-for-granted discontinuities because some form of general agreement on what constitutes a system of difference must first be established (such systems are clearly vulnerable to arbitrary decision rules, exclusion criteria, and other "incommensurables"; Collins & Pinch, 1982, pp. 9-10, introduce these ideas).

Specification has also been understood as a process that reveals the *nature* of a phenomenon by establishing a certain level of predictability of that phenomenon. Edge (1985a) points out the flaw in this kind of thinking. It does not follow that a suitable theory set up to specify a phenomenon should emerge exclusively from the fact that the phenomenon is predictable because prediction does not yield understanding, whereas a theory makes the phenomenon intelligible. (It will be seen shortly that even a good theory does not necessarily guarantee specification in any absolute way.)

Given these definitions, and the fact that attempts to specify an allegedly paranormal phenomenon by conventional means may be faulted, we need to see how a so-called *normal* phenomenon would bear up against such criteria. We can show this by

borrowing a construct from the field of physics (i.e., 'gravity') and probing into the socio-empirical derivation of that construct. The effect we know as gravity is a well-accepted (because repeatable, and therefore a predictable) phenomenon, so it came to be 'talked about' as if it had been specified at some point in scientific history. This misappropriation (*pace* Newton) emerged, first, because gravity, whatever it is, was deemed a genuine phenomenon by *consensus omnium*[7], and second, because Newton had the scientific insight to propose a good theory to explain the phenomenon. But only the effects, not the *nature* of gravity, were recognized and 'explained' by Newton's theory.[8] Still, in the twentieth century, the search for gravity waves as opposed to the equally plausible hypothesized entity, the 'graviton'—both given as possible 'causes' of gravity—only proves that, by the very standards of empirical science, unspecifiable phenomena lurk in all corners, not just in parapsychology.

For gravity, repeatability of the phenomenon was a deciding factor in its acceptance as a real phenomenon (notwithstanding the relatively powerful predictive model proposed by Newton). Yet the community of scientists then, as now, failed to realize that there is no such animal as a "common criterion" or "preconceived notion" (to use Schlitz's, 1985, p. 79, words) in science that was not, or should not be construed as anything more than a construction in the sense that it was subjectively derived by way of consensus. On this basis, scientists accepted the 'gravity' construct on the grounds of very real (i.e., repeatable and demonstrable), mathematically describable effects, and these 'facts' were set up as the criteria underlying the general acceptance of gravity as an 'entity' about which one could propose a theory. Only amongst a very limited circle of scientists in the twentieth century were these criteria and subsequent assumptions questioned.[9]

---

[7] Radin (1997, p. 44) refers to this social phenomenon as the "herd effect," which describes the way scientists herd together and come to a mutual agreement about which ideas or techniques will or will not be taken up as acceptable. This ancient principle effectively describes the formative process of the Kuhnian paradigm (see *1.2.2*).

[8] Newton, in fact, had insight into the 'problem' of gravity: "That Gravity should be innate, inherent and essential so that one body may act upon another at a distance through a vacuum without the mediation of anything else, by and through which their action and force may be conveyed, is to me so great an absurdity that I believe no Man [*sic*] who has in philosophical matters a competent a faculty of thinking can ever fall into it. Gravity must be caused by an agent acting constantly according to certain laws: but whether this agent be material or immaterial, I have left to the consideration of my readers" (Motte & Cajori, 1962, p. 364).

[9] Note how Newton's model is complemented by Einstein's theory that gravity should be perceived as geodesics in (possibly curved) spacetime. To complicate matters, masses do not 'attract' each other, they simply follow a geodesic (a path of least resistance) in warped spacetime brought about by the effects of the

To specify the phenomenon of psi in a scientific way also requires a "common criterion" by which to assess that phenomenon—a "preconceived notion." And it seems that in most cases, the 'common criterion' for assessing the existence of a phenomenon—the 'preconceived notion' that assists in pronouncing with great certainty the overwhelming evidence of a phenomenon—is the prerequisite of repeatability, because predictability ensues, and as a consequence, theory can be gainfully tested. But, how concrete is this repeatability construct? Schlitz (1985) protests that "there is no common criterion for assessing replication in science" (p. 79). In fact, to go further, it is not always possible to specify what the term 'replication' itself is supposed to mean because, as implied above, replication would be dependent on acceptability and the 'preconceived notions' that we use to interpret experimental results (Schlitz, 1985, p. 79; see also Storm & Thalbourne, 2000b, p. 348, who discuss the distinction between statistical and complete replication).

Collins (1976) confirms the above point—when replication is claimed as occurring, it tends mostly to be the end product of negotiation amongst investigators and theorists, and has very little to do with objective facts. For example, Edge (1985a) noted that replication in science does not in fact take place as often as we are led to believe. A 'replication' experiment is often the same experiment with improvements, and/or refinements, and these differences may be of such importance that they disqualify one experiment as being a replication of the other. To use Diaconis's (1991) words: "One cannot judge what 'really goes on' in most areas [of research], and it is impossible to demand wide replicability in others" (p. 386). For example, one "area," the discipline of physics (again!), delivers no more a consistency of results from experimentation than do the social sciences. And Radin (1997) gives a specific example from that discipline by asking: "How many times does an effect need to be repeatedly shown before it is accepted as a 'real' phenomenon and not an artifact?" He was referring to the 'omega-minus' particle, which has been found only twice in 200,000 observations. This particle is now considered 'real' by a majority of physicists by dint of a "sufficient" albeit "poor replication rate" of production/observation of the said particle (Radin, 1997, p. 49).

Clearly, specification and replication of a phenomenon are 'social actions' undertaken in accordance with subjective decision rules and criteria, thus throwing into question the validity of any so-called 'objective' investigation of a phenomenon intended

---

masses on the spacetime continuum. Thus, the specification of gravity (and by induction, any scientific entity) can be seen as discourse dependent.

to establish the express or definitive nature of that phenomenon. Outside the demands of conventional science and its often loosely defined criteria (like the expectation of replication), psi is no different from gravity, yet it is 'acted upon' as if it were—a purely socio-empirical response to an allegedly unspecifiable phenomenon claimed to be such by standards that are unspecifiable. Thus, the tacit assumption in science that repeatability (and its concomitants, predictability and unique identity) may lead to specifiability does not follow. In fact, the *nature* of any entity or construct is arguably unspecifiable.[10]

All the above statements may give rise to the assumption that the belief that specification has been achieved, because replication and its concomitants are somehow entity-confirming processes, is actually a defense mechanism of conventional science set up as a legitimate scientific principle for arriving at some kind of consensus (cf. McClenon, 1984, pp. 89-91). In fact, if parapsychology can be criticized as having not demonstrated a repeatable phenomenon, then perhaps *all the disciplines* may often warrant the same criticism. As an aside, and as an *apparent* contradiction to the socio-empirical approach outlined above, parapsychologists do not have to defend themselves so forcefully against the accusation that psi is not repeatable—in fact, new evidence suggests psi is repeatable. Admittedly, parapsychology has had to wait a long time for its day, but it did, in the closing decades of the twentieth century, yield statistical evidence of repeatable paranormal phenomena (see Chapters 2 and 3). Thus the argument that parapsychology has failed to show replicability can be countered by the cumulative evidence of the meta-analyses. Such findings and statements challenge, in quite compelling ways, the relevance and even the validity of the claim that unrepeatability is parapsychology's only finding (Blackmore, 1985a, p. 183).[11]

The earlier relativisation of the demand for replication, in the sense that repeatability is not necessarily the *sine qua non* of scientific principles, does nothing to debase the validity of the repeatability principle as demonstrated in the meta-analyses. It is accepted that repeatability is a generally accepted hard-line principle for confirming the

[10] Laudan (1983, cited in Rockwell & Rockwell, 1986) has demonstrated that the Aristotelean "know-why" of a phenomenon (the possibility of its being specifiable) was dropped by Galileo as a scientific criterion because he admitted no understanding of the "causes or essence of gravity" (Rockwell & Rockwell, 1986, p. 107). However, he maintained the Aristotelean "know-how" because he could demonstrate how his results were 'true'. "Scientists no longer [claim] that science offer[s] apodictic [church established] certainty" (p. 107).

[11] Other contributions to the repeatability issue have been made by Honorton (1985b) and Rao (1985). See also Shapin and Coly (1985) and Murphy (1987) for further arguments which question the assumption that repeatability is even essential to the purposes of parapsychology.

existence of phenomena. But the more important issue under consideration here is the way in which the demonstrated repeatability of a phenomenon has falsely led many researchers to (a) the illusory belief that said phenomenon is not likely to be an artifact of experimentation or demonstration simply because it is replicable, and (b) the nature of such a phenomenon has thus been specified, or indeed will be discovered, sooner or later.

Realistically, one needs to recognise the implications for science of this dichotomy where replication is accepted while at the same time can be regarded as unnecessary, and one cannot make a sound judgment about these implications by the standards of the empirical scientist or the critical sociologist alone. On the one hand, repeatability is a means of identifying (i.e., confirming the *existence* of) a phenomenon by the frequency of its occurrence, but many, if not most, inferences about the *nature* (specification) of that phenomenon might best be classified as speculative, or even philosophical. On the other hand, repeatability is not always a possibility, but that circumstance should not give rise to any suspicion about the existence of the phenomenon. Most researchers would surely acknowledge these issues, but the literature does not always suggest it. For example, see Honorton (1993) for a similar argument. Honorton identifies the 'speculative' nature of the terms ESP and PK, preferring the terms "anomalous communication" and "energetic processes," respectively. He even recommends that the 'anachronistic' term "paranormal" be "abandoned" (pp. 210-211).

In summary, when parapsychologists are expected to specify and replicate psi, yet fail to satisfy the demands of skeptics, they are criticized on the basis of technicalities that just happen to lend themselves very well to an ideology of 'skeptical scientism'[12] (to coin a term), which blazons itself as good science. Thus do the same skeptics (often the spokespersons of other disciplines) make unreasonable challenges against parapsychologists by issuing socio-empirically derived demands they would not require themselves to fulfil. (See also Irwin's, 1999, p. 314, observation that the demand for repeatability is largely a rhetorical exercise underscored by ulterior motives.)

To formulate a question based on the general statements posed in this section (viz., "Has the singular nature of psi been undermined by scientific criticism?"), it has been argued here that specification of a phenomenon is an idealistic notion falsely associated with repeatability and its concomitants. Some kind of scientific community must accept the phenomenon, identify it, name it, and make inferences about its nature. Like gravity, psi phenomena *have* been identified, and constructions and re-constructions have been

[12] I shall tentatively define 'skeptical scientism' as an 'anti-anomaly', pro-scientific position dependent upon doctrinaire scientific convention.

formulated around them within the paradigmatic frameworks of scientific discourse. Again, like gravity, psi phenomena maintain anomalous dimensions. This pluralistic aspect may forever be unrepresentable in science. To borrow from Rhine (1948/1954, p. 7) it often seems more accurate to *specify* what we collectively *believe* about a phenomenon rather than what we *know* of it. (In a later section, the influence of belief will be scrutinized further—see *1.3.4.2*.)

1.3.2 The Nature of the Paradigm in Parapsychology

From the above arguments, it can be seen that certain criticisms leveled against the subject matter of parapsychology in terms of its existence cannot be founded on objective grounds so that the field gains 'the benefit of the doubt', and therefore, psi may be workably taken as an identified phenomenon. However, the existence of possible disparities between various paranormal paradigms (or at least disparities that often exist between various theories[13]) is recognized. In fact, the Kuhnian (Kuhn, 1962) concept of 'incommensurability' between paradigms (and theories in many cases) could be seen as another threat to the integrity of the paranormal and therefore the integrity of parapsychology. A split in the parapsychological community, however, has not yet surfaced, in spite of the very real opportunity for such an occurrence (cf. Honorton, 1993, p. 209, who considers the variety of hidden agendas held by parapsychologists). Of course, other disciplines face these same difficulties, but they have the buffer of scientific hegemony supporting them when such threats arise (cf. Irwin, 1999, p. 314.)

In terms of Kuhn's (1962) notion of 'paradigm incommensurability', there may be no solution to the above-mentioned difficulties that might surface because there are supposedly no common measures by which ideas from one paradigm could be understood from within the context of another. Following Kuhn's way of thinking, a paradigm shift within parapsychology could imply a 'revolution' in the way its subject matter is to be defined.

But there seems to be a flaw in Kuhn's thinking. Kuhn led us to understand that science is a progressive enterprise, the progress being gauged by the higher predictive power of the new, superior paradigm compared with the old, less refined one. More broadly, he constructed science as a 'collective' of disciplines that are supposedly forging ahead in some way, ultimately creating newer and better technologies and more advanced societies in which to live. Paradoxically, the Kuhnian model seems to be an endorsement

[13] See Stokes, 1987, pp. 77-189.

of the aesthetic of the objectively verifiable 'trajectory' of modernism, yet it maintains the conflicting two-edged 'postmodern' premise that (a) paradigm comparison is not a plausible undertaking, which leads to the inevitable conclusion that (b) progress cannot be determined in ways other than subjective, and may not even take place at all by the standards of a relatively broader, more objective (i.e., less subjective) criterion.

The argument seems clear: If paradigms are incommensurable with each other then there is no standard by which they can be compared. Thus, to state that things can somehow be 'better', or that progress (in a discipline, for example) can be objectively substantiated, undermines the relativist argument that Kuhn puts forward. However, this ostensibly logical conclusion can be given a radical twist, as Collins and Pinch (1982) have shown. Their argument is that whatever other arbitrary system of decision-making scientists might resort to in assessing new ideas, "scientists reserve their final commitment to [these] ideas until they are forced to accept them by logic, or experiment" (Collins & Pinch, 1982, p. 16). To go further, rather than being at a loss in deciding which knowledge to keep and which to reject, scientific epistemology is evidently and productively a composite of the 'old' and the 'new'. Its diverse contents can be drawn upon like tools in a toolbox, with each tool doing a specific job. Thus, we can live with the dichotomous situations that science sometimes sets up without necessarily being frozen in our tracks.

For parapsychology, then, it is almost business as usual. There seems to be no reasonable argument that would undermine the recommendation that things proceed as normal, providing that one major concession be made. Parapsychologists may accept or reject both old and new ideas according to their findings, but with reservations. It is now not just a matter of assessing which theories do a better or worse job of explaining experimental outcomes. Judgments may need to be made about outcomes in terms of multiple paradigmatic viewpoints, since different paradigms recognize different phenomenologies according to the discursive constructs of those paradigms

The investigator will still be free to design research, and assess the results of that research, from within *or without* a specifically preferred, and/or generally approved, paradigm or theory. In the past, this state of affairs suggested a discipline that was fragmented, but it can now be regarded as a major philosophical strength suitable for any discipline. In recent decades this pluralist approach has been adopted in parapsychology (see also Irwin, 1999, pp. 318-319). Thus this approach accommodates, and is sympathetic to, the subject of the current thesis—the theory of psychopraxia—and the ways in which other theories might bear similarities and differences to the theory of psychopraxia. These

issues resurface throughout the experimental chapters, but will be brought into much finer focus in Chapter 4.

1.3.3 Socio-Cognitive Discontinuity in Parapsychology

Apart from the term 'incommensurability', Kuhn (1962) acquainted us with two more terms that also describe the schismatic relationships that can supposedly exist between paradigms: 'Incompatibility' and 'irreconcilability'. These three related terms effectively underpin the all-embracing concept: "socio-cognitive discontinuity" (Collins & Pinch, 1982, pp. 9-24, 47-65).

Socio-cognitive discontinuity describes the state of affairs that may exist between (say) two groups of scientists who use mutually exclusive epistemologies to inform their 'social actions' (Collins & Pinch, 1982, p. 9). Even when there is a consensus between groups on the facts about a phenomenon, they may differ about the ideas and conclusions that follow from those facts. Thus, in cases where groups (or, in fact, individuals) are opposed, argumentation may proceed with little chance of a resolution between the two (or more) factions.

Socio-cognitive discontinuity can be demonstrated in the field of parapsychology. When, for example, Pratt (1960, p. 25) regards psi phenomena as "precisely those psychological events which defy description in terms of any physical theory now available," we can see that he seems to 'tell' what psi is by telling what it is not—that is, the fundamental concepts of parapsychology are 'essentially negative' (to borrow from Flew, 1987, p. 93). But, to argue tentatively in support of Pratt, his characterization of psi is in harmony with the psi hypothesis: Psi may be incompatible with current scientific principles, but that should not preclude its existence when it may be the case that the problem lies with our scientific principles, not the psi hypothesis. After all, it is entirely consistent with our philosophical and scientific epistemologies to posit 'x' and therefore 'not-x', where 'x' is normal (i.e., 'regular') whereas 'not-x' is psi (i.e., 'irregular' or even 'paranormal'). Psi phenomena would thus be anomalous only by dint of their being incompatible with current definitions of what characterizes non-anomalous ('normal' and even 'abnormal') phenomena.

Of course the philosopher can repudiate this kind of formulation. First, can there be a general agreement on what constitutes 'x' and 'not-x'? Second, if 'x' and 'not-x' are in contradistinction to each other, can there be a mutually agreed upon system for drawing out this distinction? (We have just dealt with these issues—see *1.3.1*) The parapsychologist is usually oblivious to, or at least a little blasé about such questions, leaving these sorts of

problems to the philosopher.[14] And justifiably so since, as science stands, the main point of the above argument is that Pratt's position is tenable.

The skeptic, who may also be oblivious to these questions, may enter the debate at this point and start his/her argument, in agreement with Pratt, by also stating that psi phenomena are incompatible with scientific positivism. Recall Pratt's (1960) idea that the very real psi event "def[ies] description in terms of any physical theory now available" (p. 25). But the skeptic (assumed here to be one who avows a total allegiance to 'scientism' alone) will be led to state that psi cannot exist because science cannot explain it. Collins and Pinch (1982) make reference to this paradox: "These two arguments have at their core the same claims regarding the incompatibility of psi phenomena and science though they are deployed respectively by sceptics and believers . . . " (p. 49).

Believer and skeptic both agree on what psi is according to what science cannot explain. But for socio-cognitive reasons they split at this point—the one convinced that science needs to be more flexible, the other certain that there is no room for psi according to the standards of scientific truth. These arguments create the socio-cognitive discontinuity of which Collins and Pinch speak, but, to go further, both these arguments are actually epistemologically discontinuous, and therefore, unsound—both are derived from situational presumptions. Thus, the believer introduces the condition that the scientific viewpoint must change (not necessarily true, but it does square with the evidence of science as an accumulative enterprise *qua* Kuhn, therefore fully endorsing the research ethic), whereas the skeptic introduces the condition that the scientific viewpoint need not change (not constructive for science because it undermines historical process, therefore implying that all breakthroughs are explainable, and even that proof can apply without experimentation!). (There are psycho-sociological reasons underlying the social category within which one chooses to be; viz., believer or skeptic. This issue is explored in *1.3.4*.)

Socio-cognitive discontinuity, then, arises out of a failure to discern (or acknowledge) the possibility of alternative meanings that are necessarily implied in any number of scientific statements that have been posited as certainties. Such statements may be exposed for their inconsistencies. In support of the believer's argument, the only truly responsible scientific premise that would fairly represent the psi hypothesis (or any hypothesis) is that researchers proceed in their quest to accumulate an ever-increasing

[14] Thalbourne's (2000a) discussion of what may essentially be called 'the ontological question' is taken up in Chapter 4, where we shall see that psi (as both ESP and PK) can be characterized from within three mutually exclusive philosophical systems: Metaphysical Dualism, Metaphysical Materialism, and Metaphysical Idealism. The distinction between 'x' and 'not-x' will depend on one's ontological perspective.

epistemology. In recognising the implication of the skeptic's argument, researchers on both sides of the fence must not delude themselves into thinking that they have not undermined the fundamental Cartesian tenet of creating certainty from doubt. This certainty may be based on knowledge claims already established and accepted when, on the one hand, the reality is that these claims may not have been fully applied. On the other hand, they may have been carelessly applied to (say) the psi hypothesis (for examples of knowledge claims not fully applied, or carelessly applied to psi, see Collins and Pinch,[15] 1982, pp. 47-65; Meehl and Scriven, 1987). To do the latter responsibly, the practitioners of science (opposing groups of scientists) must avoid the all-too-easy task of constructing socio-cognitive and epistemological discontinuities that can often arise out of the presumptive thinking that attempts to eliminate doubt. Such constructs ultimately propagate doubt itself, not certainty. The introduction of insubstantial and/or unsubstantiated conditions and provisos (for example, rhetorical devices, over-opinionatedness, emotional appeals, possible futures, etc.) merely to peddle an argument is not, and never has been, good science, no matter how good it looks 'on the ground' (i.e., at face value.[16]). Relevant to the theme of this chapter is the fact that this type of science, inside and outside parapsychology, seems more like the social actions of the practitioners of scientism, rather than like science.

Thus, we have to be cautious about how we construct so-called 'good' science (including 'good' parapsychology). Socio-cognitive discontinuity, as well as epistemological discontinuity, should but seldom do create cognitive dissonance amongst those prone to its effects. Antithetical statements can be drawn from a single observation, yet often neither the poverty of these statements, nor the ensuing irony of their contradictory natures are acknowledged, if they are ever noticed at all.

The next section will focus on other ironies, with particular focus on the influences of experience, belief, and the scientific method as socio-empirical forces that also give form to the field of parapsychology. These factors necessarily emerge from the issues just discussed. It will be seen how it is often very difficult to discuss experience, belief, and the

[15] It is taken, for example, that psi does not obey the inverse-square law, but "the inverse-square law does not necessarily apply, even to ordinary electro-magnetic radiations" (Margenau, 1966, p. 222, cited in Collins & Pinch, 1982, p. 58). And "the inverse-square propagation of energy is seldom realized in practice" (Rush, 1943, p. 48, cited in Collins & Pinch, 1982, p. 58). Thus the "radiation analogy" can be used to show that "psi radiation" need not conform to the inverse-square law (Collins & Pinch, 1982, p. 58).

[16] The above statements do not invalidate scientific 'conjecture', or 'speculation', etc., which usually invite further research for the purposes of verification. Absolute declarations, however, are criticized here as typifying the poorest quality science.

scientific method independently of each other since there is so much conceptual overlap between them. But it will also be seen how all three factors have shaped the field of parapsychology, and how they continue to contribute to its evolution and our understanding of the anomaly called psi.

1.3.4 Experience, Belief, and the Scientific Method

Socio-cognitive discontinuity leads us back to the phenomenology of the paradigm. We may ask how paradigms are constructed, and what the key influences that maintain or disturb one's choice of social category might be (i.e., whether we choose to call ourselves believers or skeptics of psi). Kuhn (1962) proposed that new paradigms replace old paradigms because the old ones fail to answer new problems. These problems arise through observation and our capacity to reason. Scientific paradigms, therefore, are born of human reason, but reason cannot take place in a vacuum—it is influenced by the factors of experience[17] and the beliefs[18] we hold. These two factors are constantly changing—they influence each other, and they in turn underscore the processes of paradigm construction.

Thus we conduct research in characteristically human ways, which yet involves thinking from within the perspectives of prior epistemologies—we need the old techniques (experimental design, statistical methods, etc.) to test new paradigms. The scientific method, then, as an "arsenal of research tools" and a "philosophical outlook" (to use Rosnow & Rosenthal's, 1993, p. 19, terms), is a consistent and major force in its own right, thus constituting a third factor in the formation of scientific paradigms.

All the sciences must rely on these three major factors—experience, belief, and the scientific method—as the general means by which proof-oriented and process-oriented research may continue. There are undoubtedly other factors involved in scientific practice,[19] but for parapsychology, experience, belief, and the scientific method have a

---

[17] John Locke advocated experience above all things—even rationalism—as the raw material of all knowledge and reason (Drever, 1968).

[18] Quine and Ullian (1970, p. 92) hold that "in a person's web of beliefs there is no strand that does not help support some value judgement." Judgement is a subjective, non-intellectual process, like feeling-toned evaluation. The process, like thinking, can function in abstract mode, and is thus capable of arriving at sophisticated concepts and ideas about phenomena (Jung, 1987, para. 724).

[19] Rosnow and Rosenthal (1993, pp. 17-19) note that the orienting attitude of the scientist, which includes enthusiasm, open-mindedness, inventiveness, confidence, etc., has an influence on scientific practice. In fact, some of these 'variables', known as "psi-conducive states" (Braud, 1975) are claimed to have real psi influences on outcomes in parapsychological experiments. Some of these variables are tested in later chapters.

particularly unique and pivotal influence on the field, as will become clear in the following sections. These three factors also constitute the formative elements of the socio-empirical approach that an investigator may adopt. For these reasons, they are discussed next.

1.3.4.1 Experience: Ordinarily, one would hold that there is no teacher like experience. Psi phenomena usually seem the most real when they are experienced first hand, and they have a numinous quality about them that 'informs' and enhances that experience. Such an experience may derive from a spontaneous case. However, for the experimental parapsychologist, the 'experience' (or better, 'evidence') of psi is often represented only as a cumulative statistic, where individual trials in laboratory situations provide units of data that appear rather ordinary or chance-like on their own. Often, therefore, due to a failure to demonstrate psi directly (such as in the spontaneous case) or inferentially (through a statistic), many parapsychologists have probably entertained Beloff's doubt: the fleeting thought that "perhaps, the whole field had been misconceived from the start" (1990, p. 11).

Beloff's Doubt (to coin a term) ultimately implies two possible assumptions. On the one hand, psi cannot be convincingly and consistently demonstrated (or better, induced at call) in the laboratory because it does not exist. Of necessity, therefore, the parapsychologist's techniques (and even his or her own influence, otherwise referred to as the 'experimenter effect') must be ineffectual.[20] On the other hand, if psi does exist, the way we go about looking for it may be inappropriate given its phenomenology. There are many who agree with the former assumption (Hyman, 1985a; Hansel, 1966, 1980), while others have expressed sympathy or agreement with the latter (Rhine,[21] 1948/1954; von Franz, 1980).

For most researchers the idea of questioning the techniques of laboratory investigations of psi, and the existence of the psi effect itself, would be considered unreasonable, whether or not these researchers have the 'evidence' of first-hand

[20] Akers (1984) has brought attention to 'explanations' that skeptics have used to justify their doubt regarding the existence of psi, such as randomization failure, sensory leakage, cheating, and procedural errors. Added to the skeptic's list of 'explanations' of the psi event are "delusion, . . . coincidence, [and] unconscious inference" (Stokes, 1987, p. 84). The notion that psi might be an artifact of the methodologies used in parapsychology is discussed below (see *1.3.4.3*).

[21] Rhine acknowledged the general problem of the possible inappropriateness of testing in parapsychology. In speaking about the decline effect, for example, he said: "We destroy the phenomena in the very act of trying to demonstrate them. Evidently the tests themselves get in the way of the abilities they are designed to measure" (Rhine, 1948/1954, p. 161).

experience. Such loyalty is praiseworthy, but a working lifetime devoted to a single ideal can create an insurmountable prejudice. Beloff (1990) has stated a profound truth in this regard: "By the laws of cognitive dissonance the longer you commit yourself to some cause and the more effort you devote to it the harder it becomes to renounce it" (p. 11).

True words, though not true for all scientific researchers. Many investigators in the past, armed with certain prejudices, have entered a field in order to disprove that field's major claims, only to be convinced otherwise by their own experience. For example, French psychologist Michel Gauquelin (1983) sought to prove a life-long belief that astrological relationships as proposed by astrologers did not exist between the inner planets of our solar system and the professional lives of certain eminent individuals. In correlational studies spanning more than 40 years of Gauquelin's working life, statistical analyses consistently gave significant results that challenged his prior beliefs (see also West, 1992, p. 233, pp. 284-289, 302-308, who provides evidence of successful replications of Gauquelin's work).

Again, starting off with his own brand of skepticism, Walter Mischel (1968, 1973) was most emphatic in his criticism of the idea of a consistent personality (temporally and situationally) as proposed and explained by psychodynamic theories (Freudian, etc.). Not until he conducted his own longitudinal studies (showing that character traits, which developed in childhood, could persist into adulthood) did he classify the findings of this work as being akin to that of the psychodynamic theorists (Mischel & Shoda, 1995). Thus experience and the scientific method when taken together may temper our prejudices.

Experience, then, is an important component of scientific research, as it can determine the future path of an investigator. The history of parapsychology, too, has been dotted with examples of the effects that experience has had on its researchers. Susan Blackmore (1985b, 1986, 2001a), for example, joined ranks with the skeptics when she repeatedly failed to demonstrate the presence of psi in her experimental studies (probably a case of Beloff's Doubt, perhaps indicating too that cognitive dissonance needs time to take effect). On the other hand, some researchers, like the late Charles Honorton, have run consistently successful experiments (Beloff, 1990, p. 38), encouraging others to join or continue in the field.

The socio-empirical value of experience should not be ignored or conceptualized out of existence, on the grounds that its chance-like (i.e., coincidental) nature renders it unworthy as a contributing element towards scientific knowledge and the scientific enterprise. Utts (1991) has stated that "experience is a poor substitute for the scientific method" (p. 363), and she has also claimed that "there is little more [in the way of

suggestive evidence of psi] to be offered to anyone who does not accept the current collection of data" (Utts, 1995, p. 290). Perhaps she is correct for the most part, but as she herself has discovered, a large cross-section of the scientific community is not prepared to accept that which they would deem to be unexplained anomalies (Diaconis, 1978; Hyman, 1991).

Such denial, however, is typical of the outsider (specifically, the skeptic) bereft of a personal experience in the area of parapsychology (especially a direct experience of ostensible psi). So perhaps there is more to be offered to anyone who doubts the evidence. Of course, one cannot reject the fact that a novel encounter of the psi kind can wear off quickly—"every miracle, no matter how incontrovertible it may once have appeared, will lose its lustre" (Beloff, 1990, p. 33)—especially if the "miracle" is not reinforced by replication. Beloff (1990) gives examples illustrating how difficult it is to quell doubt and "sustain," or even instill a belief in psi, even when the evidence seems beyond doubt (pp. 30-32). Little wonder that the lack of an experience gives good reason for some critics to follow their doubts and criticize mostly the methods used by parapsychologists (especially the statistics, but more on this shortly), there being little recourse to do otherwise.

While the efforts of many critics do help to fine-tune scientific methods in parapsychology, the psi experience still seems to count once again as a key influence towards believing in the paranormal (especially if recent techniques such as meta-analysis are questioned). In other words, skepticism may be instrumental in strengthening the methods of parapsychology, but nothing seems to inform an individual's belief about the paranormal quite like the direct psi experience, even though many may believe without it.

1.3.4.2 Belief: Quine and Ullian (1970) describe the complex processes of belief formation and modification that many of us employ: Self-evidence, observation, testimony (i.e., vicarious observation), and hypothesis testing. Most researchers are familiar with these processes, since they may also be applied in the procedures of the scientific method. (In a formal scientific investigation, however, these processes may not necessarily result in any particular belief about the nature of the phenomenon under investigation.)

The discipline of philosophy has provided additional means of belief construction through deduction, induction, analogy, and intuition. Each of these approaches can be used independently or collectively as a means of arriving at a belief, and they may extend outside the formal disciplines to the layperson, even leading him or her to skepticism of, or belief in, paranormal phenomena.

As already noted, however, no amount of evidence may be enough to change the "closed-minded thought processes" or beliefs of the skeptic, to use Stokes's (1987, p. 84) words. To explain skepticism, Tyrrell (1945) argued decades ago that "the constitution of our minds, as a result of biological evolution, causes us to reject whatever is entirely foreign to the world of common experience." For many, such is the basis for the denial of psi phenomena too.

On the other hand, many millions the world over in modern (westernized) and pre-industrial cultures do believe in anomalous acquisition of knowledge, communication with discarnate entities, and so on. One might almost surmise, therefore, that belief formation for all peoples cannot stem from the application of any one or more of the above-mentioned processes alone. For if the opposite were true, unanimity of belief about most phenomena, including psi, would have been reached by now in all populations. However, such is not the case. Thus, there appears to be more to belief formation than we might realize.

Having reviewed much of the literature on paranormal belief, Irwin (1993, 1999) discussed four hypotheses that may serve as more general explanations for belief in the paranormal—the *social marginality hypothesis*, the *worldview hypothesis*, the *cognitive deficits hypothesis*, and the *psychodynamic functions hypothesis*:

(i) *The social marginality hypothesis:* Believers are seen as 'isolated' in some way—by age, gender, socioeconomic status, ethnicity, or culture, etc.,—so that belief acts as a buffer to the vicissitudes of a life so often associated with marginal status. (After an extensive literature review, Irwin, 1999, p. 285, reports that this hypothesis has not been supported.)

(ii) *The worldview hypothesis:* Believers hold paranormal belief as only one subjective and esoteric aspect of a much broader picture of life, humanity, and the world in general, where mental and metaphysical religious beliefs and practices are incorporated into that worldview. (There is partial support for this hypothesis, according to Irwin, 1999, pp. 285-287.)

(iii) *The cognitive deficits hypothesis:* Believers are said to have uncritical, naïve, and irrational thought processes based on deficits in education, lower average levels of intelligence and poor reasoning skills. (Evidence for this aspect of the hypothesis is inconclusive at this stage, according to Irwin, 1999, pp. 287-289.) Believers may also have overly creative imaginations.

(There is some support for this aspect of the hypothesis, according to Irwin, 1999, p. 289.)

(iv) *The psychodynamic functions hypothesis:* Believers are seen as being in some way psychologically disadvantaged or maladapted, or socially deviant in respect of their personality characteristics so that paranormal belief is seen as fulfilling a psychodynamic need (see also Radin, 1997, pp. 242-247). (There is "general support" for this hypothesis, but Irwin, 1999, p. 291, proposes that a "specific version of the hypothesis that posits more precisely the nature of the psychodynamics involved" is required.)

Overall, there is some evidence supporting the last three hypotheses, with particularly strong support from the correlational studies for the *psychodynamic functions hypothesis*. In respect of this hypothesis, and as an extension to it, Irwin (1993, pp. 21-30; 1999, pp. 291-292) notes that paranormal belief may be the result of physical abuse and childhood trauma, a mechanism whereby a psychodynamic need is fulfilled by an "illusion of control" which the belief serves (1993, p. 28). There is no evidence, however, that such an aetiology excludes paranormal ability. That is, while it might be found that some believers in the paranormal base their belief on illusory states of mind or fantasy proneness, it might also be based on genuine paranormal ability. (In fact, although controversial, it may not be all that unreasonable to propose that trauma may be a sufficient psi-conducive condition!)

In parapsychology, too, it must surely be more than a state of mind that motivates many or most investigators in their research. Irwin (1999, p. 291) actually includes "anomalous experiences" as one of two possible threats to an individual's intellectual and emotional autonomy (the other, as mentioned, being "traumatic events"), so it is feasible that experiences of genuine paranormal phenomena may be amongst those anomalous experiences. Thus may any number of parapsychologists be of a type actively engaged in the scientific pursuit of an understanding of the paranormal based on personal experience of such anomalies. Their intellectual and emotional autonomy may not be 'threatened' at all, but may be modified to include an openness to, and acceptance of, the reality of psi, based on an informed understanding, and/or experience, of the paranormal. In concurrence with this view, Irwin (1999, p. 290) proposes that belief in psi might logically stem from "the data of parapsychological research," but he does not actually say that direct experiences of psi might be sufficient causes in themselves for belief in the very same phenomena.

To continue this argument, one "specific version" of the *psychodynamic functions hypothesis*, by which at least one underlying *psychodynamic* cause of belief in psi might be explained, lies in the mental/neuro-psychological processes described in Thalbourne's (2000d) theory of transliminality. Transliminality is defined as "a hypothesised tendency for psychological material to cross (*trans*) thresholds (*limines*) into or out of consciousness" (Thalbourne & Houran, 2000, p. 853). (Transliminality is similar to Hartmann's Boundary concept and indeed the two correlate highly. Houran, Hartmann, Thalbourne, & Hughes, 2002.) Thalbourne (2000d) has found significant relationships between transliminality and a number of factors including belief in psi (see **5.3** for a review of the transliminality literature), but more to the point, transliminality has been shown to be a predictor of paranormal ability, as found and stated in Storm and Thalbourne (1998-1999):

> The highly transliminal individual was more likely to be warm and out-going (16PF Factor A), idea-oriented and imaginative (16PF Factor M), and to believe in his or her own ability to achieve a successful outcome in a psi-task, a belief which was vindicated in part by the fact that [a significant] 40% of the highly transliminal participants obtained a hit, suggesting that they . . . might be 'psi-able'. (p. 115)

In short, belief in psi need not necessarily be associated with pathology, but may follow from psi ability.

The argument does not end here: It may have been shown that paranormal belief can engender a feeling of control (Irwin, 1999, p. 292), but again, such control may not necessarily be an illusory one. Even if we were to disregard Storm and Thalbourne's (1998-1999, p. 115) results, there is still stronger evidence that belief in ESP is a psi-conducive factor, as has been borne out in the many "sheep-goat"[22] studies since 1947 (Lawrence, 1993). Lawrence found that belief in ESP correlated positively with ESP scores. (See also Haraldsson, 1993, where various forms of belief other than belief in ESP also correlated with ESP scores.) Certainly, the socio-empirical effects of belief can determine the direction of scientific research, but belief also appears to have real motive force as well (psychologically and parapsychologically) so that it is indeed a variable of quite considerable impact worthy of continued investigation as a psi-conducive factor.

Belief, then, appears to have manifest potency, and this fact may give reason for the reverence in which we hold our beliefs. It is not surprising, therefore, that our beliefs are

[22] Sheep are believers in ESP, and they tend to score significantly better than goats, who are disbelievers in ESP.

the very foundations of our truths. But however 'true' our truths may seem, the often personal nature of such truths must not be overlooked. We run the risk of deceiving ourselves if the so-called truths, which arise from our beliefs, can be argued as misguided due simply to their being nothing more than something we firmly believe to be true (i.e., are found not to be substantiated). In such cases Irwin's (1993) *cognitive deficits hypothesis* holds certain validity.

A few years after Tyrrell wrote the words quoted above (see *1.3.4.2*), Jung (1960, para. 821) drew a distinction between the unique, "just so" quality of the chance-like (coincidental), personal truth and the more scientifically acceptable "statistical truth." Jung claimed:

> Absolutely unique and ephemeral events whose existence we have absolutely no means of either denying or proving can never be the object of empirical science; rare events might very well be, provided that there was a sufficient number of reliable observations [cf. the omega-minus particle]. The so-called possibility of such events is of no importance whatever, for the criterion of what is possible in any age is derived from that age's rationalistic assumptions. (1960, para. 821)

(These thoughts touch on an earlier discussion concerning scientific consensus—see *1.3.1*.) One "rationalistic assumption" that can be made in order to verify empirically the possibility of a certain phenomenon is to show that its probability is not zero. One can go further and establish the actual probability of the phenomenon (achieved statistically), which in our "age" involves considering the frequency of the event in terms of its relation to chance. Thus enters the need for statistics, and more generally, the scientific method in parapsychology.

1.3.4.3 The scientific method: In the nineteenth century, science became defined by its methodology—the scientific method—and in many disciplines, statistical analysis, became one of the fundamental components of the scientific method. Having successfully made the transition to the behavioural sciences, the scientific method naturally became a crucial part of experimental parapsychological research, and has served the field remarkably well for over a century.[23] But in the process, the effects of which statistics is not designed to

[23] Richet (1884) and Edgeworth (1885, 1886) were the earliest to propose the use of statistical approaches to the analysis of data generated in psi experiments. Fisher (1924, 1929) followed with improved methods that credited "close" guesses in card-guessing experiments as evidence supporting the psi hypothesis. See also

explain or give account (particularly psi effects) have been too easily labeled as artifacts, based on the assumption that statistical and methodological procedures may be fundamentally flawed.[24] For example, for some time there was the assumption that true randomness cannot be achieved in any finite system (Spencer-Brown, 1957), but Scott (1958) refuted this claim (Spencer-Brown overestimated the importance to the argument of true randomness). Gilmore (1989) later warned that researchers in psi "should take account of the possible inappropriateness of classical inferential statistics" (p. 338).

Once again, the average parapsychologist would have no difficulty in rejecting as untenable the consequent idea that these effects, which they usually take as suggesting the presence of psi phenomena, are to be taken as the perennial artifact of fundamental flaws. Such a rejection seems warranted, given the vast improvements in methodologies over the last 150 years or so (see Pratt, 1973; Braud, 1991). Braud, in particular, objected to the possibility that a "singularly powerful flaw" or "more subtler confounds" could be responsible for the "positive outcomes" in psi studies (1991, p. 58). To argue the case, he cites a "vast number of successful experiments" by many different investigators, all of whom have used "very different methodologies" (for example, ganzfeld, autoganzfeld, forced-choice, etc. These and other domains are introduced in Chapters 2 and 3).

To expand the range of techniques and statistical tools for assessing psi, and to help avoid hackneyed experimental and statistical approaches, Rosenthal (1986) and others (e.g., Schmeidler, 1968; Morris, 1991) have advised the use of additional methodologies, and even an alternative to significance testing of hypotheses. Rosenthal, for example, recommends "sub-division of studies as a function of different experimental procedures or individual difference variables such as sex, age, degree of belief in psi effects, and the like" in order to arrive at "moderator variables" (p. 322). He also recommends the use of multiple dependent variables, and replicability considerations based on effect size (p. 322).

These advances would be (or have been) constructive for parapsychology, but a larger problem may lurk in the wings. If statistical procedures are flawed because classical statistics does not offer a valid model of reality, then *all* the claims of science, not just those of parapsychology, are in jeopardy in so far as they employ just such procedures. To use Mosteller's (1991, p. 396) words: " if there is no ESP, then we want to be able to carry out null experiments and get no effect, otherwise we cannot put much belief in work on small effects in non-ESP situations."

Burdick and Kelly (1977) for outlines of the statistical analyses of forced-choice and free-response data (defined in Chapter 3).

[24] In defense of the statistical procedures used in parapsychology, see Mauskopf and McVaugh (1980).

Like researchers in other fields, the parapsychologist must continue as before in the belief that the "point null hypothesis makes some sense" (Utts, 1991 p. 401). But, to be realistic, is such a criterion enough to satisfy even the most hardened skeptic? The likely answer is generally a resounding "No!" because often the skeptic believes only when everybody else believes. In order to justify this seemingly facetious remark, one has to consider the social ramifications of statistical findings. In some cases particular methodologies are not even given due recognition by reviewers of scientific journals, or the scientific community in general, regardless of the claimed validity of these methodologies (see Greenhouse, 1991).

The issue of prejudice surfaces once again when "prevailing substantive beliefs and theories held by scientists at any given time" work against the parapsychologist (Greenhouse, 1991, p. 387). When Jung (1960, para. 821) implied that "the possibility of [rare] events" was a "rationalistic assumption" decided by the *consensus omnium*, he failed to mention that the process could sometimes be a biased procedure. Our age's 'rationalistic assumptions' are based on the epistemological foundations of our sciences, such as Newtonian physics, causality, etc. They form the commonsense basis by which we think and form our ideas, and we must remember, as Beloff (1990) claims, that a commonsense approach means doing "justice to the evidence while, at the same time, seeking to do the least violence to our reason and our general knowledge" (p. 13). (As it happens, relativity theory and discoveries in quantum physics are slowly modifying our commonsense approach, as well as our reason and general knowledge.)

Psi phenomena, therefore, are irrational according to generally held scientific precepts, and that adds to their anomalous nature. But it is to the highest degree ironic how the irrationality of emotions and other prejudices can lend support to arguments that criticize psi on the basis that psi is an irrational phenomenon (see Stokes, 1987, p. 84-90). Thus do some critics, perhaps many, try to have their cake and eat it too. It is even more ironic when we realize that the reasoned approach of statistical law can suggest the psi anomaly in the first place. Taking all the foregoing statements into consideration, we can see how experience, belief, and the scientific method may all intertwine and give us much food for thought.

At the end of this necessary first chapter, any responsible investigator who considers these issues and takes them seriously must be left with the resounding feeling that social factors form the undercurrent of a seemingly unimpeded flow of scientific progress. These factors combined with the business of science as usual, give a socio-empirical structure to parapsychology as much as they do to science in general. The whole

field of science is a human endeavour that requires of us that we give heed to that fact consistently and thoughtfully.

1.4 Conclusion

In this introductory chapter it was shown that *specification* of psi phenomena—the scientific account of the nature of the subject matter of parapsychology—is no more or less a burden for parapsychologists than it is for investigators and theorists in any other discipline, when they honestly confront their own subject matter. The general feeling from skeptics, though, is that this is not the case. For them, invalidation of a phenomenon (e.g., psi) stems naturally and rightfully from the fact that preconceived notions must need serious scrutiny, but they rarely apply a reflexive criticism to the phenomena they hold as being manifestly self-evident and self-specifying—their own preconceived notions remain unscrutinized.

The principles of *replicability* and repeatability were introduced as conventional means of arriving at agreement about the *existence* of a phenomenon—more correctly, its statistical reality only—but the principles themselves and allegiance to them were actually shown to create the false impression that less replicable phenomena were unspecifiable and probably even artifacts. Since that conclusion is not always true in science, it follows that a socially constructed 'myopism' or prejudice exists against anomalous phenomena based on principles that are anything but scientific. For all that, the principle of replicability still holds a certain validity and workability so long as we can arrive at some kind of consensus as to what it is that has been replicated. It was concluded, therefore, that when a phenomenon is specified in some way, that way of specification is inherently constructed, and is therefore a socio-empirical process of definition making.

It was further considered that thought must be given to the nature of any given paradigm. It is only an assumption that disparate systems of thought must necessarily be irreconcilable or incomparable with each other. Scientific knowledge consists of ideas and theories that complement each other as much as contradict. Very often they are merely discursively different—they even speak from the same core ideas and values. Given these considerations, it was argued that scientists must put their concepts in the balance and identify the discontinuities for what they may be—the result of social actions based more on opinion and personal or in-group prejudice rather than fact.

This chapter also discussed the ways in which three factors—experience, belief and the scientific method (statistics in particular)—have given shape and form to the field of

parapsychology and its paradigms (including the psi construct). These three factors also influence the researcher's attitudes and prejudices and therefore, have a direct influence on the future of parapsychology. The social repercussions, then, of experience, belief, and the scientific method are of direct concern to the field of parapsychology because these factors ultimately influence the field's social status as a discipline.

The same three factors were also claimed to be persuasive mechanisms that can influence 'closed-minded' skepticism (and indeed any belief system). Skepticism can thereby be somewhat ameliorated through the processes of a direct experience in the field (preferably of the paranormal type), a more centred consideration of belief formation (incorporating a critical or common sense approach to the nature of belief), and improvements in scientific methodologies.

To resurrect an earlier theme, it was noted again how the elusive nature of psi gives cause for skeptics to undermine parapsychological research—they believe that the study of 'unqualified anomalies' is a contradiction in terms, so that psi must therefore be unbelievable. While skepticism, wherever it exists, was just argued as being a somewhat myopic outlook on all kinds of anomalous phenomena, skepticism can indirectly be a productive force in its own right—even in parapsychology—because its critical approach to the procedures and methodologies used to investigate those phenomena serves to identify faults that may exist in parapsychological practice.

1.5 Epilogue

Partly by way of a criticism of hard-line skepticism, this first chapter has been an attempt to lay the foundations for the main theme of this thesis: the investigation of psychopraxia. I have withheld a full-scale launch into the theory of psychopraxia until Chapter 4 because a failure to pinpoint the particular issues that have faced parapsychology in the past, and still do in the present, would be to present the theory of psychopraxia without a context. That would be an injustice to the theory for one very important reason: the theory of psychopraxia (a theory that describes the nature of the 'psyche at work' manifesting its 'goals' by way of a 'pro attitude' and concomitant necessary conditions) wrestles with the same problems that parapsychology has long wrestled with—Chapter 1 is about some of those problems. It would not provide the theory of psychopraxia with the opportunity of fair representation, when put through the rigors of an investigation, if the investigator failed to make sense of the major problems facing parapsychology in general, and the theory of psychopraxia in particular.

These problems challenge the parapsychologist, and through that challenge emerges one of the active ingredients of science—change. Science, as an evolutionary and dynamic process, depends on, and is represented by, constant change, and change characteristically comes about because new goals are constantly being set. The theory of psychopraxia emerged out of the same processes—like any theory it exemplifies and crystallizes the idea of science as a process of necessary change brought about by the challenges set by problems such as those addressed above.

The next two chapters continue the case for the psi hypothesis, and give justification for the investigation of the theory as psychopraxia. Those two chapters will be a review of the meta-analytic literature in parapsychology. The broader, conceptual and practical issues of meta-analysis—a relatively new and distinct form of statistical procedure—will also be reviewed. Like statistics in general, meta-analysis has maintained the evidence for an anomalous effect that we may call psi, though we must also acknowledge, as already stated, that such evidence should not do "violence to our reason and our general knowledge" (Beloff, 1990, p. 13).

CHAPTER 2
META-ANALYSIS IN PARAPSYCHOLOGY: I. THE GANZFELD DOMAIN

At any point in time some judgment can be made.
Robert Rosenthal (1986, p. 333)

We have to wait for future attempts at replication to see if a replicable psi effect is at hand.
Ray Hyman (1994, p. 24)

2.1 Introduction to Meta-Analysis

Meta-analysis, a term coined by Glass (1976), is "the statistical analysis of the summary findings of many empirical studies" (Glass, McGaw, & Smith, 1981, p. 21). By combining the data from a number of different studies, meta-analysis produces an increased data set that may be analyzed to yield a more powerful result. Glass et al. (1981, pp. 22-23) describe the characteristics of meta-analysis:

(i) "Meta-analysis is quantitative" (it organizes and extracts "information from large masses of data that are nearly incomprehensible by other means") (Glass et al., 1981, p. 22).

(ii) "Meta-analysis does not prejudge research findings in terms of research quality." This policy runs counter to the actions of critics of meta-analysis who would like to exclude many studies on the methodological grounds that they are poorly designed or treatments are badly implemented, even though "evidence is never given to support the assumption that these deficiencies . . . influence their findings" (Glass et al., 1981, p. 22).

(iii) "Meta-analysis seeks general conclusions," and "aims to derive a useful generalization that does not do violence to a more useful contingent or interactive conclusion" (Glass et al., 1981, pp. 22-23).

(Naturally, these assumptions have been criticized, and these assumptions will be addressed in the next section.)

The importance and usefulness of the 'meta-analytic' approach was demonstrated decades ago by prominent statisticians, psychologists, and medical researchers (Cochran, 1954; Edgington, 1972a, 1972b; Fisher, 1932; Mantel & Haenszel, 1959; Mosteller & Bush, 1954; Tippett, 1931). The flexibility and diversity of applications of meta-analysis have also been well reviewed and described.[25]

Bullock and Svyantek (1985) point out that meta-analysis is "objectively verifiable, using measured concepts, quantitative data, and statistical analysis" (p. 112). As opposed to the traditional literature review, where the reviewer is prone to his or her own "subjective interpretation of results across studies" (p. 112), the reviewer's selection bias in a meta-analytic review is more explicit, since the inclusion and exclusion of particular domains (i.e., 'experimental types' in the context of the present study) is so readily apparent. The advantage for the reader is that the philosophical persuasion and/or general intention of the author can be determined with ease so that the overall merits of the work can be more easily assessed.

As mentioned, meta-analysis produces results that are not attainable by other means (at least, not without difficulty), but even these results may not end the debate over a specific area of controversy. Meta-analysis can never be conclusive. It provides a critical examination of the current status of research in a given area. However, by finding flaws in current meta-analytic techniques, and by identifying the more successful experimental domains through up-to-date meta-analytic procedures, newer, more focussed directions for parapsychology can be established, which, in the case of the parapsychologist, may help "settle the question" of whether or not psi exists (Broughton, 1991, p. 284).

2.2 Parapsychology and Meta-Analysis

As will be seen in the next section and Chapter 3, meta-analysis has been a godsend for parapsychologists. Specifically, the direct benefits of meta-analysis for parapsychology become clear once it is understood that meta-analysis is the means by which the "signal" of

[25] For examples, see Glass et al.'s (1981, pp. 25-26) meta-analyses of studies relating to the efficacy of psychotherapy in treatment of asthma and alcoholism. See also the meta-analyses by Cook et al. (1992) of juvenile delinquency studies, and studies on psychoeducational care with adult surgical patients, among others.

the psi effect can be 'distilled' from the "noise" of chance with greater sensitivity than in any single study (Broughton, 1991, p. 281-282).[26]

Closely related to meta-analysis is the concept of replication (introduced in Chapter 1, see *1.2.1*). Rosenthal (1986) lists three main reasons for the so-called failure of many single studies to elicit significant psi effects, after previous successes with the same experimental design:

(i) Pseudo-failure to replicate due to a poor consideration of the appropriateness of the statistical test(s) used. (See the illuminating example of a comparison between effect size and significance levels given by Rosenthal, 1986, pp. 317-318.) Essentially, the investigator may be too dependent on the significance (or not) of the *p* value.

(ii) Pseudo-successful replications, where *p* values are less than .05, but the effect sizes are significantly different, and therefore have not been replicated.

(iii) Successful replication of Type II error (due to small sample size, or very weak effects).

Consideration of the replication issue has led to an expansion of the meta-analytic methodology and a refinement of its technique (see Chapter 1 for an alternative 'take' on the replication problem—*1.3.1*). Replication cannot be guaranteed, but it does come in forms other than rejection of the Null hypothesis (for example, replication of effect sizes, z scores, etc.).

Meta-analysis is fast becoming the *only* acceptable evidence that psi might well be an "anomalous effect in need of an explanation" (Utts, 1991, p. 363). However, there is now a growing suspicion of meta-analysis, not simply because it provides suggestive evidence of psi, but because in controversial non-parapsychological fields, such as psychotherapy, meta-analysis is also providing favourable evidence (Smith & Glass, 1977). Thus have some critics (Bandura, 1978; Eysenck, 1978; Oakes, 1986; Shapiro & Shapiro, 1977) focused on the construct validity of meta-analysis. Arguments are made

[26] In fact, the presence ('signal') of a paranormal effect can never be specifically demonstrated as and when it occurs, even if 100% of trials are hits. For example, an experiment by Rhine (1937/1950, pp. 74-76) showed 25 hits in 25 trials, but in such an experiment there is no means of distinguishing the successful calls that suggest psi effects from those calls attributable to chance. The statistical inference of a psi effect is always made in the knowledge that the effect is produced within a stochastic framework.

that data from methodologically flawed experiments are thrown into the 'soup' with the data from better-designed experiments, therefore corrupting the result. Different procedures in ostensibly the same kind of experiment may also yield a tainted result.

Of course, the general implication derived by critics (that meta-analysis is a flawed procedure) also applies to meta-analysis in parapsychology. But these criticisms have not been ignored, and improvements have been made to control for the confounding factors. Rosenthal (1984) advocates differential weighting as an effective way of dealing with "variation in the quality of research" (p. 127). Hence the 'blocking' procedure is used to code experiments according to their 'quality' and type (i.e., methodology, hypothesis, etc.).

Sample size and the population from which the sample is drawn are also critical considerations. Credit is given to studies if sample size is specified in advance, as well as the nature of the analyses—pre-planned or post hoc. Acceptable randomization methods are also credited, and even the date of the experiment and the identity of the investigator are now important criteria in meta-analytic studies (Broughton, 1991, p. 283). Although these procedures may be seen as subjective, some degree of qualitative assessment can be made about studies, and these assessments are converted to numerical values to arrive at a more objective, albeit pseudo-precise, numerical result that is still seen as a gain over previous methods which did not consider study quality.

One important criticism made against meta-analysis in parapsychological research (applicable to meta-analysis in general) is that significant results are inevitable, since the majority of studies used in the analysis have significant results already. On the other hand, studies with nonsignificant results are never or are rarely published and, therefore, cannot be included in the meta-analysis (Hyman, 1985a). There are two facts that belie this criticism.

First, parapsychology journals go to great efforts to publish studies with nonsignificant results—they tend not to end up in the 'file-drawer'. In 1975 the Parapsychological Association Council adopted a policy of opposing the exclusive publication of studies with positive outcomes. Thereby, "negative findings have been routinely reported at the association's meetings and in its affiliated publications" since that date (Bem & Honorton, 1994, p. 6; see also Honorton, 1985a, p. 66.)

Second, estimates can be made which account for unpublished studies. Usually, the number of nonsignificant studies that would be needed to reduce a significant meta-analytic result to a chance outcome is shown to be far in excess of that which would be possible for the few researchers in the field of parapsychology (Broughton, 1991, p. 286; Utts, 1991, pp. 370, 372, 375-376).

Rosenthal (1984) has also addressed many criticisms leveled at meta-analysis, including exaggeration of significance levels, which can be a problem since increasing the number of studies in a meta-analysis increases the probability of rejecting the Null hypothesis. Rosenthal argues that when the Null is false it ought to be rejected, but notes that if such a characteristic of meta-analytic procedure increases its accuracy and decreases the likelihood of Type II errors, it must be an advantage. Alternatively, should it be possible that meta-analysis increases the likelihood of Type I errors (i.e., when the Null is really true), an increase in the number of studies does not increase the probability of rejecting the Null, or the size of the estimated effect.

Finally, Oakes (1986) questions the validity of a procedure that doesn't really test for a 'directional hypothesis' (e.g., "do men perform better than women?" Oakes, 1986, p. 162). For parapsychology, at least, domain-specific meta-analyses clearly refer to directional hypotheses. Generally, what are sought in the meta-analyses of these specific domains are significant effect sizes, significant differences in psi performance (above or below mean chance expectation), or significant differences in scoring between comparison groups.

The dubious value of "meta-meta-analysis," criticized by Oakes (1986, p. 162), but endorsed by Glass (1981, cited in Oakes, 1986, p. 162) as a reasonable undertaking, is also acknowledged (Glass, McGaw, & Smith, 1981, p. 218, believe it is a valid exercise to mix studies on "apples and oranges" if your hypotheses are about fruit!). Such a 'method' may be superfluous and even of no validity, furnishing (it would seem) a rather vague, nebulous, and therefore, ambiguous finding, whether significant or not.

As it happens, parapsychology is bereft of such adventurous undertakings. The idea, for example, that ganzfeld studies could be meta-analyzed with dice-throwing studies is untenable in the extreme. For example, a significant result in one single meta-analysis of both domains combined would not draw out the possible effect size difference between these two domains. Should there be a significant effect size difference, it would only draw out the distinction that either two types of psi were in operation, or the paradigms are more conducive to psi in one domain, and less conducive in the other. Thus there are limits to how far we can apply Glass et al.'s fruit analogy.

Having described the current status of meta-analysis, and argued that meta-analysis has a relative degree of validity, and therefore, that there is scientific value in its processes, the following section is a presentation of the findings of the meta-analyses of ganzfeld studies dating back to the mid-1970s.

2.3 Review of the Ganzfeld Meta-Analyses

The meta-analyses now reviewed used as their sources bibliographic databases for parapsychology, and all the parapsychological journals, including publications of technical reports, conference proceedings, and manuscripts, etc., and, in some cases, physics journals (e.g., Radin and Nelson, 1989).

2.3.1 The Ganzfeld Procedure

The ganzfeld is a form of free-response test—'free response' being a term that "describes any test of ESP in which the range of possible targets is relatively unlimited and is unknown to the percipient" (Thalbourne, 1982, p. 28). The target is not restricted to a few choices, but can be almost anything, thus hopefully reducing the risk of boredom so common in forced-choice experiments because free responses "more nearly resemble the conditions of spontaneous psi occurrences" (Burdick & Kelly, 1977, p. 109).

The ganzfeld is a "special type of environment (or the technique for producing it) consisting of homogeneous, unpatterned sensory stimulation" to the eyes and ears of the participant who is usually in "a state of bodily comfort" (Thalbourne, 1982, p. 29). A number of investigators pioneered the technique in the 1970s (Braud, Wood, & Braud, 1975; Honorton & Harper, 1974; Parker, 1975).

Procedurally, the eyes of the participant are covered with halved ping-pong balls illuminated by a uniform source of light (usually of a single wavelength, such as red light). A uniform auditory signal of "white" noise (full-range audio signal; see Thalbourne, 1982, p. 29, and Utts, 1991, p. 369) or "pink" noise (high-frequency filtered sound; see Stanford, 1979, p. 253) is channeled through headphones to the ears. The participant reclines on a chair or lies on a bed. This technique has remained essentially the same since the 1970s.

2.3.2 The Ganzfeld Experiments

The first major meta-analytic study in parapsychology started in 1981 when Hyman (1985b) began evaluating 42 ganzfeld psi studies conducted during the period 1974 to 1982. Hyman initially chose the ganzfeld studies because they supposedly held a "high level of research sophistication and rigor" (Hyman, 1985b, p. 4)—a claim that Hyman was to criticize heavily.

A drawn-out debate ensued between Hyman and Honorton, since they arrived at conflicting conclusions from the same data set. Hyman first argued that the "alleged" 55% success rate of 42 studies determined from a vote-count made by Honorton (Hyman,

1985b, p. 5) was inflated due to the fact that many of the studies were not independent (they were more like subsets of ongoing experiments).

Hyman also cited evidence that suggested bias in how the studies were reported. For example, some studies were not planned as such, but were "given this status retrospectively just because they yielded significant results" (Hyman, 1985b, p. 16). Hyman reduced the success rate to 31% (he actually argued for less than 30%, given that there must be unsuccessful but unknown ganzfeld studies yet to be considered—the 'file-drawer' problem mentioned above).

Hyman (1985b) further criticized many of the studies for their multiple analyses (e.g., use of a number of measures of ESP), which gave increased opportunity for a good result, especially since investigators were not adjusting their criterion significance levels according to the number of statistical tests they performed. He also claimed that independence had been violated in some meta-analytic studies because "agents were friends of the percipient . . . [or were even] members of laboratory staff" (1985b, p. 26).

Of interest is Child's (1986, pp. 337-343) comment on procedural flaw where pooling results based on groups or conditions can actually conceal an effect rather than erroneously identify one. Optimal randomization could not be assumed for such studies. Child indicated that hit-rates can vary systematically from individual to individual, or group to group, so that "genuinely high performances of some may well be buried by the chance performance of many others" (1986, p. 339).

Honorton (1985a) accepted the criticism of multiple analysis, and he applied a Bonferroni correction across all studies. He found that only 45% of the 42 studies were significant—not 55% which he originally claimed (but 45% was still higher than Hyman's lower estimate of 31%). Honorton then used the proportion of direct-hits as a common index, since it was the most common measure in the studies (also the most conservative). A total of 28 studies using direct hits alone were thus employed in the meta-analysis, 12 (43%) of which were significant at $p \leq 05$ (see Table 2.1).

Honorton (1985a) noted that of the 28 studies, 23 (82%) had positive z scores. (The probability of this outcome is shown in Table 2.1, along with the number of studies, also expressed as percentages, and their significance levels.) Honorton reported a composite Stouffer[27] *Z* score of 6.60 across the 28 studies.

[27] "Stouffer's *Z* is found by dividing the sum of the *z* scores for the individual studies by the square root of the number of studies" (Rosenthal, 1978, p. 6).

Table 2.1
Meta-Analysis[28] *of the 28 Direct-Hit Ganzfeld Studies (and their Subgroups)*

Number of studies	Independent *p* value	Proportion of Hits	Effect Size (π)	Stouffer *Z*	Probability (*p*)
7 (25%)	.01	0.47	0.73	8.63	9.80×10^{-9}
12 (43%)	.05	0.46	0.72	10.46	3.50×10^{-9}
23 (82%)	n/a[a]	0.40	0.67	8.42	4.60×10^{-4}
28 (100%)	n/a[b]	0.38[c]	0.65	6.60	2.10×10^{-11}

Notes: [a] Studies with positive *z* scores (Exact binomial test with $p = q = .5$)
[b] Includes five studies with negative *z* scores
[c] Rosenthal's (1986, p.333) more conservative estimate of the proportion of hits is "about 1/3." Bem and Honorton (1994, p. 8) give a value of .35, and thus calculate π as .62

Table 2.1 also includes effect size measures as π values because of their ease of interpretation—π "depends simply on *k*, the number of alternative choices available, and *P*, the raw proportion of hits"[29] (Rosenthal & Rubin, 1989, p. 333). Using the mean effect size (*ES*) formula, $\Sigma[z/\sqrt{n}]/k$, the *ES* for the 28 studies was .26.

Honorton (1985a, p. 59) also calculated a more conservative estimate of significance by including 10 additional blind-judging studies that did not report direct-hit information. Assuming a mean *z* score for these 10 studies of zero, the Stouffer *Z* becomes 5.67, $p = 7.30 \times 10^{-9}$. Such a probability still indicates how extremely unlikely it would be that these successful ganzfeld studies were all the result of chance.

Using the blocking technique, six of the ten independent investigators who produced these studies achieved significant results, so that neither a specific investigator, nor a specific laboratory was single-handedly responsible for the significant results. The suggestion of a file-drawer problem was also rendered less plausible by the fact that 15 nonsignificant and unknown studies would have to exist for every one of the 28 direct hit studies to reduce the result to a chance outcome.

But other problems had to be addressed. Hyman (1985b, pp. 30-35) found that flaws correlated positively with significant results. He identified 12 major flaws, such as inadequate randomization of targets, and failure to use a duplicate set of targets for judges. Hyman (1985b, pp. 35-36) used cluster and factor analysis on these 12 flaws, combining them into 3 new variables: General Security, Statistics, and Controls—upon which were

[28] Most of these data come from Honorton (1985a), or are calculated from the data provided in that study.

[29] $\pi = P(k - 1)/[1 + P(k - 2)]$. Bem and Honorton (1994, p. 8) point out the advantage this measure has in providing a "straightforward intuitive interpretation" of the effect size, because π is the "proportion correct, transformed to a two-choice standard situation" so that $P_{MCE} = P_{test} = .50$ (Rosenthal & Rubin, 1989, p. 333).

conducted several analyses. The most detailed (factor) analysis was one consisting of supposedly 17 variables, from which emerged four factors (actually, there were only 16 variables, according to Saunders, 1985, p. 97). Utts (1991) paraphrased the findings of this analysis:

> From these [four factors], Hyman concluded that security had increased over the years, that the significance level tended to be inflated the most for the most complex studies and that both effect size and level of significance were correlated with the existence of flaws. (p. 371)

Hyman's adjusted figure for the number of successful studies which would not possess these flaws was 27% of all the studies considered—"well within the statistical neighborhood of the 25% chance rate" (1985, p. 37). Honorton acknowledged the problems with the studies, but Saunders (1985), on behalf of Honorton, repudiated Hyman's "meaningless" analysis and its "logical problems" (Saunders, 1985, p. 87). Saunders found a violation in statistical procedure in Hyman's factor analysis: "the size of the available database marginally suffices to support [only] one factor" (p. 87).

Hyman had performed a multiple analysis that included the three flaws just mentioned, but out of nine potential flaws (giving 84 sets of three) Hyman (conveniently) selected the set that correlated highly with effect size. Thus the impression was given that, as Hyman implied (1985b, p. 37), effect size was a function of procedural flaws (the more flaws in an experiment, the higher the effect size). Saunders noted that Hyman's multiple correlations that resulted from selective testing should be regarded as nonsignificant, rendering Hyman's adjusted figure of 27% meaningless.

Rather than continue the debate, Hyman and Honorton produced a "Joint Communiqué" (Hyman & Honorton, 1986) addressing fundamental issues in parapsychological experimentation. The Communiqué recommended that "more stringent standards" be implemented in experiments, which should also be conducted by a "broader range of investigators" (p. 351). Utts (1991 listed these standards as including:

> controls against any kind of sensory leakage, thorough testing and documentation of randomization methods used, better reporting of judging and feedback protocols, control for multiple analyses [and statistics] and advance specification of number of trials and type of experiment. (p. 371)

Hyman and Honorton (1986) also believed that meta-analysis had a growing role in the evaluation of "research quality and the assessment of moderating variables" (p. 361).

A number of researchers commented on the "Joint Communiqué" and most were in general agreement with its recommendations, though all had unique points to make about the state of affairs of parapsychology. Hövelmann (1986, p. 366) felt that the participant should be left alone at the judging stage (no presence of the experimenter), since even the non-verbal behavior of a non-blind experimenter may have an influence on the judging outcome. Usually, the use of 'blind' experimenters avoids this problem.

Palmer (1986, p. 379) argued that the presence of Hyman's identified flaws in a ganzfeld experiment does not mean replication of positive results will continue in the future. He added that the absence of flaws would not necessarily guarantee positive results either. Nor should it be assumed that failure to replicate when the flaws are removed means that past successes were due to the presence of the flaws.

Stanford (1986, p. 384) expressed his unease about both Hyman's and Honorton's readiness to make a *cause célèbre* out of the ganzfeld by yielding it up to the National Science Foundation for extensive replicability studies. The worst-case scenario of dismal failure could damage the field of parapsychology, not to mention the careers and professional lives of parapsychologists. It is too early for parapsychologists to be so confident when ganzfeld-ESP success looks more like "art" than "science" (p. 386).

Utts (1986), as a statistician, felt that power considerations must be undertaken more often in experimental design since the "replicability" problem in so many experiments can be due to poor consideration of the sample sizes needed in certain experiments. Real effects can be lost if N is too small, whereas a larger N (as in meta-analytic studies) only increases the chance of getting a significant result.

Rosenthal's (1986) commentary on the meta-analysis also focused on replication, and so included a consideration of effect size. He used Cohen's h, which is the transformed proportion of direct hits.[30] Rosenthal calculated that 23 (82%) of the direct hit studies had effect sizes greater than zero. The mean effect size h was 0.28, corresponding to the significant direct-hit-rate of 0.38 reported in Table 2.1 (where $P_{MCE} = .25$). Rosenthal recommended that effect size be considered a more favorable measure of replication success over and above that of significance testing which has "nothing to do with success of replication" (1986, p. 334).

[30] Cohen's $h = 2(\arcsin \sqrt{p'} - \arcsin \sqrt{p})$, where p' is the proportion of observed direct hits, and p is the proportion of expected number of direct hits.

Milton (1997) conducted one recent meta-analysis relevant to the ganzfeld domain. She meta-analyzed 46 free-response[31] studies (including 42 ganzfeld studies) to determine which measure—direct hits or sums of ranks—was the more sensitive of the two. The cumulative result of the 46 studies was significant, suggesting a psi effect. In considering only effect sizes and *p* values, sums of ranks "outperformed" direct hits (Milton, 1997, p. 227). However, there was no statistically significant degree of difference between the two measures. Milton called for caution until further research might be more conclusive about which of the two techniques should be considered 'superior', since the nonsignificant difference may have been a chance result.

2.3.3 The Autoganzfeld Experiments

The automated ganzfeld (i.e., "autoganzfeld") procedure was adopted as a more rigorous approach to psi testing, while still maintaining the ganzfeld paradigm. It came into being as a proactive response to the "Joint Communiqué" recommendations. Thus some strict guidelines were implemented in the autoganzfeld, the major one being the introduction of a computer-controlled, randomly-selected, presented, and scored target, which was therefore unknown to all those involved in the experiment except the sender. Feedback is eventually given to the receiver in the form of the correct choice. As in the ordinary ganzfeld, targets can be "dynamic" (short scenes from movies, cartoons, and documentaries, etc.), or "static" (photographs, art prints, advertisements, etc.).

A series of 11 autoganzfeld experiments was conducted by eight experimenters during the period 1983-1989 (see Honorton et al., 1990). As reported in Bem and Honorton (1994) there were 106 direct hits in 329 trials for 10 of these studies yielding a 32.2% hit-rate (P_{MCE} = .25), Stouffer $Z = 2.61$, $p = 4.50 \times 10^{-3}$ (mean $ES = 0.117$; $\pi = 0.59$). The π value for this series of experiments is comparable to the π value of .62 given by Bem and Honorton (1994, p. 8) for the 28 direct-hit non-automated ganzfeld studies. The eleventh study, which used dynamic targets exclusively, had the highest hit-rate (54%) and was in fact significantly higher than any of the other ten studies. The study was rejected by Bem and Honorton (1994) due to "response biases" (pp. 11-12).

The hit-rate for all the 'dynamic' target studies (164 sessions) was 37% ($\pi = 0.74$), a considerable difference compared to the hit-rate for all the 'static' target studies (165 sessions) of only 27% ($\pi = 0.52$). Bem and Honorton (1994, p. 12) note these differences

[31] The free-response method describes any test of ESP using a relatively unlimited range of possible targets, thus permitting the participant to "respond freely with whatever impressions come to mind" (Thalbourne,

as suggestive evidence that, generally, dynamic targets may be "more effective than static targets." There was also a distinct difference in success rate between experienced participants and novices, suggesting that experienced participants (those previously tested) yield better results than novices (experienced participants' hit-rate: 37%, $\pi = 0.64$; novices' hit-rate: 32.5%, $\pi = 0.59$).

When Rosenthal (1986, p. 333) adjusted for the flaws in the earlier ganzfeld studies, he arrived at a conservative estimate of "about 1/3," thus reducing the original 38% hit-rate to a hit-rate roughly equivalent to the 10 autoganzfeld studies of 32.2%. The ordinary ganzfeld and the autoganzfeld appeared to be equally effective, since they produced effect sizes in roughly the same vicinity ($\pi = 0.62$, $\pi = 0.59$, respectively). After two nonsignificant performance comparisons (on effect sizes and *z* scores) between Honorton's (1985a) database of 28 studies and the new Honorton et al. (1990) database of 11 autoganzfeld studies, Honorton et al. (1990, p. 99) combined the two databases into a 39-study database, Stouffer $Z = 7.53$, $p = 9.00 \times 10^{-14}$ (Cohen's $h = 0.28$).

One criticism leveled at the autoganzfeld meta-analysis was that the eleven experiments were conducted by only eight experimenters, all of whom were at the same laboratory. Consequently, Milton and Wiseman (1999) conducted a meta-analysis of new ganzfeld studies dating from 1987 to 1997. Studies prior to 1987 were not used because it was assumed that investigators needed time to familiarize themselves with Hyman and Honorton's (1986) guidelines so that earlier studies would be too flawed for serious consideration in a meta-analysis.

Milton and Wiseman (1999) deemed suitable for analysis thirty studies by "10 different principal authors from 7 laboratories" (p. 388). They calculated a Stouffer *Z* of 0.70, $p = .24$, one-tailed ($ES = 0.013$), and concluded that a significant psi effect for the ganzfeld had not been replicated by a "broader range of researchers" (p. 391).

Storm and Ertel (2001) singled out Milton and Wiseman's (1999) main finding of a nonsignificant *ES* of 0.013 and disputed its derivation. Storm and Ertel argued that Milton and Wiseman did not adopt a 'responsible' attitude in their meta-analysis. A thorough meta-analysis requires a comprehensive literature search and an accumulative approach to the available databases (this approach was not adequately demonstrated in Milton & Wiseman's paper). Arbitrary exclusion rules and unjustifiable, a posteriori periods of analysis should be considered unacceptable in any meta-analysis.

1982, p. 28). The participant may, for example, respond by drawing a pictorial representation of the target.

Storm and Ertel (2001) found 11 pre-Communiqué studies not previously meta-analyzed, and after step-by-step performance comparisons, combined them with the three ganzfeld databases currently extant: Honorton's (1985a) database of 28 studies, Bem and Honorton's (1994) databases of 10 studies, and Milton and Wiseman's (1999) database of 30 studies. The resulting 79-study database had a significant mean *ES* of 0.14 ($Z = 5.66$, $p = 7.78 \times 10^{-9}$). Milton and Wiseman's negative conclusion about the failure of the ganzfeld to replicate is rather misleading and premature as it is pertinent to a limited pool of only 30 studies.

In reply to Storm and Ertel (2001), Milton and Wiseman (2001) argued that the 11 pre-Communiqué studies used in their meta-analysis should not have been used at all because they were poor in quality due to their ostensible "methodological problems" (p. 434). Thus, Milton and Wiseman clearly ignored Storm and Ertel's (2001) performance comparisons of (a) pre-Communiqué authors with post-Communiqué authors, and (b) pre-Communiqué studies with post-Communiqué studies, both of which yielded no statistical evidence that the guidelines in the Communiqué had any "influence on effect size outcomes" (p. 430). Logically, there was no indication that the mean effect size of the pre-Communiqué database was 'inflated' (i.e., an artifact of flaws) because it compared favourably with the allegedly 'flawless' post-Communiqué studies. It follows that there was no evidence that the mean effect size of the post-Communiqué database was 'deflated' due to the removal of these flaws. Apropos to these findings, Palmer (1986) warned that false conclusions can be drawn on account of, and by appeal to, the Communiqué's guidelines—it should not be assumed that "past successes were due to the presence of the flaws" (p. 379).

Milton and Wiseman's (2001, p. 436) only other major criticism concerned the lack of conservative calculations of some *z* scores for studies in the 11-study database. In fact, only 3 of the 11 studies needed adjustment, thus reducing the quality-weighted mean *z* score from 0.32 (*ES* = 0.14; Stouffer $Z = 1.06$, $p = .144$) to 0.26 (*ES* = 0.13; Stouffer $Z = 0.87$, $p = .192$). The 11-study database is still not significantly different from Honorton's (1985) 28-study database, $t(37) = 0.61$, $p = .543$, two-tailed. The Old Ganzfeld Database can still be formed. It has a mean *z* of 0.97 (*ES* = .225; Stouffer $Z = 6.05$, $p = 7.24 \times 10^{-10}$; cf. Storm & Ertel's, 2001, p. 429, original data for the Old Ganzfeld Database: mean *z* of 0.99 (*ES* = .227; Stouffer $Z = 6.15$, $p = 3.93 \times 10^{-10}$).

The 'Old' (Pre-Communiqué) and the 'New' (Post-Communiqué) ganzfeld databases are significantly different, $t(77) = 3.04$, $p = .003$, $\omega^2 = 0.09$, but the omega squared value (9%) is now exactly that of the critical value stipulated in Storm and Ertel's

(2001) paper. The difference might be considered important, but Cohen's (1988) test, as originally applied by Storm and Ertel, was again not significant. When the two databases are combined the 79-study database has a mean z score only slightly reduced from 0.64 to 0.63 ($ES = 0.14$; Stouffer $Z = 5.59$, $p = 1.14 \times 10^{-8}$).

This larger database has not only unified the ganzfeld database, but also indicates that over two decades of ganzfeld/autoganzfeld work has not been in vain. The ganzfeld might even prove to be the ideal paradigm that Honorton hoped it might be for finding "strong evidence for psychic functioning" (Milton & Wiseman, 1999, p. 391).

2.4 Conclusion

This chapter reviewed the ganzfeld meta-analytic literature and found that significant results were obtained in all but one ganzfeld meta-analysis—that of Milton and Wiseman (1999). However, with combinatorial re-construction of the available databases and the uncovering of 11 studies overlooked by Milton and Wiseman, Storm and Ertel (2001) reconfirmed that the ganzfeld was still the paradigm that delivered one of the highest effect sizes of all the experimental domains in parapsychology.

Thus, excluding the one disputable meta-analysis, there has not been one meta-analysis of the ganzfeld domain to date that has not supported Honorton's (1985a, p. 81) claim that the ganzfeld demonstrates a "significant psi effect." Even more importantly, results from the meta-analyses have indicated, as Honorton also believed, that the ganzfeld represents an 'encouraging' step toward replicability of psi effects. In the next chapter, a review of the meta-analyses continues into non-ganzfeld domains.

CHAPTER 3
META-ANALYSIS IN PARAPSYCHOLOGY: II. NON-GANZFELD DOMAINS

The overall evidence [from the meta-analyses] indicates that there is an anomalous effect in need of an explanation.
Jessica Utts (1991, p. 363)

Either something spooky is going on, or it is possible for a field to exist on error and artifact for over 100 years.
Persi Diaconis (1991, p. 386)

3.1 Meta-Analyses of Non-Ganzfeld Domains

This chapter is a review of the meta-analytic literature for the non-ganzfeld domains: biological systems, forced-choice (precognition and clairvoyance), free-response (GESP/remote-viewing), dice-throwing, and micro-PK (RNG). Lawrence's (1993) meta-analysis of the sheep-goat effect, and Haraldsson (1993) meta-analysis of various belief measures that correlate with ESP are reviewed. The meta-analyses of the psi-conducive moderator variables—hypnosis, perceptual defensiveness, and extraversion—are also reviewed. Finally, Jessica Utts's (1991) recommendations for future meta-analysis and parapsychological research are also considered.

3.1.1 Biological Systems

Braud and Schlitz (1991) conducted a 13-year-long series of studies that looked at eight living target systems—electrodermal activity (participant's *influence* on target system's skin resistance), electrodermal activity (participant's *attention* on target system), ideomotor reactions (reactions associated with thought), muscular tremor (measured by the movement of a hand-held metal stylus in a small aperture), blood pressure, fish orientation, mammal locomotion (gerbil activity in a wheel), and rate of hemolysis of human red blood cells. The goal in each case was to influence these systems to bring about "increments" or "decrements" in the "monitored systems' activities" (Braud & Schlitz, 1991, p. 2). Experiments in this domain were described as testing participants' "direct mental influence on living systems" or DMILS (Braud & Schlitz, 1991, p. 3).

Thirty-second "influence epochs" in a session were reduced to a single score (the unit of analysis). A *percent influence score* was then calculated, which was a percent measure of the "total activity that occurred in the prescribed direction during the entire set of influence (decremental or incremental aim) periods" (Braud & Schlitz, 1991, p. 5). A score of 50% (*MCE*) set the baseline for influence outcomes (no effect) and the *t* test was used to compare actual percent influence scores against this baseline.

Influence on remote biological systems was generally found to be significantly above chance on all target systems except muscular tremor, although a total of 19 sessions only were run, whereas the next lowest was 40 sessions, and the average number of sessions was over 65 (see Table 3.1 for results of individual studies). Meta-analysis of the eight studies showed a significant mean *ES* of 0.42 ($Z = 7.72$, $p = 2.58 \times 10^{-14}$). Braud and Schlitz (1991, pp. 31-34) discussed rival hypotheses that might explain these successful results, such as external stimuli, common internal rhythms, recording errors and biased misreading of records, participants' prior knowledge of when influence was to take place followed by appropriate responses, and even fraud. All these 'explanations' were disconfirmed.

Table 3.1
Summary of Results of Direct Mental Influence Experiments

Living target system	No. of sessions	Mean *z* score	Mean *ES*	Stouffer *Z*	*p*
Electrodermal activity (influence)	323	1.05	0.058	4.08	2.25×10^{-5}
Electrodermal activity (attention)	78	0.84	0.095	1.68	4.65×10^{-2}
Ideomotor reactions	40	1.72	0.272	2.98	1.44×10^{-3}
Muscular tremor	19	–0.42	–0.096	–0.59	7.22×10^{-1}
Blood pressure	41	1.35	0.210	1.91	2.81×10^{-2}
Fish orientation	40	1.88	0.297	3.78	7.84×10^{-5}
Mammal locomotion	40	1.90	0.300	3.81	6.92×10^{-5}
Rate of hemolysis	74	2.43	0.282	4.20	1.33×10^{-5}
All systems combined	655	1.34	0.420	7.72	2.58×10^{-14}

Note: Table reproduced from Braud and Schlitz (1991, p. 30, *Table IX*)

The overall conclusion was that "effect[s] appear to occur in a 'goal-directed' manner" (p. 41) because influencers were able to bring about effects without a specific understanding or awareness of how the physical or physiological processes brought about

the desired outcomes. Braud and Schlitz (1991) argued that participants' "intentionality" was the "key factor" that brought about the significant results":

> Maintaining a strong intention of a desired goal event, focusing attention upon the relevant aspect of the target system, and filling oneself with strong imagery of the desired biological activity are, under certain conditions, accompanied by a shift in the target system's activity in the intended direction (p. 41)

The idea of goal-directed or goal-oriented behavior is used in the theory of psychopraxia, and Braud and Schlitz's theoretical explanation for the above-chance effects on the above-mentioned biological systems, supports the 'pro attitude serving' hypothesis of Thalbourne's theory (see Chapter 4).

3.1.2 Forced-Choice Precognition

Forced-choice experiments require that the participant "guess a target which is one of a limited range of possibilities which are known to them in advance [such as in the card-guessing experiment]" (Thalbourne, 1982, p. 28). Forced-choice precognition[32] experiments from as early as 1935 up to 1987 were meta-analyzed by Honorton and Ferrari (1989).

Honorton and Ferrari (1989) used only the studies where the procedure was to select the target "randomly after the subject [*sic*] had attempted to predict what it would be" (Utts, 1991, p. 374). A total of 309 studies (62 of which were from "senior authors") were analyzed, amassing a phenomenal 50,000 participants and approximately two million individual trials. The $z/\sqrt{n}$ formula was used as the measure of effect size (*ES*). The *ES* was 0.02 (mean $z = 0.65$, all studies).

Ninety-two studies (30%) showed significant hitting at the 5% level. When outlier studies contributing to the heterogeneity of the database were removed the *ES* fell to 0.012 ($Z = 6.02$, $p = 1.10 \times 10^{-9}$). (See Table 3.2 for other results.) The "fail-safe *N*" (Rosenthal, 1984) was 14,268 studies, which would be needed in order to reduce the significant effect to a chance outcome (requiring 46 unreported and unsuccessful studies for every successful study). Naïve participants (where $h = 0.008$) did not perform as well as experienced participants ($h = 0.051$).

[32] Precognition is a form of ESP "in which the target is some future event that cannot be deduced from normally known data in the present" (Thalbourne, 1982, p. 55).

Table 3.2

Meta-Analyses of Nine Psi Domains: DMILS, Forced-Choice, Free-Response, Dice-Throwing, RNG, Clairvoyance, Precognition, Ganzfeld, and Autoganzfeld.

Domain meta-analyzed and author(s)	No. studies (k) and period of analysis	Mean z score ($\Sigma z/k$)	Mean ES ($\Sigma[z/\sqrt{n}]/k$)	Stouffer Z ($\Sigma z/\sqrt{k}$)	p
DMILS (Braud & Schlitz, 1991)	8 (1979-1991)	1.34	0.420	7.72	2.58×10^{-14}
Forced-Choice (Honorton & Ferrari, 1989)	248 (1935-1987)	0.38	0.012[a]	6.02	1.10×10^{-9}
Free-Response (GESP/Remote Viewing) (Milton, 1998)	75 (1964-1993)	0.68	0.170[a]	5.85	2.46×10^{-9}
Dice-Throwing (Radin & Ferrari, 1991)	59 (1935-1987)	0.42	0.003[b]	3.19	7.16×10^{-4}
RNG (Radin & Nelson, 1989)	490 (1959-1987)	0.42	0.0003 [b]	9.74	$\approx 10^{-23}$
Clairvoyance (Steinkamp et al., 1998)	22 (1935-1997)	0.60	0.009	2.81	2.50×10^{-3}
Precognition (Steinkamp et al., 1998)	22 (1935-1997)	1.02	0.010	4.78	8.80×10^{-7}
Ganzfeld (Storm & Thalbourne, 2000a)	71 (1974-1999)	0.71	0.154	5.98	1.12×10^{-9}
Autoganzfeld (Storm & Ertel, 2000)	17 (1983-1997)	0.58	0.117	2.35	9.39×10^{-3}

Note: The most conservative values were used in the table. DMILS = direct mental influence on living systems; *ES* = effect size; GESP = general extra-sensory perception

[a] Calculated from a homogeneous data set

[b] Quality-weighted value calculated from a homogeneous data set

Honorton and Ferrari also found that precognition forced-choice experiments, although demonstrating a weak effect, produced consistent ("robust") and highly significant results across a time span of more than 50 years. Unfortunately, Honorton and Ferrari (1989) felt it was "bad news" that the effect sizes had not increased over this time since the stable *ES* meant investigators had not "develop[ed] sufficient understanding of the conditions underlying the occurrence (or detection) of these effects to reliably increase their magnitude" (p. 295). However, Honorton and Ferrari's meta-analysis revealed that the largest effect sizes were found in experiments using (a) experienced participants, (b) independent testing (one participant at a time) as opposed to group testing, and (c) trial-by-trial feedback—important factors for future researchers to consider.

A later study by Steinkamp, Milton, and Morris (1998) meta-analyzed forced-choice studies for the period 1935-1997, while at the same time comparing clairvoyance[33] with precognition in order to ascertain statistical evidence of a phenomenological difference between the two. They hypothesized that clairvoyance studies would have a significantly higher effect size because precognition had an extra "calculational step," involving "real-time ESP" (clairvoyance) and then extrapolation from that information "to make an informed prediction about future events" (p. 193).

Steinkamp et al. (1998) used a total of 22 comparable study-pairs in their meta-analysis, where procedures were effectively the same in both types of studies. Effect sizes for precognition and clairvoyance were almost identical (see Table 3.2). Being such a small sample (N = 22 study-pairs) N-weighted effect sizes were calculated, again with essentially no difference in outcome (precognition: 0.034; clairvoyance: 0.030). Steinkamp et al. felt that their coding method may have been responsible for this nonsignificant result, and that a different method for coding study comparability may yield different results. They concluded that the burden of proof rested with those "who argue for a difference between effect sizes under real-time and future ESP" (p. 209).

3.1.3 Free-Response[34]

The free-response experiment tests participants in a "normal, waking state of consciousness," thus differentiating it from the ganzfeld experiment (Milton, 1998, p. 31). Milton (1998) meta-analyzed all available free-response studies published during the period 1964-1993 (these studies included remote viewing studies, in which the percipient "attempts to describe the surroundings of a geographically distant agent," Thalbourne, 1982, p. 67).

Milton found a mean effect size for 78 studies of .16 (Z = 5.72, p = 5.33 x 10^{-9}) (see Table 3.2 for other results). A file-drawer of 866 studies would be necessary to reduce this significant result to a chance outcome. An homogenized database of 75 studies had an *ES* of .17 (Z = 5.85, p = 2.46 x 10^{-9}). She also wondered if the large effects sizes in the ganzfeld domain were due to the ganzfeld technique or to the free-response methodology itself as used in ganzfeld research. When she compared mean effect sizes between her database and Honorton's (1985a) database, there was no significant difference, $t(104)$ = 1.49, $p > .05$. This result suggests that there is no basis as yet for assuming that the

[33] Clairvoyance is "paranormal acquisition of information concerning an object or contemporary physical event" (Thalbourne, 1982, p. 11).

[34] See p. 38, n31, for a definition of the free-response test.

ganzfeld condition provided a more superior method of eliciting psi than the standard free-response protocol.

3.1.4 Dice-Throwing

The dice-throwing experiment is one of a number of experiments designed to test whether consciousness can influence physical systems at the 'macro' ('greater than molecular') level. Radin and Ferrari (1991) examined dice-throwing studies spanning more than 50 years (1935 to 1987). There were 148 experimental studies and 31 control studies considered. A total of 2,500 participants attempted to influence over 2.5 million dice throws.

Forty-four percent of the 148 experimental studies gave results significantly above the 5% level. The weighted mean *ES* for the experimental studies was 0.012, which was "19 standard errors from chance" (Radin & Ferrari, 1991, p. 79). The control studies' weighted mean, however, was a low 0.00093, which was within one standard error from chance (p. 79). The combined Stouffer *Z* for the experimental studies was 18.20, but the control studies gave a low 0.18. The fail-safe *N* was 17,974 (121 nonsignificant studies to every one significant study). (See Table 3.2 for other results.)

Given that die faces are rarely equal in mass due to scooping out of the die face to mark the numbers, biases would have existed in many of the 148 studies. Radin and Ferrari (1991, pp. 74-76) took into consideration the fact that only 69 studies used protocols where targets were evenly balanced among all six die faces. A conservative quality-weighted *ES* of 0.007 was calculated ($Z = 7.62$, $p = 1.30 \times 10^{-14}$). Eliminating the outlier studies that contributed to the heterogeneity of the database resulted in a database of 59 studies with an even more conservative, but still significant, quality-weighted *ES* of .003 ($Z = 3.19$, $p = 7.16 \times 10^{-4}$).

Radin and Ferrari (1991, p. 68) found no evidence that the overall effect size was due to a "few exceptional investigators." Of note was their finding that methodological quality improved over time, but they also found, in a first analysis, that quality correlated negatively and significantly with effect size, suggesting that design flaws present in low quality studies were contributing to the success of earlier experiments. However, analysis of a homogeneous subset of the original database (from which outliers were removed) found no suggestive evidence for a possible 'regression to the mean effect' in the "perfect" dice-experiment. The general conclusion, based on the "homogeneous subset of balanced protocol studies," was that, if not strong, the mean effect size for the dice-throwing

experiments was still significant and consistent over time, indicating a "genuine mental . . . intention effect on dice" (Radin & Ferrari, 1991, pp. 79-80).

It is important to note that, after RNG experiments (reviewed next), dice-throwing experiments produce the weakest effect sizes on average of all the domains. This fact must be seriously considered by investigators wishing to conduct research in the areas of micro-PK and macro-PK since the probability of a 'pseudo-failure to replicate' (i.e., Type II error) would be high.

3.1.5 Micro-PK (Random Number Generators)

Paranormal influence on physical systems at the 'micro' (subatomic, atomic, molecular) level can be tested experimentally using random number generators (RNGs), which are also referred to as random event generators (REGs). These machines are similar to electronic coin-flippers. RNG experiments are designed to test the hypothesis that "the statistical output of an electronic RNG is correlated with observer intention in accordance with prespecified instructions" (Radin & Nelson, 1989, p. 1502).

Radin and Nelson (1989) meta-analyzed 832 RNG studies, comprising 597 experimental and 235 control studies. There was an initial analysis of 597 experimental studies. The mean effect size for the experimental studies, although very small (*ES* = 3.18 x 10^{-4}), was significantly higher than the negative (indicating psi-missing) mean *ES* for control studies (*ES* = −1.00 x 10^{-5}; Note that 54 experimental studies and 33 control studies were nonsignificant, and 6 experimental studies and 2 control studies were eliminated because effect sizes could not be estimated). The fail-safe *N* was 54,000 (approximately 90 unreported, nonsignificant studies to every one successful study reported in the literature). Radin and Nelson (1989) then trimmed the database down to 490 studies to produce a more homogeneous distribution of effect sizes, but the mean effect size was still significant (*ES* = 3.00 x 10^{-4}; see Table 3.2). They concluded that "under certain circumstances, consciousness interacts with random physical systems" (p. 1512).

Results supporting Radin and Nelson's (1989) meta-analysis were cited in Braud (1994, p. 82, Table 1), who lists four new RNG studies (study period: 1988-1993) undertaken since Radin and Nelson's meta-analysis. Taken together, these four studies

have an overall significant Z of 2.86, p = .002 (chief author being Helmut Schmidt in each case).[35]

Although replicability and robustness were demonstrated in this meta-analysis, statistical power must be considered in RNG experiments since their effect sizes are typically very low. In fact, RNG experiments to date represent the domain that delivers the lowest mean effect size of all domains meta-analyzed (see Table 3.2). Broughton (1991, p. 290) stresses the point, however, that these results are still dramatic evidence that consciousness could have an effect on micro-systems. The extreme odds against the RNG effect sizes being the result of chance represent the most convincing evidence so far, after dice-throwing, for an anomalous effect of the psychokinetic kind.

3.1.6 The Sheep-Goat Effect

Schmeidler (1945, 1960) introduced the term 'sheep' to describe a person who believes in the possibility of ESP (sheep) under the given experimental conditions, and 'goat' as one who rejects this possibility. This definition has since been broadened in numerous studies. Lawrence (1993) looked at the sheep-goat effect (SGE), which refers to the consistent finding that sheep score significantly better than goats on paranormal tasks (see also the reviews by Palmer, 1971, 1977). From the 73 forced-choice studies analyzed (4500 participants, 685,000 guesses) Lawrence calculated an *ES* of 0.029, with a highly significant Stouffer Z of 8.17, $p = 1.33 \times 10^{-16}$ (18 studies [24%] showed a significant SGE at p = .05). The mean z was 0.956, and mean *ES* per investigator was 0.026. He also found that study quality and effect size had not changed in 46 years. The file-drawer estimate was 1726 (23 unreported, nonsignificant studies for every one successful study).

Lawrence (1993, p. 82) determined that of the 65 studies that reported the order of administration of the belief scale and ESP task there were 15 studies where the sheep-goat scale was given after the ESP task (*ES* = 0.044), and 50 studies where the scale was given before the ESP task (*ES* = 0.028). He thus found suggestive evidence of an artifact (an order effect), though the difference between effect sizes was not significant. Although the number of studies stating the kind of feedback (i.e., 'no feedback' and 'trial-by-trial feedback') given was "insufficient" (p. 81), Lawrence nevertheless advised that measures

[35] Particular attention is drawn to these four studies because each experiment was supervised by "independent observers" to avoid accusations of fraud, error, and experimental negligence. This type of improvement in methodology speaks volumes about the current direction of parapsychological research in the RNG domain, particularly when significant psi results can still be produced under the critical eye of possibly skeptical supervision. This same 'witness' protocol will be employed in the '*I Ching* Experiment' (see *5.6.4*).

of belief be given before the ESP task, since outcome feedback from the ESP task could have a biasing effect on belief responses.[36]

3.1.7 Religiosity, Belief in Afterlife, Psychic Phenomena and ESP

Haraldsson (1993) conducted a meta-analysis of eight forced-choice ESP experiments to determine correlates, if any, of three belief variables with ESP performance. The Icelandic Sheep-Goat Scale (Thalbourne & Haraldsson, 1980) was administered in six of the experiments. Seven of the experiments were tests of clairvoyance and precognition, and one was a test of precognition only.

The three belief variables were religiosity, belief in an afterlife (both of which correlated significantly with ESP performance), and belief in psychic phenomena (which failed to correlate significantly with ESP performance). The z value was highest for religiosity, second highest for belief in an afterlife, and lowest for belief in psychic phenomena (see Table 3.3).

Table 3.3
Possible Predictors of ESP Performance (Haraldsson, 1993)

Predictor variable	*z*	*ES*	*p*
Religiosity (*n* = 383).	2.48	0.13	.007
Belief in an Afterlife (*n* = 382).	1.92	0.10	.028
Belief in Psychic Phenomena (*n* = 448).	0.73	0.04	.232

Note: *p* values are one-tailed

Finally, a regression analysis revealed that religiosity was the only significant predictor of ESP performance. Haraldsson (1993) concluded: "religiosity might be a more efficient predictor of ESP performance than the traditional sheep-goat variable" (p. 270, since sheep-goat scores did not correlate significantly with ESP performance.

However, Haraldsson added that inferences from these findings may only be applicable to Iceland's population. Religiosity has "differences in content" in Iceland compared to the USA, "which may be the reason for the less consistent relationship in the US between religiosity and belief and reporting of psychic experiences than in Iceland" (Haraldsson, 1993, p. 270).

[36] Lawrence's (1993) finding has been applied to all the experiments in this thesis that involve measures of belief.

3.1.8 *Hypnosis/Comparison Condition ESP*

Hypnosis has long been associated with paranormal events, and Dingwall's (1967; cited in Stanford & Stein, 1994, p. 235) review of the literature showed an association between hypnosis and paranormal events, including ESP performance. Twenty-five studies which tested hypnosis and comparative conditions (controls) for their effects on ESP performance, were meta-analyzed by Stanford and Stein (1994). An unweighted cumulative *Z* score of 8.77 (highly significant) was found for the 25 studies deemed suitable for analysis, whereas for comparison conditions, $Z = 0.34$ (not significant). Effect size was rather small for hypnosis ($\pi = 0.52$), but was essentially at chance for comparative conditions ($\pi = 0.51$).

While the hypnotic state appears conducive to psi performance (judging from the cumulative *Z* values) further statistical analysis showed a tendency only toward psi-hitting among both hypnotic participants and the comparative condition participants when consideration was given to the chief investigator. Some investigators were better than others at inducing an effective hypnotic state. Nevertheless, Stanford and Stein (1994, pp. 260-261) reached the conclusion that hypnosis generally speaking may still enhance psi performance, as long as the expectations of the investigator, and the skill and personal attributes necessary in the participants and the investigator, are present or can be implemented in the experimental situation.

3.1.9 *Perceptual Defensiveness and ESP*

The Defense Mechanism Test (DMT; Kragh, 1969) is a projective test of subliminal perception and preconscious processing designed to predict performance during stressful situations. A subliminal threat is said to cause anxiety that leads to defensive responses in perception, such as are measured by the DMT. The DMT is thus based on the Freudian concept of the 'defense mechanism': "When under pressure from extreme anxiety, the ego is sometimes forced to resort to mental strategies, called mechanisms of defence, to relieve this pressure" (Summers, Borland, & Walker, 1989, p. 701). Conscious and unconscious cognitive processes govern how perceptions are formed, and these vary from individual to individual. The DMT gives an indication of how the perceptions take place sequentially and, as a consequence, the DMT is used in the Swedish armed forces for personnel selection, since levels of defensiveness can be determined from the number of perceptual distortions measured in the test.

Ten experiments (1977-1991) were meta-analyzed by Haraldsson and Houtkooper (1995). In all studies the DMT was administered in one session and the ESP test in another.

Generally, the ESP task consisted of 40 clairvoyance trials, followed by 40 precognition trials (except one experiment where the order was reversed). These ten experiments were planned in advance to test the replicability of the DMT-ESP relationship. For all ten experiments the DMT-ESP correlation was confirmed, $Z = 2.60$, $p = 4.66 \times 10^{-3}$, one-tailed ($ES = 0.12$).

There was no decline effect in the correlation over the series of ten experiments. However, a "meta-analysis demolition effect" was found, said to result from the experimenter's focussed attention on the combined result of a meta-analysis—a form of negative experimenter effect (Haraldsson & Houtkooper, 1995, p. 267). Haraldsson and Houtkooper explored the possibility of a "DMT-correlated scoring pattern" within the psi task (e.g., position effect) but no pattern was found. They concluded that overall Z score appeared to be the best indicator of ESP performance to date.

3.1.10 Extraversion-ESP

A number of studies done in the past give firm evidence that psi hitting and extraversion are positively correlated (Eysenck, 1967; Palmer, 1977; Sargent, 1981). Honorton, Ferrari, and Bem (1998) conducted a meta-analysis using a total of 60 independent studies involving 2,963 participants and 17 independent investigators. Rather than look at psi performance *per se*, Honorton et al. focused on the relationship between psi-performance and extraversion as a predictor variable. Thus the effect size index was a correlation coefficient calculated on these two variables.

Overall, the mean weighted r for the 60 studies indicated an extremely weak relationship, but nonetheless a significant one, $r = 0.09$, $Z = 4.63$, $p = 4.00 \times 10^{-5}$. The extraversion-ESP correlation was even smaller for the 45 forced-choice studies, $r = 0.06$, $Z = 2.86$, $p = 4.20 \times 10^{-3}$, but for the 14 free-response studies the correlation was higher, $r = 0.20$, $Z = 4.32$, $p = 1.5 \times 10^{-6}$, though conventionally signifying a weak relationship. (For all three extraversion-ESP meta-analyses χ^2 tests were significant at $p < .05$.)

One problem raised in Honorton et al.'s paper was the fact that the extraversion measure was administered after the ESP task in some of the forced-choice studies, meaning that there was a conceivable inflation in extraversion scores for those participants successful in the ESP task:

(i) ESP task before extraversion assessment, $r = 0.17$—a weak extraversion-ESP relationship.

(ii) ESP task after extraversion assessment, $r = -0.02$—no extraversion-ESP relationship.

A significant difference between these two correlations was found.

Since extraversion scores were also available for 221 participants who had participated in auto-ganzfeld studies, Honorton et al. were able to calculate the correlation between extraversion scores and ganzfeld direct-hit scores, $r = 0.18$, $p = .008$. This result was not significantly different from the estimate in their free-response/extraversion test study ($r = .20$), and was also confirmatory of the fact that extraverted participants in free-response experiments (including the ganzfeld) appear to produce higher ESP scores than the extraverted participants in forced-choice studies.

It seemed that Honorton et al.'s findings contradicted those of Eysenck (1967), Palmer (1977), and Sargent (1981), that there is a relationship between extraversion and forced-choice psi-performance. Results based on data generated from psi tasks conducted before the personality test appeared not to be reliable. However, Palmer and Carpenter (1998) criticized this conclusion and argued that personality tests are designed not to be susceptible to situational bias (their high test-retest correlation suggests their resistance to mood changes, which may be caused by psi feedback).

Also, Palmer and Carpenter pointed out that an order effect could not be considered an un-confounded variable, but was an artifact of the group-testing status ("test setting"), since order was found to be conflated by individual and group testing in the forced-choice data. In fact nine of the 18 forced-choice studies with the extraversion test administered after the ESP test used individual testing, whereas only eight of 19 studies with the extraversion test administered before the ESP test used individual testing. A confound between test-setting and order of testing was found.

When correlation coefficients were calculated for the 19 individual-testing 'before' studies, $r = 0.21$, and the group-testing studies, $r = -0.05$, a significant difference was found, $t(17) = 2.17$, $p = .045$. Regression analysis revealed that test-setting made a significant contribution as a predictor to the transformed extraversion-ESP correlations, but order did not. Palmer and Carpenter (1998) claimed that group-testing status was the "true mediator" (p. 280) of the extraversion-ESP correlation, and not the order of testing. Also, elimination of the group testing studies showed extraversion-ESP relationships of comparable magnitude in the forced-choice and free-response ESP tests.

Palmer and Carpenter (1998) considered it possible that other "unrecognized confounding variables" may have contributed to the problem. Thalbourne and Storm

(1999), for example, note that the types of extraversion tests used in the forced-choice studies should have been reported, as was the case for the free-response studies (Honorton et al., 1998, p. 267). Thalbourne and Jungkuntz (1983) reported two significant negative extraversion-ESP correlations when they used the Minnesota Multiphasic Personality Inventory (MMPI) Social Introversion Scale. If Honorton et al. are proved incorrect in their negative assumption about forced-choice experiments then the forced-choice paradigm might still serve in ESP-extraversion research.

3.2 The Review by Jessica Utts (1991)

Utts (1991) has already given a comprehensive review of most of the foregoing meta-analytic studies (pre-1991). She also stressed the importance of replicability, power analysis, and effect size. In discussing the general state of parapsychology she concluded that:

(i) "The recent focus on meta-analysis in parapsychology has revealed that there are small but consistently [significant] nonzero effects across studies.

(ii) "It may be that the [significant] nonzero effects observed in the meta-analyses can be explained by something other than ESP, such as shortcomings in our understanding of randomness and independence.

(iii) "A promising direction for future process-oriented research is to examine the causes of individual differences in psychic functioning.

(iv) "If ESP does not exist, there is little to be lost by erring in the direction of further research, which may in fact uncover other anomalies.

(v) "If ESP does exist, there is much to be lost by not doing process-oriented research, and much to be gained by discovering how to enhance and apply these abilities to important world problems." (p. 377)

All these issues are familiar to parapsychologists, and some implications of those issues, to varying degrees, have been addressed above. Skeptics will continue to criticize the validity of the significant nonzero effect-sizes, while proponents may base their conviction that psi is real, based exclusively on the statistical evidence, although the very low effect sizes may raise serious doubts from both skeptics and believers about the ultimate utility of psi. As it

happens, the problem of low effect sizes is not unique to parapsychology. This problem is quite common in other fields (see the examples from the field of medicine in Greenhouse & Greenhouse, 1988, and Rosenthal, 1990).

Studies in extraversion and other personality variables address (iii) above ("a promising direction for psi research . . ."), but it is difficult to imagine how we might entertain the idea that the 'wild goose-chase' hypothesis implied in (iv) above ("ESP does not exist . . .") would win a legion of motivated investigators. The safest bet for practicing parapsychologists is to maintain (v) as a goal and a certainty ("ESP does exists . . ."). This way the futurity of parapsychology is assured by (1) the conviction of its investigators, (2) the promise of a plausible paradigm that appeals to many, if not all, in its 'explanation' of paranormal phenomena, and (3) the belief that psi has some practical value.

Some of the comments made for and against the inferences of Utts's article are worthy of note. Diaconis (1991, p. 386) feels that parapsychology is "worth serious study," but efforts at producing valid and reliable results that allegedly indicate psi can be taken seriously if and only if "magicians" and/or "knowledgeable" psychologists "skilled at running experiments with human participants" are used. This unwarranted claim flies in the face of the rigorous methodologies that form the basis of studies in parapsychology. Experimenters in psi are often "knowledgeable" psychologists, many of whom consult magicians, or use them to scrutinize experimental proceedings (Thalbourne, 1995a, pp. 362-363; Utts, 1991, pp. 396-397).

Greenhouse (1991, p. 386) praises Utts's presentation, but does not believe that "anyone's views regarding the existence of paranormal phenomena" will change. He makes it clear that only until certain prejudices are dissolved can new paradigms be adopted. Greenhouse (1991, p. 367) cites the aspirin/heart-attack study used by Utts (1991, p. 367) to clarify this point. In spite of the very low effect size reported in this study, the relationship was easily accepted because the biological mechanism that explained this relationship is a causal process that is easily understood, even by the layperson. The point to be noted here is that parapsychology must focus on process studies that aim to explain the mechanism of psi, rather than just demonstrate its presence.

Hyman (1991, p. 392) yields to Utts's conclusion that there is an "anomaly in the parapsychological findings." His only reservation in this regard is that the effect sizes vary so much, both within experimental domains ("the effect sizes obtained by [Robert G.] Jahn are much smaller than those obtained by [Helmut] Schmidt with similar experiments on random number generators," Hyman, 1991, p. 392), and between domains (for example, meta-analysis shows that the RNG domain yields the lowest mean effect size, while the

biological systems domain yields the highest). Hyman asks what it is we are supposed to account for when there are "nonchance departures from a statistical model" (p. 392) concomitant with such varied effect sizes. He uses an anecdotal and misleading astronomical example to illustrate his point (Uranus's anomalous orbit is explained as being due to another planet, which led to the discovery of Neptune), which implies that non-parapsychological science can always give account of its anomalies. However, unaccounted anomalies abound in orthodox science, where in many cases, it is "empirical departures from expectation [that have] led to important findings or theoretical models" (Utts, 1991, p. 400). Parapsychologists have merely followed the example of science. Utts's (1991) response to Hyman is to propose "intensive investigation" as the means by which we may arrive at explanations for psi effects (p. 400).

Finally, Morris (1991) notes that meta-analysis allows the parapsychologist the opportunity of identifying the moderator variables that need to be incorporated into experiments in order to yield the strongest effects. For example, it appears to be better to test individuals rather than groups, better to use dynamic rather than static targets, and better to use experienced participants rather than novices. Currently, though, there are doubts about the last two of these three recommendations (Milton & Wiseman, 1999).

3.3 Conclusion

This chapter continued the main theme of Chapter 2—a review of the meta-analytic literature in parapsychology. Though meta-analysis is a generally accepted methodology, it does have its critics. Those critics of meta-analysis, and statistics in general, will assure us that statistical procedures can "mislead and be misused" (Diaconis, 1991, p. 386). There seems little doubt that such an assurance is based on fact, but to avoid such pitfalls a basic tenet of science has always been to proceed with caution.

Nevertheless, the evidence to date from the meta-analyses—that there are significant effect sizes for the nine major domains featured (see Table 3.2)—suggests a very real (albeit statistical) anomaly worthy of continued investigation. Thus the field of parapsychology justifies itself as a discipline, and in the spirit of Rosenthal's (1986, p. 333) words—"at any point in time some judgment can be made"—it is hoped that the chapters to follow, including the findings from those experiments described therein, will foster sound judgements that might help strengthen the integrity of the field by reaching a deeper and more critical understanding of the nature of psi. The theory of psychopraxia

will be the means by which this task will be operationalised. The next chapter (Chapter 4) introduces this theory.

CHAPTER 4
THE THEORY OF PSYCHOPRAXIA[37]

Psychology essentially cannot be exhausted by causal methods only, because the mind [= psyche] lives by aims as well.

C. G. Jung (1960, para. 843, n.38)

What can a cause be if no specification can be written for it?

Edwin G. Boring (1966, p. *xxi*)

4.1 The Traditional Perspective on the Paranormal

In the previous two chapters nine major paranormal domains were reviewed: Biological systems, forced-choice, free-response (including remote viewing), dice-throwing, micro-PK (RNG), clairvoyance, precognition, ganzfeld and autoganzfeld. These and other domains are used to test various forms of ESP and PK, both of which can be represented in tabular form (see Tables 4.1 & 4.2, adapted from Thalbourne, 1981). Table 4.1 shows that ESP target-events are traditionally classified as either mental in nature, or physical in nature.

Table 4.1
The Traditional 3 x 2 Taxonomy of ESP according to Temporal Location and Nature of the Presumed Target-Event (SOURCE: Thalbourne, 1981, p. 35)

Nature of Presumed Target-Event	Temporal Location of Presumed Target-Event		
	Past	Present	Future
Mental	Retrocognitive Telepathy	Contemporaneous Telepathy	Precognitive Telepathy
Physical	Retrocognitive Clairvoyance	Contemporaneous Clairvoyance	Precognitive Clairvoyance

[37] Parts of this chapter were published in a refereed article (see Appendix AD).

Table 4.2 shows that PK target-events are restricted to physical events.

Table 4.2
The Traditional Taxonomy of PK according to Temporal Location

Nature of Presumed Target-Event	Temporal Location of Presumed Target-Event		
	Past	Present	Future
Physical	Retroactive Psychokinesis	Contemporaneous Psychokinesis	Preactive Psychokinesis

J. B. Rhine (1934) coined the terms extra-sensory perception (ESP) and psychokinesis (PK) as ostensibly useful categories for describing two seemingly different paranormal phenomena. Even though, by the 1940s, Rhine (1948/1954) regarded ESP and PK as "one single fundamental two-way process" (p. 112), both terms are still used as two basic categories of paranormal phenomena. Parapsychologists have grown accustomed to the terms ESP and PK, and generally believe that there is a clear distinction between the two phenomena. This distinction or 'dichotomy' is discussed next.

4.2 The ESP-PK Dichotomy

4.2.1 ESP and PK as a Continuum

A closer look at certain types of paranormal experiments reveals that an investigator can sometimes be uncertain about the paranormal category under investigation. In other words, one can find oneself asking if the experiment is testing for ESP or PK, or both. If it is both, then the distinction is not so clear. This degree of latitude, at the very least, undermines the ESP-PK dichotomy, and researchers are aware of this problem. Heath (2000), for example, claims that "the outward differences [between ESP and PK] may have blinded us to their similar core essence, and that ESP and PK may be the same process" (p. 70).

On the basis that there might be some degree of overlap between ESP and PK it becomes possible to propose an ESP-PK continuum, as Schmeidler (1987, p. 38) did. She suggested that at one extreme there may be pure PK (no measurable ESP). Movement along the continuum would see an increase in ESP fused with PK, until only ESP with no measurable PK would be evinced at the other end of the continuum. Schmeidler (1994)

later claimed that supporting evidence for a difference between ESP and PK would come from experimentation that might reveal a distinction in scoring between the two, but if "procedural conditions for success are found to be similar in PK and ESP, and if their physiological and psychological correlates are also found to be similar, then the two are functionally indistinguishable" (p. 199).

One problem with testing Schmeidler's proposal would be in determining just how the procedural conditions of the ESP task and the PK task are to be considered the same when the tasks would have to be procedurally different from each other just by design. A further problem, as Schmeidler (1994, pp. 198-199) recognized, is whether the tasks can be exclusively described as ESP or PK in the first place, and she admits there are "gray areas" in this regard. (See Knowles, 1967, 1968; Osis, 1953; Schmidt, 1969, 1973; and Storm & Thalbourne, 1998-1999, for examples of these gray areas, where it is possible to interpret the paranormal effects as either ESP, or PK, or both in combination.) In fact, some theorists suggest that ESP and PK are so similar to each other that they might very well be the same thing, or that one may be a form of the other (S. E. Braude, 1979; Nash, 1983; Rhine, 1974; Rhine, cited in Rao, 1983; Schmidt & Pantas, 1972; Stanford, 1977a, 1977b).

Although it seems that the psychophysiological state-trait of the participant has an influence on paranormal performance, there is no consensus yet as to whether ESP and PK function best under mutually exclusive conditions. For example, some studies seem to suggest that PK works best when the participant is tense or aroused (Braud, 1985; Steilberg, 1975), whereas other studies report a negative correlation between PK performance and anxiety (which is related to tension and arousal;[38] Broughton & Perlstrom, 1986, 1992). Still other studies suggest that ESP can function similarly under conditions of arousal or lack of arousal (see Stanford, 1977a; Van der Sijde & Snel, 1992). Therefore, Schmeidler's (1987) schema and similar schemata only raise the question: "What grounds are there for dichotomizing what may be a single paranormal process?" For some theorists, it appears there are no grounds, and a few of them have attempted a more complete unification of ESP and PK than has been offered to date. Their theories are now reviewed.

[38] Clark, Beck, and Beck (1994) found, through discriminant-function analysis, that anxiety was discriminated by subjective anxiety and tension as measured on standard anxiety disorder scales. Hoehn, Rudolf, Pourmotabbed, and McLeod (1997) found that increased muscle tension and heightened arousal are consistent symptoms among anxiety-prone individuals.

4.2.2 The Thouless-Wiesner Hypothesis

The first major consideration of the ESP-PK dichotomy predates Schmeidler's attempts at unification and comes from Thouless and Wiesner (1947). They proposed a unitary process underlying ESP and PK, but paradoxically, they went on to use the Greek symbols Ψ_γ (psi-gamma) and Ψ_κ (psi-kappa) as labels for ESP and PK, respectively, in order to indicate that ESP and PK were "different aspects of one process" (Thouless & Wiesner, 1947, p. 179). Rhine accepted the term 'psi', and it has been used ever since in parapsychology to designate "paranormal causation" or "paranormal process" (Thalbourne, 1982, p.56).

Thouless and Wiesner (1947), therefore, never quite escaped the dualistic argument that ESP and PK were qualitatively different phenomena. Nevertheless, they then proposed that "normal thinking and perceiving," which are endosomatic processes (i.e., they take place within the body) are established by the same means as an ESP event (e.g., clairvoyance), which has an exosomatic target (a target outside the body). They also proposed that motor activity of the body is conducted by the same means as psychokinetic activity. So under the Thouless and Wiesner schema there are fundamentally two phenomena: Ψ_γ and Ψ_κ, manifesting endosomatically and exosomatically, thus re-formulated as a bi-modal 'diasomatic'[39] process (see Figure 4.1).

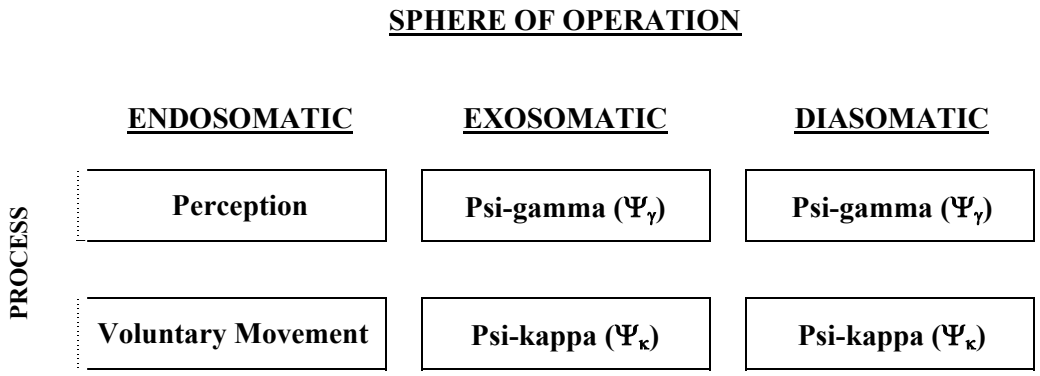

Figure 4.1. The Thouless-Wiesner reduction of four phenomena to two. (SOURCE: Thalbourne, 2000a, p. 52)

[39] "From the Greek *dia* 'through' (as in 'diaphanous') or 'throughout', + *soma*, meaning 'body'. The term is intended to mean something like 'capable of occurring either endosomatically or exosomatically, the body being like a 'permeable membrane'." Thalbourne, 2000a, p. 35).

Any kind of mental event (thinking a thought, or a telepathic mind-to-mind communication) is Ψ_{γ}, and any kind of motor action (moving an arm, or paranormally influencing the fall of coins) is Ψ_{κ}. Again, note that Thouless and Wiesner maintained the ESP-PK distinction, even though they seemingly opted for a unitary psi process. It is noted too that psi has invariably come to be associated more with paranormal phenomena and less, if at all, with normal phenomena (Beloff, 1985, p. 226). Few parapsychologists today seem to be familiar with the Thouless and Wiesner theory.

4.2.3 Jung's Synchronicity Theory

In the 1930s and 40s, the Swiss psychiatrist C.G. Jung became familiar with Rhine's work, corresponded with him, and knew of the terms ESP, PK and psi. In the early 1950s, Jung (1960) published two major works on the paranormal, and viewed ESP and PK as essentially one process, which he called 'synchronicity'. Two or more events constitute synchronicity when a meaningful connection—a meaningful association—can be made between the events, but it is only synchronicity when meaningfulness is the connecting principle between the events, with no mechanically causal connection between the two (Jung, 1960, para. 849-850).

Jung (1960, para. 965) saw synchronicity as being contingent upon an archetype (archetypes are nodal points or structural components of the unconscious that govern or influence our modes of perception and patterns of behaviour). The archetype forms the substructure of the synchronicity, connecting at least two events (an exo-psychic one, and an endo-psychic one) with a common theme, and acting as a defining quality throughout the experience, thereby intensifying the meaningfulness (Jung, 1960, para. 912).

Mansfield, Rhine-Feather and Hall (1998) argued that synchronicity is characteristically different from psi (i.e., ESP and PK) because psi is 'scientifically' causal (for example, it has been repeatedly demonstrated in the laboratory), whereas synchronistic phenomena are too "sporadic and unpredictable" (p. 20) to be considered scientifically causal (i.e., experimentally, they escape detection). Storm (1999, pp. 262-264), however, argued that experimental synchronicity is possible, and gave examples.

Not only may synchronicity and psi have much in common (Storm, 1999, pp. 264-266), but Jung's theory also draws attention to psychological variables that have heretofore been ignored in parapsychology, particularly the meaningfulness in paranormal events, whereby emotional and/or cognitive states of mind can be represented in the physical world as meaningfully similar events (either ESP or PK). Thus, although Jung used Rhine's categories, he helped move theoretical parapsychology forward somewhat by proposing a

fundamental similarity between ESP and PK that was overlooked by Thouless and Wiesner. However, the only way of determining that ESP and PK might be a unitary, synchronistic process, as Jung's theory suggests, is to show how they are consistently contingent upon coincidence and meaningfulness. Thus far, evidence is limited, but certainly the experimental approach to synchronicity (e.g., see Braud, 1983) does not require that ESP and PK terminology be evoked.

4.2.4 Stanford's PMIR and 'Conformance Behavior' Models

Decades after Jung, Stanford (1974a, 1974b) came close to merging ESP and PK together with his psi-mediated instrumental response (PMIR) model, although at that stage he did not see ESP and PK as the same process. In describing the PMIR model, Stanford (1974a) proposed that:

> The organism nonintentionally uses psi to scan its environment for need-relevant objects or events or for information crucially related to such events, and when obtained, the organism tends to act in ways which are instrumental in satisfying its needs in relation to the particular object or event in question. (p. 35)

(Note that the "scanning component" was later discarded from the PMIR model—see Stanford, 1990, pp. 57-59.) Stanford (1974b) extended the above definition to include PK phenomena. He then developed his *conformance behavior* model, which "subsumes both ESP and PK" events under one model (Stanford, 1978, p. 198). Any paranormal process was thereby typified as a "conformance" (i.e., a disposition-serving outcome) respondent to the "need or other disposition" of an individual (p. 207). Conformance behavior theory regards both ESP and PK as "goal-oriented rather than information-based" (pp. 203, 208).

At a later date, Stanford (1990) was "less than fully convinced about the goal-oriented character of psi processes" (pp. 58–59), but maintained it for the most part in his resurrected PMIR model. Stanford (1990) also abandoned the conformance behavior model, but even though the PMIR model now maintains that the same underlying elements are necessary for ESP and PK effects (e.g., a "disposed" system) he still discusses them as discernibly different processes (cf. Stanford, 1990, on ESP, with Stanford, 1974b, on PK).[40] For example, Stanford (1974b, pp. 326–328) holds that "implicit extrasensory guidance" is "part and parcel" of many PK events and that "active-agent telepathy" may really be a special form of PK (p. 344). He thus appears to identify qualitative differences

[40] Stanford (1990, p. 57) regards Stanford (1974b) as the definitive reference for PMIR in terms of psychokinetic effects (see especially Stanford, 1974b, pp. 348–351).

between ESP and PK, thereby failing to merge acceptably the two into a truly unitary process.

4.2.5 Decision Augmentation Theory

May, Utts, and Spottiswoode (1995a) later proposed decision augmentation theory (DAT) as a "phenomenological, . . . logical and formal extension of Stanford's elegant PMIR model" (p. 456). Guided by the PMIR model, they proposed that "humans integrate information obtained by anomalous cognition [e.g., precognition] into the usual decision process" (May, Utts, & Spottiswoode, 1995b, p. 195). The RNG participant, for example, using anomalous cognition, opportunistically scans the environment for information in order to determine the appropriate moment to "press the button" (p. 195). The scanning component even extends to anomalously choosing the most opportune time and day to attend a test session, or to anomalously choosing a preferred experimenter, and so on (May, Utts, & Spottiswoode, 1995b, p. 200).

May et al. (1995a) effectively reduced a conventional PK[41] effect, in the RNG case, to an ESP effect (viz., anomalous cognition). However, DAT does not fully achieve—or, in fact, attempt—a unification of paranormal effects (e.g., May et al., 1995b, p. 200, acknowledged that DAT does not explain levitation). Furthemore, Dobyns and Nelson (1997) found empirical evidence that the process in RNG experiments involved "an actual change in the operation of the experimental devices, taking place at the level of elementary binary decisions . . . [which] are inconsistent with the process of biased sampling required by the DAT model" (p. 256). DAT, then, while showing a bias toward an ESP-like causality behind most paranormal effects, is not yet able to eliminate successfully the need for the PK category (see also Braud, 1990; Braud & Schlitz, 1989; Stevens, 1998-1999).

4.3 The Three Metaphysical Ontologies

The major theories just reviewed have dealt with the ESP-PK dichotomy either directly or indirectly. In some respects, most theorists have regarded the dichotomy as a real problem in need of a solution, with the exception of Jung, who never criticized the dichotomy, yet ironically, identified a common phenomenology between ESP and PK. According to Thalbourne (2000a, pp. 15-25), theorists (such as those identified above)

[41] There are traditional-minded researchers who would regard the RNG effect as a form of PK. The RNG effect has been referred to as *micro-PK* since 1969, although Schmidt (1969) preferred *precognition* (i.e., ESP).

have failed to identify the fact that they were making specific philosophical assumptions about the phenomena. In this next section, the ontological question, and its bearing on the ESP-PK dichotomy, will be addressed.

Scientists for the most part depend heavily on terminology. The kind of terminology they use depends on the philosophical world-view they hold. The investigator, therefore, is discourse-dependent in regard to interpreting, describing, defining, and otherwise explaining the phenomena they 'discover' or 'invent'. Likewise, a careful consideration of the ESP-PK problem reveals that ultimately any interpretation of paranormal phenomena will depend absolutely on one's metaphysical outlook, which may logically include normal phenomena in this consideration, as will be seen.

Traditionally, there are three fundamental metaphysical perspectives on the ontology of normal phenomena: Metaphysical Dualism, Metaphysical Materialism, and Metaphysical Idealism (see Table 4.3). Thalbourne (2000a, pp. 15-25) applies these same perspectives to paranormal phenomena. The terminological problem (or discursive problem), just raised, has to do with the way in which one prefers to characterize or interpret phenomena. The difficulty for science that this problem raises is readily apparent, since scientific consensus about any phenomenon also requires some implicit agreement about the philosophical structure underlying the terms to be used in describing that phenomenon. Often there is a tacit understanding amongst scientists regarding the use of the various ontologies. For example, in the case of dualism, there may be a mixed use of terms that entails a switching to and from materialistically derived, 'entity'-like words, to idealistically derived, conceptual words.

Thalbourne (2000a) argues that it is rare that the above actions, and the implications of the various ontological foundations, are brought forward for analysis. As a case in point, the ongoing debate in parapsychology over ESP and PK is essentially driven by the three metaphysical perspectives (see Table 4.3). The idealist may prefer ESP-type definitions, while the materialist may prefer PK-type definitions. Of course, the dualist has the dubious advantage of wrestling with the category they find the most parsimonious! (i.e., ESP or PK; see Edge, 1985b). This conundrum forms the basis of the theoretical ideas postulated earlier by Thouless and Wiesner (1947), and revived by Thalbourne (2000a), and these ideas are discussed next.

Table 4.3

Interpretation of Normal and Paranormal Phenomena in Terms of the Three Major Metaphysical Ontologies: Dualism, Materialism, and Idealism

Ontology	Perspective on Normal Phenomena	Perspective on Paranormal Phenomena
Dualism[a]	a) All events are considered possible so long as they do not defy the known physical laws, or known mental processes. There may be other physical laws yet to be discovered. The absolute limits of the mind are yet to be determined. b) Normal human events are classified as physical events (e.g., motor function, etc., originate in the brain) or mental events (e.g., cognition, volition, etc., originate in the mind).	a) The processes underlying paranormal events would be described in terms of ESP if mind-to-mind or mind-to-object information transfer is involved (telepathy, clairvoyance, and precognition), or PK when mind influences matter. b) ESP and PK are distinct categories.
Materialism[b]	a) All events are considered possible so long as they do not defy the known physical laws. There may be other physical laws yet to be discovered. b) Normal human events (i.e., cognitive and motor functions) are mediated by the brain.	a) The processes underlying paranormal events would be described in terms of PK, where 'ESP' is actually a brain state otherwise called a mental state (telepathy may work like a brain-to-brain field transmission; clairvoyance may work like radar or sonar). b) ESP reduces to PK.
Idealism	a) All events are considered possible so long as they can be experienced as sense-data. The absolute limits of the mind are yet to be determined. b) The source of all normal human events (i.e., cognitive and motor functions) is attributed to mind.	a) The processes underlying paranormal events would be described in terms of ESP, where 'PK' is actually a paranormal transformation of sense-data otherwise called 'the perceived physical object' (telepathy is mind-to-mind sense-data; clairvoyance works via mind-to-external sense-data or information). b) PK reduces to ESP.

Notes: [a] Assumes Interactionism
[b] Assumes Identity Theory

4.4 The Theory of Psychopraxia

4.4.1 Origin and Definition of the Theory of Psychopraxia

The theory of psychopraxia, put forward by Thalbourne (1981, 1982, 2000a), goes beyond the ESP-PK dichotomy, and Thouless and Wiesner's endosomatic/exosomatic dichotomy. But the theory was also developed to clear up another conceptual muddle—the nature of psi. Psychopraxia is a neologism based on two Greek words: *psyche*, which means 'soul' or 'mind', and *praxia*, from which we get our word 'practice' (derived from *prattein*, which means 'to accomplish' or 'bring about'; Thalbourne, 2000a, pp. 73-76). Thus psychopraxia is the psyche at work effecting a goal.

Thalbourne (2000a, pp. 15-25) proposes that the first and possibly major advantage of this neologism is that it avoids the debate over whether psi is ideational by nature, or material, or both. Thus the theory of psychopraxia is conservative in the sense that it makes no ontological judgement about the nature of psi by avoiding the use of the term altogether: Psychopraxia refers only to organisms that are in the practice of bringing about goals, whether these goals are to be defined as normal or paranormal.

The theory of psychopraxia also relativizes ESP (or more broadly, Ψ_γ) and PK (or more broadly, Ψ_κ), by identifying three commonalities. The first is that they are both diasomatic, where diasomatic means 'throughout [and beyond] the body' (see p. 60, n39). Thalbourne (2000a, p. 35) believed that there was little purpose in maintaining a distinction between the inside and outside of the organic structure.

4.4.2 The Pro Attitude

The second commonality is the "pro attitude," which covers conscious goals and "motivational states not currently in awareness." Thalbourne (2000a) states that "a person may be said to have a pro-attitude towards state **S** when they would consciously prefer **S** rather that **–S** [**not S**] if those two alternatives were to be brought to their attention" (p. 65). In fact, the theory of psychopraxia prefers the term 'pro attitude', seeing it as a refinement of the apparently ambiguous term 'goal-orientation' (as used, for example, by Stanford, 1978), since often the way in which a desirable outcome is fulfilled may be completely beyond the ken and capacity of the organism, such that no "volition or motive or need" (p. 66; i.e., goal) may be present, but a pro attitude is present.

4.4.3 *ESP and PK as 'Actions'*

The third and final commonality is that ESP and PK are both instances of action, whether cognitive or motor. "Mind and matter may, in the final analysis, be ontologically different substances (as the Dualists believe), but the more important fact is that *from the point of view of the active agent, mind and matter are manipulated in fundamentally the same way*" (Thalbourne, 2000a, p. 75). Psychopraxia is thus defined as:

> A . . . principle underlying all interactions between the self, or ego, and the realm consisting of mental and physical events, whereby under certain conditions (as yet unspecified, but probably psychophysiological) the adoption of a desire automatically results in its fulfillment in reality; *psi* is thus seen as a special instance of psychopraxia being those manifestations of goal-achievement which are *exosomatic* rather than *endosomatic*, that is, which are not mediated by the normal sensory-motor apparatus. (Thalbourne, 2000a, p. 76)

Thalbourne (2000a) later drops the term "desire," replaced it with "pro attitude" (p. 96), and referred to other "certain conditions" as being "necessary" (p. 96) in the sense that they were sufficient in bringing about a psychopractic effect. This important stipulation eliminated the "wish-fulfillment" allegation that might be made against the theory of psychopraxia ("someone *might* be tempted to describe 'psychopraxia' as simply a more formal name for 'magic'," p. 95). In the present thesis, a 'sufficient condition' shall refer to a necessary condition that was sufficient in bringing about a psi effect at the level of the sample or a group formed from that sample, but certainly not at the level of each and every individual. Were that the case, we would expect perfect correlations, and 100% hit-rates.

By noting that exosomatic and endosomatic processes have the same underlying cause, the theory of psychopraxia is thus an attempt to unify normal and paranormal phenomena, cognitive and motor (Thalbourne, 2000a, pp. 87-88).

4.4.4 *The Concept of the Necessary Condition*

By the above definition, psychopraxia only takes effect when certain necessary conditions are present during the task at hand. These conditions are deemed 'necessary' because the "pro attitude is *not*, in and of itself, sufficient to bring about a goal" (Thalbourne, 2000a, p. 96). Conditions, therefore, are necessary and sufficient if the paranormal effect cannot occur without them. There may be any number of these conditions. Listed below are some conditions hypothesized as being conducive to paranormal effects (or psychopraxia, to use Thalbourne's term), and may be shown to be necessary and sufficient:

1. Honorton et al. (1990) identified three psi-conducive 'features' (i.e., conditions):

   (i) Individual testing (as opposed to group testing).
   (ii) Dynamic targets (as opposed to static targets).
   (iii) Experienced participants (as opposed to naïve participants).

   All experiments in this thesis involve individual testing, so that (i) above has been optimized as a psychopraxia-conducive condition.

2. Certain personality factors may also be conducive to psychopraxia. For example, Extraversion scores, as measured on Eysenck's (Eysenck & Eysenck, 1965) EPQ scales, have been found to relate to ESP scores, although there are no clear trends on Eysenck's scales (see Palmer, 1977, p. 185). Also, Extraversion as measured on Cattell's (Cattell, Eber, & Tatsuoka, 1970) 16PF scales, and the component factors of Extraversion on the 16PF, have been used to predict ESP scores (e.g., Storm & Thalbourne, 1998-1999).

   The 16PF is featured in the experiment with the *I Ching*. See Chapters 5 and 6, where attempts at replicating Storm & Thalbourne's (1998-1999) findings of significant correlations between psi performance and four 16PF factors, Extraversion (Factor EX), and its components—Liveliness (Factor F), Social Boldness (Factor H), and Self-Reliance (Factor Q_2)—were undertaken.

3. More recently, belief (or disbelief) in the paranormal process, and high (or low) scores on the Transliminality Scale (Form B; Thalbourne, 1998; see **5.3**) have been correlated with psi performance (in the case of belief, see Lawrence, 1993, and in the case of transliminality, see Sanders, Thalbourne & Delin, in press, and Storm & Thalbourne, 1998-1999). Belief in paranormal processes (usually given as ESP) and transliminality may therefore be taken as conditions that are conducive to psychopraxia. Belief is tested as a psychopraxia-conducive condition in the experiments featured in Chapters 5, 7, and 9. Transliminality is also tested in Chapter 5.

4. Positive attitudes toward gambling, belief in good luck, and intuition as a personality trait (Thalbourne, 2000a, p. 117) may also be conducive to psychopraxia. Thus, the Gambling Attitude Scales (Kassinove, 1998), the Belief in Good Luck Scale (Darke & Freedman, 1997a, 1997b), and Intuition on the Singer-

Loomis Type Deployment Inventory (Singer & Loomis, 1996), were each tested as possible correlates of paranormal performance (see Chapter 8).

5. Braud (1975) compiled a list of seven psi-conducive 'states', which he saw as components of the psi-conducive "syndrome":

    (iv) physical relaxation.
    (v) reduced 'physical' arousal or activation.
    (vi) reduction in sensory input and processing.
    (vii) increased awareness of internal processes, feelings and images.
    (viii) receptive mode/right-hemispheric functioning as opposed to action mode/left-hemispheric functioning.
    (ix) an altered view of the nature of the world. (For example, a perception of timelessness in precognition tests.)
    (x) psi (or what might be accomplished through psi) must be (at least) momentarily important.

Some of the last six of these conditions can be corollaries of the first (i.e., physical relaxation), particularly (v) and (vi) above. It is hypothesized that relaxation is conducive to psychopraxia (see Chapter 9).

6. Thalbourne (2000a, p. 96) identified six 'state of mind/body' factors that are conducive to normal (i.e., psychopsychological) functioning, and given that psychopraxia is a principle hypothesized as applying to normal as well as paranormal functioning, these same 'state of mind/body' factors may be conducive to exosomatic psychopraxia:

    (xi) 'fresh' (rather than 'tired')
    (xii) 'sober' (rather than 'intoxicated' or drugged)
    (xiii) 'attentive' (rather than 'distracted')
    (xiv) 'confident' (rather than 'diffident')
    (xv) 'fit' (rather than 'out-of-condition')
    (xvi) 'young' (rather than 'advanced in age'; Thalbourne, 2000a, p. 96).

These six 'state of mind/body' factors were tested as conditions conducive to psychopraxia (see Chapter 11).

7. Finally, based on the low effect size found in the RNG meta-analysis (Radin & Nelson, 1989; see Chapter 2), it is hypothesized that perceived simplicity of the apparatus/mechanical set-up and perceived ease of the paranormal task (as opposed to perceived complexity of the apparatus/mechanical set-up and difficulty of the task) would be conducive to psychopraxia (see Chapter 11). This hypothesis was based on the assumption that the low effect size in the RNG meta-analysis came as a result of participants' being intimidated by the experimental set-up, and thus performances were worse than might ordinarily have been expected in comparison with other paranormal domains (see Chapter 10 for details).

There would be many other conditions conducive to psychopractic effects, but not all of them are necessary. For example, Vernon Neppe (V. Neppe, personal communication, 31st December, 2000) gave examples of altered states of consciousness that are hypothesized as being conducive to "subjective paranormal (psi) experiences":

(a) classical hypnosis.
(b) given amounts of prescribed or recreational chemicals—ranging from opiates to barbiturates to hallucinogens to benzodiazepines to alcohol.
(c) certain transition stages of a mental illness.
(d) dream states.
(e) temporal lobe twilight states.
(f) ganzfeld states.

Note that the ganzfeld state, (f) above, includes relaxation, which is tested in Chapter 9.

Stanford (1990, pp. 138-148) lists conditions ('factors') that may be aversive to psi (i.e., may limit or prevent PMIR), including 'rigidity', 'inhibition', and 'stereotopy'. The polar opposites of these factors (namely, 'flexibility', 'openness', and 'spontaneity'), which would not put "constraints on behavior" (p. 139), and thus would not diminish the possibility of PMIR, may also be seen as conditions that bring about psychopraxia. These factors are not specifically tested in this thesis, but the dimensions, 'openness' vs. 'inhibition', and 'spontaneity' vs. 'stereotopy', recall Factor H on the 16PF (i.e., 'Social Boldness', which has been polarized as "adventurous vs. withdrawn" by Kanthamani & Rao, 1971, p. 194). This factor has correlated with psi performance (Storm & Thalbourne, 1998-1999), and has been retested in the replication study with the *I Ching* (see Chapter 5).

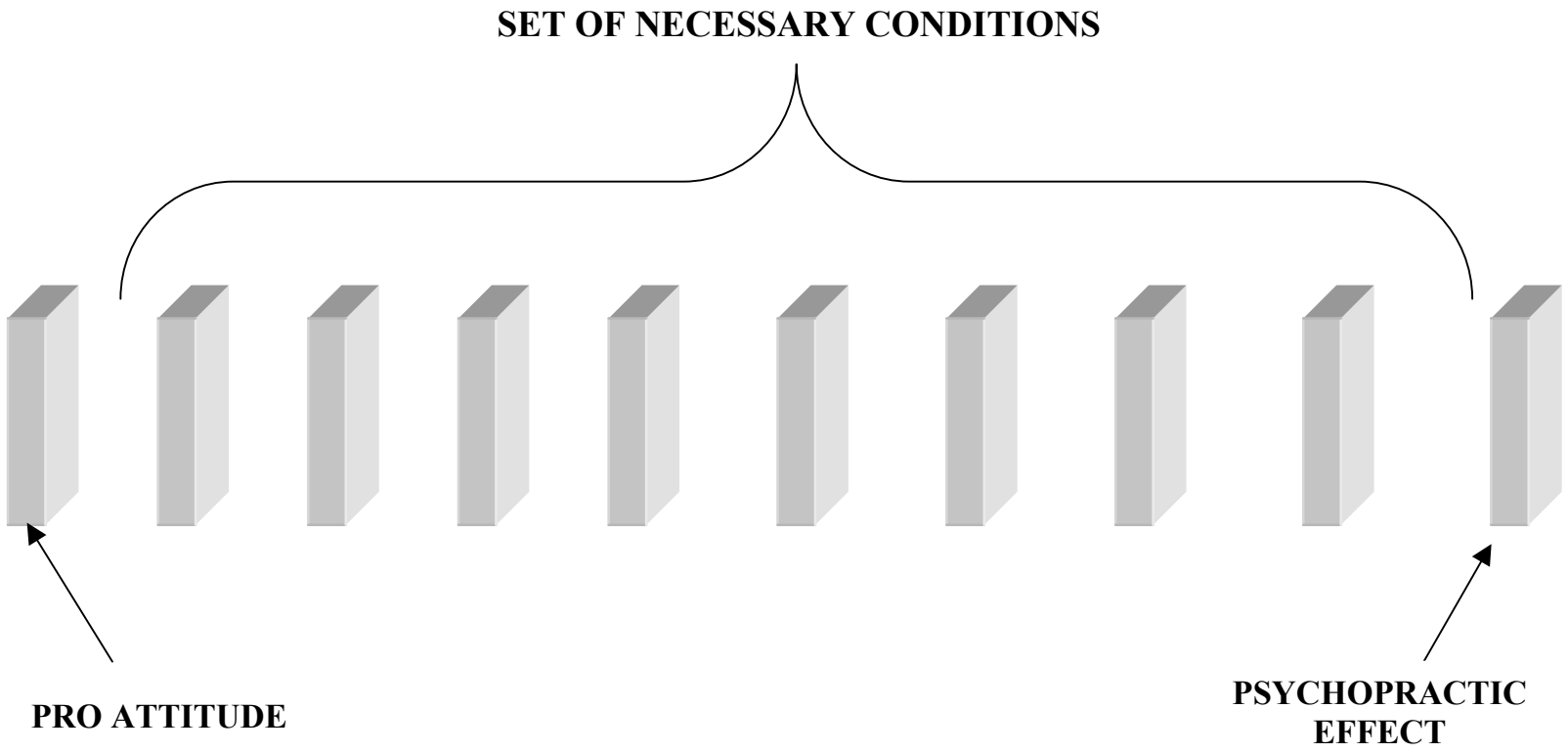

Figure 4.2. The domino analogy of psychopraxia.

For Thalbourne, various combinations of the above-mentioned psychopraxia-conducive conditions work like dominoes lined up in a series (see Figure 4.2)—take one condition out of the 'line-up' and a psychopractic effect may fail to eventuate. Note that the mechanical analogy shown in Figure 4.2 is a valid description of the psychopractic process only insofar as it is a model, but it does introduce a chronological and mediational dimension to the psychopractic process that may not exist. In fact, psychopractic causation may also depend on a 'concomitancy' of conditions that must be effected simultaneously in 'real' time, and not necessarily contiguously as in the domino analogy. Furthermore, a necessary condition in one situation may not be necessary in another (see Thalbourne, 2000a, p. 96, where conditions are also referred to as "appropriate" for an effect).

Although the theory of psychopraxia subsumes normal and paranormal actions under one rubric, it does not follow that a condition necessary or conducive to exosomatic psychopraxia (i.e., a paranormal effect) should be necessary or conducive to endosomatic psychopraxia (i.e., a normal effect), or some other manifestation of exosomatic psychopraxia. Thus, in all cases where necessary conditions bring about psychopraxia, it should be understood that the necessary condition or conditions needed on a given occasion are task-dependent, irrespective of whether the task may be classed as normal or paranormal.

Thalbourne (2000a, pp. 99-102) argues that investigators may appeal to the theory of psychopraxia in order to appreciate better just how close ESP and PK really are in nature, and how both are underscored by one singular process. Psychopraxia achieves parsimony by blocking off the old ESP-PK dichotomy, and in fact avoids the ESP-PK debate altogether.

Thalbourne (2000a) introduced the theory of psychopraxia as the means by which the nature of the paranormal might be described with less ambiguity than is currently the case due to the concomitant use in science of the three ontologies just described, but it does not conflict with any of them. Since the theory circumvents the ontological debate over whether a psi effect is a case of ESP or PK, or both, more time and opportunity become available to the investigator for the pursuit of an understanding of the locus and content of pro attitudes, and the conditions that affect goal achievement, whether that goal is physical or mental. An increased understanding of these factors may result in a more precise operationalization of parapsychological experiments, which may ultimately manifest as increased effect sizes for the relevant paranormal domains.

4.5 Experimental Psychopraxia

As yet, exosomatic psychopraxia (or, "*exo-psychopraxia*," Thalbourne, 2000a, p. 87) specifically has not been empirically tested under that name. For the theory to be tested adequately three factors must be implemented in the experimental design: (1) *the diasomatic hypothesis*, (2) *the 'pro attitude serving' hypothesis*, and (3) *necessary conditions*.

1. *The diasomatic hypothesis:* Psychopractic effects that are endosomatic (i.e., normal cognition and motor control) are givens (i.e., proved by common experience, psychological verification, etc.), so testing the diasomatic hypothesis within a parapsychological paradigm implies testing the extent to which psychopraxia may function exosomatically. Exo-psychopraxia may operate in two ways:

   (a) In a *compensatory* or substitution mode, taking the place of endosomatic psychopraxia. In the old language, psi operates when the psi-on-the-mind/body-complex does not (Thalbourne, 2000a, p. 50).
   (b) As a "spill-over" effort brought about by the organism as a result of high levels of activity "generated by volition on the nervous system" (Thalbourne, 2000a, pp. 49-50).

2. *The pro attitude serving hypothesis:* In most cases of paranormal experimentation, the goals of the participants in the sample, as instructed by the experimenter, are assumed to be task performances that result in the measurable achievement of paranormal information acquisition (ESP), or measurable paranormal changes in a physical system (PK). The theory of psychopraxia subsumes these effects under one, ontologically neutral rubric. The theory maintains that the self achieves goals in the exosomatic domain, by holding a pro attitude towards the stated outcome.

3. *Necessary conditions:* It is important for the theory of psychopraxia that "certain conditions" (Thalbourne, 1982, p. 96) that result in the fulfillment of a pro attitude be identified. To demonstrate the effect of these conditions in the experimental situation, the conditions must be measurable and have variability. In the present thesis, certain conditions are hypothesized as being 'necessary'. This terminology tentatively refers to conditions that are also to be termed 'sufficient' *in the given experimental situation* in bringing about a paranormal effect, otherwise called exo-psychopraxia. Where identified, for example, as being psi-conducive variables in

correlations that are found to be significant in the planned analyses, these conditions will be re-tested (post hoc through median-split analyses, etc.) to determine whether they are sufficient as well. If they are sufficient, then they will be classed as circumstantially necessary.

It is the aim of the present thesis to demonstrate aspects of the theory of psychopraxia by incorporating these three factors into the designs of five experiments, which form the main content of the thesis. The experiments lay particular emphasis on (i) the ambiguity of the traditional ESP-PK dichotomy, (ii) the changeability of the pro attitude, (iii), compensation, and (iv) necessary conditions.

4.6 Conclusion

The traditional view of paranormal processes is that they can usually be placed in either one of J. B. Rhine's two categories, ESP or PK. Thouless and Wiesner (1947) attempted a unification of these two categories, subsuming them under the super-ordinate category of psi, and the psi process has since been used to refer to anomalous influences on mental and physical systems. Since then, a number of theorists (Schmeidler, 1987, 1994; Jung, 1960; Stanford, 1974a, 1974b, 1978, 1990; and May, Utts, & Spottiswoode, 1995a, 1995b) have tried to subsume ESP and PK more fully into one or other category, but attempts have been impeded, either by theoretical and philosophical biases, or by insufficient experimental evidence.

The theory of psychopraxia is an attempt by Thalbourne (2000a) to resolve the ESP-PK dichotomy, first, by recognizing that the distinction exists only on the basis of philosophical assumptions, and second, by unifying the two formerly categorical effects as one singular psychopractic effect. To paraphrase Thalbourne's (1982, pp. 76, 96) definition: Psychopraxia underlies all interactions between the individual and mental/physical events, whereby under 'certain conditions' the adoption by the self of a pro attitude towards a certain goal results in the achievement of that goal.

Endo- and exo-psychopraxia only take effect if certain conditions are met, and these mandatory conditions are regarded as necessary and sufficient. There would be many other psychopraxia-conducive conditions similar to those listed above (see *4.4.4*), and some set of them would be hypothesized as sufficient for a given psychopractic effect. A series of five experiments have been conducted to illustrate the theory of psychopraxia. These experiments are presented chapter by chapter, starting with the next, Chapter 5.

CHAPTER 5
AN EXPERIMENT WITH THE *I CHING* (A REPLICATION)[42]

The method of the I Ching *does indeed take into account the hidden individual quality in things and men [*sic*]*.
C. G. Jung (1989, p. *xxviii*)

It is easy to see, hard to foresee.
Benjamin Franklin (circa late-18th Century, cited in Schick & Vaughn, 1995, p. 25)

5.1 The '*I Ching* Experiment'

This chapter is a report of a replication experiment conducted in 1998 by Storm and Thalbourne (1998-1999). They carried out a study with the *I Ching* (an ancient Chinese system of divination) to determine, amongst other things, whether or not transliminality[43] "might function as a connecting principle between paranormal effects and other personality variables" (Storm & Thalbourne, 1998-1999, p. 100). Participants were required to complete the Transliminality Scale (TLS; see **5.3**), and Cattell's 16PF Questionnaire (see **5.4**). (See **5.6** for a more detailed description of the experiment.) The experiment was deemed successful in the light of statistically significant (or marginally significant) confirmations of six out of eleven hypotheses (details of their results are reported in *5.2.5*, *5.3.4*, and *5.4.2*).

The aims of the present study were to:

1. Replicate the significant effects found in Storm and Thalbourne's (1988-1999) initial *I Ching* experiment.

2. Find experimental evidence that might substantiate Thalbourne's (2000a, pp. 1-15) claim that the traditional ESP-PK division can be unworkable in practice.

[42] A shorter version of this chapter was published as a refereed article (see Appendix AE).

[43] Recall from Chapter 1 (see *1.3.4.2*) that Transliminality is defined as "a hypothesised tendency for psychological material to cross thresholds into or out of consciousness" (Thalbourne & Houran, 2000, p. 853).

3. Show that exo-psychopraxia is the result of a pro attitude (Thalbourne, 2000a, pp. 93, 95-96). In the present study, the 'necessity' of the pro attitude was tested.

4. Show that psi is not so much "goal-oriented" as "pro attitude serving" (Thalbourne, 2000a, p. 66). Irwin (1999, p. 160) describes goal-orientation as *non-mediational* (i.e., not requiring a flow of information by various means such as is posited in the "cybernetic" theories—see also Stanford's, 1978, comments in *4.2.4*). In the present study, exo-psychopraxia was tested as a possible *mediational* process.

5. Test Thalbourne's (2000a, p. 96) claim that various 'conditions', such as level of transliminality, belief in paranormal processes, and specific personality traits are 'necessary' (see **4.5**) in bringing about effects, which Thalbourne describes as psychopractic.

The first aim is self-evidently necessary—replication studies in parapsychology are not only proof-oriented toward confirmation of the psi hypothesis, but also serve the purpose of confirming the consistency of the paranormal process. The other four aims are put forward to test the theory of psychopraxia. The principles underlying these aims are explained next.

5.1.1 The ESP-PK Dichotomy (Reprise)

In Chapter 4 it was stated that the theory of psychopraxia finds fault with, and indeed, rejects the ESP-PK dichotomy. In conventional parapsychology, it was shown that at least three possible classifications may be used to describe one and the same psi effect—it may be (i) ESP, or (ii) PK, or (iii) ESP *and* PK. Thalbourne (2000a, pp. 73-76) argues that nothing is served by this ambiguity, and thereby prefers the term psychopraxia to describe, in this case, the paranormal process. It was also claimed that for psi experimentation each of these three possible 'explanations' may be found wanting, and indeed there is no consensus amongst parapsychologists as to which explanation(s) should be preferred.

Thalbourne (2000a) presented the theory of psychopraxia as an alternative viewpoint on the alleged paranormal process taking place in the typical psi experiment, and drew attention to a possible lack of parsimony in any one (or all three) of the above-mentioned parapsychologically conventional explanations. He (Thalbourne, 2000a, pp. 1-2) even implied that the social position and therefore credibility of a "revolutionary" parapsychology could be undermined by those already skeptical of the discipline's basic tenet (i.e., that there are established in the field two unexplained statistical anomalies, both

described as psi, and theorized as paranormal). Thus, the theory of psychopraxia draws attention to the problems that are inevitable if parapsychological experimentation proceeds with the ESP and/or PK constructions in mind. This chapter, in part, considers the paranormal process from a 'psychopractic' (i.e., non-dichotomous) point of view in an attempt to demonstrate the hypothesized redundancy of the terms ESP and PK.

5.1.2 Testing the Concept of the Pro Attitude

The *I Ching* experiment is suitable for testing the concept of the pro attitude because of the nature of the *I Ching* process. There are always two groups in the *I Ching* experiment—those participants designated 'first-hexagram-hitters' because the first hexagram they generate is one of 16 hexagrams preselected on the descriptor form, and those designated 'first-hexagram-missers' because the first hexagram they generate is *not* one of 16 preselected hexagrams. In each group, there are subgroups of participants who generated changing lines, so that a second hexagram is generated (see *5.2.4* for an explanation of this process).

Some second hexagrams will match one of the 16 preselected hexagrams. This kind of match is called 'second-hexagram-hitting'. However, participants are never required to generate a second hexagram that matches one of the 16 preselected hexagrams, so that, according to the theory of psychopraxia, participants should never hold pro attitudes for second hexagrams (this issue is taken up again in *7.1.1*).

These two groups ('first-hexagram-hitters' and 'first-hexagram-missers') are each tested in the study described in the present chapter to see if their second-hexagram hit-rates are significant (Note that there is a statistical reason for having two groups—see *5.2.6*.) Results from these tests may be consistent with the psychopractic hypothesis that paranormal success depends on a pro attitude. The argument follows thus: *Given sufficient other conditions, if there is significant hitting on first hexagrams, then a sufficient number of pro attitudes were present; therefore, if there is no significant hitting on second hexagrams, then a sufficient number of pro attitudes were not present*. The absence of a pro attitude toward hitting on the second hexagram should be reflected in scoring that does not differ significantly from *MCE* because these hexagrams should not be targeted by participants in accordance with the task requirement.

5.1.3 Psychopraxia as a Mediational Process

Stanford (1977b) has observed that the results of many PK experiments indicate that psi might not work by means of information transfer or processing, as do normal cognitive

processes. Kennedy (1978, 1995) also found that ESP and PK do not necessarily involve information processing. There appears, then, to be some justification for models of paranormal 'process' that can be described as non-mediational. That is, psi may not necessarily take the form of information flow between 'object' and 'subject'—environment and individual—which may, for example, depend on wave-forms, particles, or so-far undiscovered forces (see Stokes, 1987, pp. 111-135, for a review of some of these theories). In fact, using theories or models that suggest information processing in one form or another to explain paranormal knowledge acquisition or action may simply be a prejudice of conventional parapsychological thinking.

While Thalbourne (2000a, p. 55) states that psi is a goal-oriented process, he later prefers that psychopraxia be termed a 'pro attitude serving' process (p. 66). The theory of psychopraxia allows for information processing to the extent that it fulfils the obligation of being a sufficient condition in a given situation. Some paranormal phenomena merely present in such a way as to suggest the goal-orientation hypothesis, leaving open the possibility that subtle information processes may well occur in psychopractic action, but may not be detected in any measurable way.

It is possible to test the hypothesis that psychopraxia may be a mediational (i.e., informational, or cybernetic) process in the *I Ching* experiment if so-called 'hexagram-hitting' in Sample 1999 is significantly lower than it was in Sample 1998. This hypothesis is plausible because the new Hexagram Descriptor Form (Appendix D) used in the follow-up study does not contain the 64 hexagram symbols, as does the old Hexagram Descriptor Form (Appendix E). If paranormal mediation is based on information processing then the new descriptor form, being harder to use (since it has no symbols), would mean an additional, *anomalous* step must be taken by the participants to reach their goals of successfully achieving hits. If there is a significant difference in hitting between the two samples, then exo-psychopraxia might be a process involving information processing that is pro attitude serving rather than goal-oriented.

It was hypothesized, therefore, that the hexagram hit-rate for Sample 1999 is lower than the hexagram hit-rate for Sample 1998. This hypothesis was tested using Rosenthal and Rubin's (1989, p. 334) procedure for testing the statistical significance of the "heterogeneity" of the obtained π values by means of a chi-square test on the two independent studies.

5.1.4 The Concept of the 'Necessary Condition'

The necessary condition in the theory of psychopraxia can be rendered by analogy as being like one domino in a series of dominoes, parenthetically embraced by the pro attitude on one side and the psychopractic effect on the other (see Figure 4.2). Each 'domino' must be strategically 'placed' in that series so that, prior to the pro attitude, it fulfils its obligation of maintaining a scientifically causal[44] chain of events, where each domino initiates the mandatory next step, and is itself a step towards psychopraxia (see *4.4.4* for further clarification of this model).

Chapter 4 more fully details the nature of the necessary condition, and includes a partial inventory of psi-conducive conditions already identified in the literature (see *4.4.4*). Added to that inventory were some more recently identified psychopraxia-conducive conditions that are hypothesized as being necessary and sufficient *in a given experimental situation* for bringing about the psychopractic goal. (Note that this switch in terminology, from psi-conducive to psychopraxia-conducive, serves the primary purpose of avoiding the narrowly defined term 'psi-conducive' that has come to connote only paranormal process.[45])

Theoretically, psychopraxia-conducive conditions bring about effects that can be classed as either normal or ostensibly paranormal (see again, *4.4.4*, for a specific clarification on the issue of 'task-dependence'—conditions identified as necessary because they are sufficient in one task, normal or paranormal, may not be necessary in another).

Storm and Thalbourne (1998-1999) found that transliminality and certain 16PF factors correlated with success at a psi task. An attitudinal variable—belief in one's "psi ability"—was reported as an indirect predictor of psi success so that belief (possibly in a variety of forms) may also qualify as a possible psychopraxia-conducive condition (see *5.3.3* for details). Assuming the successful replication of the above correlates, the same conditions may again be regarded as psychopraxia-conducive. These conditions will also be tested to determine whether they are necessary and sufficient in the *I Ching* experiment.

---

[44] See Storm (1999, pp. 260-261) for an explication of scientific causality. Essentially, in scientific circles a causal connection between cause and effect is recognized as long as the requirement 'if A, then B' is fulfilled.

[45] Psychopraxia and psi are not interchangeable terms, but since the present thesis is fundamentally a parapsychological work, psychopraxia-conducive conditions that are instrumental in normal cognition and motor function will not be scrutinized here. Therefore, in this thesis both terms—psi and psychopraxia—will describe, or otherwise imply, the same ostensible paranormal processes as suggested by the significant findings.

Given that the hypotheses in these four sections are falsifiable and can be tested experimentally, it was therefore deemed a suitable undertaking to attempt a replication study of the *I Ching* experiment. The following three components form the basis of the experiment: (1) the *I Ching* divinatory system, (2) the concept of transliminality, and (3) Cattell's 16PF Questionnaire (Cattell & Kline, 1937; Cattell, Eber, & Tatsuoka, 1970; Russell & Karol, 1994). These three components are now described.

5.2 The *I Ching*

5.2.1 Description of the I Ching *and its Introduction to the West*

As a divinatory system the *I Ching* consists of a book of 64 hexagrams (six-line symbols) and their corresponding 'readings'. The user poses a question, throws three coins six times (the modern technique), which generates a hexagram reading depicting a scenario (with commentary) that allegedly describes the past, present, and future life situations of the user in the context of the question.

The *I Ching* was first introduced into the English-speaking world by James Legge (1899, cited in Jung, 1989, p. *xxi*), who translated the original Chinese text. Legge's text gives evidence of a rationalist, positivistic skepticism about any possible paranormal process underpinning the *I Ching* system. Although his version was directed towards an academic audience, rather than a general readership, it does offer a rather pragmatic, albeit anachronistic, interpretation of the text.

By comparison, the Wilhelm/Baynes (1989) joint-effort is a more stoical and comprehensive interpretation, but is also complex and difficult to use.[46] In modern times the *I Ching* has become one of the major divinatory systems in current use throughout the West and its continued use suggests that the system "produces the intended effect" in the user—an effect that cannot be explained by chance alone (Brier, 1974, p. 5).

5.2.2 The Major Premise of the I Ching

The fundamental premise of the *I Ching* is that an anomalous (ostensibly paranormal) process underscores its function, with the implication that the structure of the system is somehow attuned to certain individuals (see *5.2.3* for details). Jung (1989, p. *xxii*) argued that chance is given a free rein in the process and that an acausal principle

[46] The Wilhelm-Baynes version and Hazel's (1990) 'New Age' interpretations of the 64 hexagrams were used to derive the *I Ching* Hexagram Descriptor Form in the present study (see *5.6.2*).

underlies the outcome, but his claims are driven by certain assumptions made under the rubric of synchronicity theory (see Storm, 1999).

It is equally valid to argue the case that the rules of chance can actually be subverted by intention (conscious or unconscious) so that the effects of chance are reduced. The outcome may then be less coincidental and more 'veridical' in nature. That is, the pre-disposed participant introduces his or her own influence into the system, so that human involvement—the crucial part of the *I Ching* process—may anomalously generate the appropriate reading. Thus, the system may be dependent on other than normal processes (i.e., a paranormal process).

5.2.3 Origins of the I Ching *Philosophy and its Structural Dynamics*

It is claimed that King Fu Hsi (2953-2838 BC) laid down the basics of the *I Ching* system (Legge, 1971). King Wên, the founder of the Chou dynasty (1150-249 BC), wrote most of the original text, and his son Chu Kung, the Duke of Chou, expanded the text further. Buddhist and Taoist philosophies are also evident in the *I Ching*, as is the influence of Confucius, who authored the commentaries that supplement the readings.

The major principle of the *I Ching*, which instills it with a meaningful structure, is the binary or yin/yang philosophy. It is this philosophy that activates the *I Ching*'s meaningful structure, in the sense that change is inevitable and comes about as a result of the resonance between two polar opposites—yin and yang—a resonance epitomized as the dynamic interplay of the "archetypal" opposites of yin and yang (Capra, 1988).

Traditionally, yin represents the principles of Earth, passivity, the 'feminine', the negative, and darkness, while yang represents the principles of Heaven, activity, the 'masculine', the positive, and light. According to Taoist philosophy, the two actually 'depend' on each other since together they 'hold the universe in balance'. The principle of two-ness or duality then is the dynamic or resonating effect that two opposing or antithetical principles embody in their 'relationship' (Capra, 1988). The two elements of the yin/yang binary system are represented as 'broken' and 'unbroken' lines in the *I Ching* (see Figure 5.1).

Figure 5.1. The yin and yang lines.

Four ‘duograms’ are then generated through various combinations of the two lines: yin/yin, yin/yang, yang/yin, and yang/yang (see Figure 5.2).

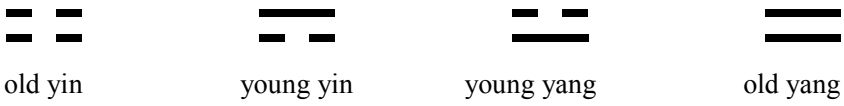

Figure 5.2. The four duograms.

The duograms represent a four-step continuum, moving from old yin to old yang—a gradual shift from one polar opposite to the other. By the addition of a yin or a yang line to each of the duograms, eight trigrams (the *pa kua*) are formed (see Figure 5.3).

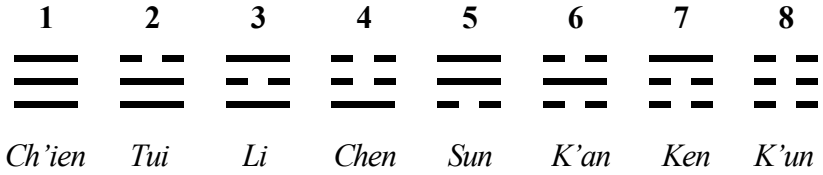

Figure 5.3. The eight trigrams (*pa kua*).

The trigrams also represent a continuum, but in eight steps, from Heaven (*Chi’en*) through to Earth (*K’un*). The individual trigrams are magnified in meaning above and beyond the simple degrees of yin and yang, or the duograms. For example, Heaven (☰) symbolizes creativity, masculinity, activity, and the Father.

The trigrams were also given a basic characteristic of the world of nature and social experience. From the social world the eight trigrams were also associated with Mother, Father, three sons, and three daughters (the ‘sons’ represent the principle of movement in its various stages—‘beginning of movement’, ‘danger in movement’, and ‘completion of movement’, while the daughters represent devotion in its various stages: ‘gentle penetration’, ‘clarity and adaptability’, and ‘joyous tranquility’ (Wilhelm, 1989; see Table

5.1). Therefore, natural phenomena, such as water, thunder, and earth, etc., appear in the symbolism of the trigrams.

Table 5.1
Meanings and Associations of the Eight Trigrams.

Ch'ien	☰	heaven	Creative, male, active	Father
K'un	☷	earth	Receptive, female, passive	Mother
Chen	☳	thunder	Movement, perilous	1st son
K'an	☵	water	a pit, danger	2nd son
Ken	☶	mountain	Arresting, progressive	3rd son
Sun	☴	wood, wind	Gentle, penetration	1st daughter
Li	☲	fire	Brightness, beauty	2nd daughter
Tui	☱	marsh, lake	Pleased satisfaction	3rd daughter

[SOURCE: Barrett, D. V. (1992). *Destiny and your dreams*. London: Treasure Press.]

Finally, pairing the trigrams produces the 64 hexagrams (see Appendix E). For example, the combination of two trigrams, *K'un* plus *Ch'ien*, gives Hexagram #11, which is called *T'ai* and means 'Prosperity' (see Figure 5.4). In most cases the hexagrams have new names that describe their unique qualities.

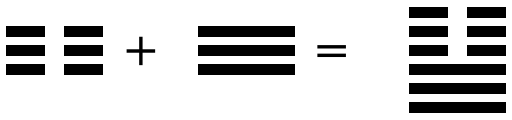

Figure 5.4. Constructing Hexagram #11: 'Earth (*K'un*) + Heaven (*Ch'ien*)' = 'Prosperity' (*T'ai*).

The hexagrams (the paired trigrams) also range as a continuum from the most yang (Hexagram 1: 'Heaven above Heaven') to the most yin (Hexagram 64: 'Earth above Earth'; see Figure 5.5). The 64 hexagrams each have their own unique reading, and it is the readings that form the basic text of the *I Ching*.

Chi'en/The Creative *K'un/The Receptive*

Figure 5.5. Hexagram 1: Heaven above Heaven (*Chi'en*) and Hexagram 64: Earth above Earth (*K'un*).

Covello (1977) found that there was a rationally derived, systematic substructure to the *I Ching*, and he claimed:

> One of the more fascinating aspects of the *I Ching* is that it can be viewed as a mathematically ordered cosmology. Its surprisingly systematic structure renders many of its assertions amenable to controlled investigation. (p. 115)

On this basis, Covello investigated the various symbolic systems in the *I Ching* for their meaningfulness. He found, by analyzing the line structures making up the hexagrams, and classifying the words used in the text, that "concreteness" (associated with "earthy, pictorial, and regressive") was characteristic of the "Yin principle," and "abstractedness" (associated with "heavenly" and "potency") was characteristic of the "Yang principle." Within the abstract-concrete dimension exists a continuum or spectrum of meanings as determined by the number of yin or yang lines in the hexagrams.

It is important to recognize that yin and yang are not either good or bad, as Von Franz (1980) makes clear:

> The Chinese were detached and philosophical enough to say that even if it is bad for me it might be good as a whole. From the beginning they had a wiser or more objective view of what we call good and bad, and saw it more as something in the ensemble of existence. (p. 47)

Earlier, Jung (1949, cited in Jung, 1989) claimed that "the method of the *I Ching* does indeed take into account the hidden individual quality in things and men [*sic*], and in one's own unconscious as well" (p. *xxviii*). This claim was partly prompted by Jung's use of the *I Ching*, but he was also responding in terms of an ancient Chinese philosophical way of thinking where "natural laws are merely statistical truths and must necessarily allow for exceptions" (Jung, 1989, p. *xxii*).

5.2.4 *Use of the* I Ching

To consult the 'oracle' of the *I Ching*, one must first generate a hexagram, which involves, first, posing a meaningful question to the *I Ching* (yes/no questions are excluded), followed by the repeated casting of 64 yarrow sticks to generate all six lines of the hexagram. (The modern method involves throwing three coins six times. The 'coin method' was used in the present *I Ching* experiment.)

In the coin throwing method, each throw of the coins produces one of the six yin or yang lines, so that, throw by throw, the hexagram is built from the bottom up. The coins may show three-of-a-kind on some occasions, which means the yin or yang line is a 'changing' line (with its own unique reading: the so-called *yao* text) that changes into its opposite—either yin to yang, or yang to yin. In other words, a second hexagram is produced. One reads the first hexagram reading (the 'judgement') in the context of past and present events, and the second hexagram reading is said to be a forecast or prognostication of the scenario discussed in the first reading to be read in the context of the question.

5.2.5 *Experimentation with the* I Ching

It should be noted that the *I Ching* is not a standard piece of psychological (or even parapsychological) apparatus. Nevertheless, a number of studies have shown that ostensible paranormal effects may be at work in the *I Ching* process. The following is a review of the literature on experimentation with the *I Ching*.

One of the earliest experiments with the *I Ching* was by Rubin and Honorton (1971, 1972). Participants were asked to generate a hexagram in response to a specific question. They were each given two readings: (a) the correct one, and (b) a control reading. Both readings were rated on a scale of 1 to 10 by the participant according to how relevant or accurate the participant felt them to be (the difference between the two ratings being the dependent variable). There was no overall significant result for the whole sample, but a sheep-goat rating scale had also been administered, and those participants who believed in ESP scored significantly higher than those who disbelieved.

A replication of Rubin and Honorton's (1971, 1972) experiment was conducted by Thalbourne, Delin, Barlow, and Steen (1992-1993). Participants followed the same procedure of generating a hexagram and rating two readings (the correct one and a systematically selected control). There were three planned "parapsychologically-relevant" hypotheses, and the following results were found:

(i) Although the mean difference-score was above chance, it was not significantly so.

(ii) There was no significant positive correlation between scores on the Australian Sheep-Goat Scale (Thalbourne & Delin, 1993) and difference-scores.

(iii) Those who believed in the efficacy of the *I Ching* scored significantly higher on difference scores than those who disbelieved.

Regarding finding (ii), Lawrence (1994) used Rosenthal's (1991, p. 63) Formula 4.2 for testing the significance of the difference between effect sizes to test the nonsignificant finding of Thalbourne et al. (1992-1993) against Rubin and Honorton's (1971, 1972) significant finding. When the original mean difference scores were converted to effect size estimates, Lawrence found no significant difference between the two. The conclusion was that Thalbourne et al. had effectively replicated Rubin and Honorton's result.

Thalbourne (1994) conducted another experiment using the *I Ching*, but this time focussed on changing lines exclusively. Using his own ecologically sound database, Thalbourne found that the number of changing lines in mid-1992 was significantly higher than mean chance expectation (1.66 changing lines where *MCE* = 1.5 changing lines). Later, in early 1994, the number of changing lines dropped to 0.70 lines, which was significantly lower than *MCE*. Thalbourne (1994) concluded that "familiarity" with the *I Ching* system had over time reduced his "information-hunger." Thus the *I Ching* process seemed to be influenced by an experimenter effect, so that more "static" hexagrams (i.e., no changing lines) were generated in the later period compared to the earlier period (p. 133). This "motivational hypothesis" (p. 133) may explain the earlier result.

While these studies have, to varying degrees, supported the hypothesis that the *I Ching* process may involve anomalies, the study by Storm and Thalbourne (1998-1999) went further by looking for relationships between paranormal ability in the *I Ching* 'setting' and two dimensions not previously tested in any *I Ching* experiment—transliminality (see *5.3.4*) and other aspects of personality (see *5.4.2*).

In relation to paranormal performance in the *I Ching* experiment, Storm and Thalbourne's (1998-1999) sample of 93 participants (mainly university students) produced an overall hit-rate of 32% on hexagram-hitting (P_{MCE} = 25%), which was described as "marginally significant" (p. 109; p = .067). In a follow-up research note (Storm [&

Thalbourne], 2001a), a direct significance test of the effect size was found to be significant ($\pi = .59$, $p = .048$). Number of changing lines was at chance.

5.2.6 Hexagram-Hitting Expressed as Effect Size π—Recent Findings

As just reported above (see *5.2.5*), a hit-rate of 32% was achieved for the sample of 93 participants (Storm & Thalbourne, 1998-1999, p. 109). Storm and Thalbourne considered it wise to remove the possibility that this result may have been an artifact. They tested the hit-rate of 'second-hexagram-hitters', but the overall hit-rate did not reach significance ($P = 0.30$; $p = .187$; Storm & Thalbourne, 1998-1999, p. 109). Note the fact that all participants ($N = 93$) were not instructed in any way to regard the second hexagram as a target. This same protocol applied to the participants in the replication study.

However, the *MCE* of a hit on the second hexagram (given as $P_{test} = 0.254$) for that group of 79 participants, applied only to the group of second-hexagram-hitters who did not get a hit on their first hexagrams ($n_1 = 52$). Their hit-rate was not 30% (as reported for all second-hexagram-hitters), but was actually 27%, which was still not significant ($p = .463$).

The same hit-rate of 30% (as reported for all second-hexagram-hitters) did not apply to the subgroup of second-hexagram-hitters who *did* get a hit on their first hexagrams ($n_2 = 27$). However, they had only 15 chances out of 63 of getting a hit on the second hexagram. Thus, $P_{test} = 15/63 = 0.238$. The hit-rate for this group of 27 second-hexagram hitters was 37%—also not significant ($p = .082$).

As was discussed above (see **5.1**, Aim 3, and *5.1.2*), these results have important implications for the concept of the pro attitude as proposed in the theory of psychopraxia (see also *5.5.2*, Hypotheses 2 and 3).

5.3 The Concept of Transliminality

5.3.1 The Origins of Transliminality

At the beginning of this chapter, transliminality was defined as "a hypothesised tendency for psychological material to cross (*trans*) thresholds (*limines*) into or out of consciousness" (Thalbourne & Houran, 2000, p. 853). This definition is in current use, and essentially describes the nature and process of transliminality. While many representations in consciousness ultimately have their origins in preconscious processes, the transliminal material represented in consciousness is recognized as having three specific sources: the

'subliminal' mind (the 'subconscious' or 'unconscious'), the 'supraliminal' mind (for example, states of consciousness and conscious activity), and the external environment.

The origins of transliminality trace back to 1991, when it was first conceptualized as "an openness or receptiveness to impulses and experiences whose sources are in preconscious (or unconscious) processes" (Thalbourne, 1991, pp. 181-182). This conceptualization very closely mirrors the above definition, although it does not account for the loss of this material to the subconscious or unconscious mind.

A similar two-way process of "leakage" of information between the subliminal mind and supraliminal mind was described as "transliminal" by Usher and Burt (1909, cited in Thalbourne, 2000d, p. 3). Rugg (1963, cited in Thalbourne, 2000d, p. 4) used the same word to describe the attraction of non-conscious creative material into consciousness. MacKinnon (1971; cited in Thalbourne, 2000d) referred to "transliminal experiences." These were the experiences in consciousness of material that came "out of the non-conscious into the vestibule of the conscious mind," but which could also go back again (Thalbourne, 2000d, p. 4).

These various descriptions of the 'transliminal' tend to converge in meaning—they all refer to material either coming into, or going out of consciousness, or both. Having established the origins and nature of transliminality, we now turn to the mainly correlational studies that bear on the transliminality concept.

5.3.2 Potential Psychological Constituents of Transliminality

Empirical evidence for a factor that was to become known as transliminality arose from a study by Thalbourne and Delin (1994). They administered a series of tests to samples of manic-depressives, schizophrenics and university students. It was found that scores on belief in the paranormal correlated with scores on measures of creative personality (four items constructed for that study, and five from Torrance's, 1971, Creative Motivation Inventory). Also, mystical experience (Thalbourne, 1991), magical ideation (Eckblad & Chapman, 1983), and history of manic-like and depressive experience (Thalbourne, Delin, & Bassett, 1994) correlated with belief in the paranormal.

Principal components analysis (a form of factor analysis) was conducted using the student data only (since it was the largest subset of data), and a single factor was identified. This was called transliminality (Thalbourne & Delin, 1994, p. 23). The above-mentioned variables were also classed as constituents of transliminality because they correlated significantly and positively with each other.

Another constituent of transliminality—'attitude to dream interpretation', from the MMPI (Dahlstrom, Welsh, & Dahlstrom, 1972)—was later identified by Thalbourne and Delin (1999).

Finally, Thalbourne, Bartemucci, Delin, Fox, and Nofi (1997) identified a further four potential constituents: schizotypal personality (Claridge & Brooks, 1984), 'fantasy-proneness' using the Inventory of Childhood Memories and Imaginings (Form C; Myers, 1983), 'absorption', which is "the ability to become totally absorbed in the object of attention" (Tellegen & Atkinson, 1974), and 'hyperæsthesia' (hypersensitivity to environmental stimulation; see Thalbourne, 1996b).

5.3.3 Psychological Correlates of Transliminality

Many correlates of transliminality have also been found (these are referred to as correlates because, unlike constituents, they do not correlate with all constituents). Thalbourne et al. (1997) found one correlate with transliminality ('depressive experience', which had dropped out of the factor analysis) and they also identified another two potential constituents: 'general religiosity' and 'frequency of dream interpretation' (Thalbourne & Delin, 1999).

Thalbourne et al. (1997) found three correlates: Eysenck's P (Psychoticism), Eysenck's E (Extraversion), and Eysenck's N (Neuroticism). (See Eysenck & Eysenck, 1991.) Age also correlated negatively with transliminality. Thalbourne (1998) then found another eight correlates:

(i) The Dissociative Experiences Scale (Bernstein & Putnam, 1986).
(ii) The Questionnaire of Experiences of Dissociation (Riley, 1988).
(iii) The Launay-Slade (1981) Hallucination Scale.
(iv) Unusual Experiences (short version) from the Oxford-Liverpool Inventory of Feelings and Experiences (Claridge, 1997).
(v) Reading about Eastern religions (Haraldsson, 1981).
(vi) Belief in life after death (Thalbourne, 1996a).
(vii) Vivid religious or spiritual experience (Haraldsson, 1981).
(viii) Depressive Experience (Thalbourne & Bassett, 1998).

In the same study, Thalbourne devised the Transliminality Scale (TLS)—a 29-item scale of true/false questions derived from nine constituent variables (see Appendix B).

Given that the Launay-Slade Hallucination Scale (Launay & Slade, 1981) already correlated with transliminality, Thalbourne (1999) hypothesized that vivid visual imagery

would also correlate with transliminality. The hypothesis was not confirmed. However, in a study on creativity, Thalbourne's (2000b) hypothesis that transliminality would correlate with the Creative Personality Scale (see *5.3.3*) was replicated, although transliminality failed to correlate with the Revised Barron-Welsh Art Scale (Welsh & Barron, 1963).

Looking further into the nature and correlates of transliminality, Thalbourne and Houran (2000) used two measures: the Mental Experience Inventory (Kumar & Pekala, 1992) and the AT-20 Scale (MacDonald, 1970) to measure tolerance of ambiguity. They found that participants scoring high on transliminality also tended to:

(i) "believe in more psi related and unusual events,
(ii) "to report more paranormal and unusual experiences,
(iii) "to have a greater sense of being high and that the mind is tremendously powerful,
(iv) "to be more given to introspection and daydreaming,
(v) "to experience altered consciousness."

There was no significant correlation with age or tolerance of ambiguity.

Transliminality has also been found to correlate with experience of Kundalini ("energy-like sensations . . . which surge upwards through the body," Thalbourne & Fox, 1999), "general afterlife belief" (Thalbourne, 1998-1999); and Openness to Experience as measured on the Quickscales (J. M. T. Brebner, personal communication, June, 2000)—a 30-item short version of the NEO PI-R personality inventory (Costa & McCrae, 1992; Thalbourne, 2000c). Storm and Thalbourne's (1998-1999) experiment especially focussed on transliminality, which correlated with five factors from Cattell's 16PF Questionnaire (see Table 5.2).

Table 5.2
Correlations between Transliminality and Five 16PF Factors

Variable	*r*	*p*
Factor A (Warmth)	0.21	.040
Factor G (Rule-Consciousness)	–0.26	.013
Factor M (Abstractedness)	0.41	<.001
Factor TM (Tough-Mindedness)	–0.33	.001
Factor SC (Self-Control)	–0.29	.005

Notes: $N = 93$, $df = 91$; p values are two-tailed

Highly transliminal participants were more likely to be 'warm' and 'outgoing' (Factor A), 'expedient' and 'nonconforming' (Factor G), 'imaginative' and 'idea-oriented' (Factor M), 'receptive', 'open-minded', and 'intuitive' (Factor TM), and 'unrestrained' (Factor SC). Transliminality also correlated with a belief measure ("Do you believe in your own abilities to predict the outcome hexagram, or influence the fall of coins so that the outcome hexagram matches one of your sixteen choices?").

Transliminality went through three changes in definition during these six years of experimental research. At first it was identified as an exclusively inwardly generated process, in the sense that the material seemed to originate in the subliminal mind only. As other sources were identified, particularly from the work of Thalbourne et al. (1997), the definition changed in order to include supraliminal sources deemed responsible for experiences such as absorption and active fantasy-proneness, and also sensitivity to external stimuli, such as hyperæsthesia.

Finally, the current definition of transliminality (see *5.3.1*) was arrived at through recognition of the fact that the transliminal process was analogous to a two-way street giving passage to material coming and going in both directions, across a number of thresholds, into and out of consciousness.

5.3.4 Transliminality and Paranormal Performance

With the current definition in mind, where transliminal processes may originate in external sources, and from the suggestion in Thalbourne and Delin (1994) that transliminality may correlate with paranormal processes, Thalbourne (1996b) hypothesized that transliminality scores would correlate with scores on a precognition task. The result was not significant. However, Sanders (1997; Sanders, Thalbourne, & Delin, in press) showed that, for senders, telepathic transmission of emotional states correlated significantly with transliminality.

Storm and Thalbourne (1998-1999) found a significant positive correlation between transliminality and hexagram-hitting. They also found a weak but marginally significant correlation between transliminality and changing lines. It was concluded from these findings, as suggested in Thalbourne and Delin (1994, p. 24), that transliminality might be related to an anomalous dimension of human behaviour that not only underlies paranormal belief (recall the belief-question/transliminality correlation in *5.3.3*), but also indicates the possibility of ostensible paranormal ability (Storm & Thalbourne, 1998-1999, p. 113).

More recently, however, Parker (2000a), using dynamic targets (short film excerpts) in the ganzfeld condition failed to demonstrate any correlation with paranormal

ability and an earlier version of the TLS (Form A)—the transliminality "scores of receivers making direct hits on the film clips were not discriminated from those making misses" (Parker, 2000a, p. 10).

In the next section the third and final component of the *I Ching* experiment will be explored: Cattell's 16PF.

5.4 Cattell's Sixteen Personality Factor (16PF) Questionnaire

5.4.1 Description of the 16PF

Raymond Cattell's (Cattell & Kline, 1937) 16PF is unique among personality tests and inventories in that it was derived from English-language adjectives that describe human behaviour and personality (Cattell used English language dictionaries to source these adjectives). He used factor-analytic techniques to arrive at the 16PF as it is currently known: a personality test of no less than 183 questions that produces a profile of 16 Primary Factors, and 5 Global Factors (which are derived from the 16 Primary Factors) (Russell & Karol, 1994).

The sixteen Primary Factors are Warmth (A), Reasoning (B), Emotional Stability (C), Dominance (E), Liveliness (F), Rule-Consciousness (G), Social Boldness (H), Sensitivity (I), Vigilance (L), Abstractedness (M), Privateness (N), Apprehension (O), Openness to Change (Q_1), Self-Reliance (Q_2), Perfectionism (Q_3), and Tension (Q_4). The five Global Factors are Extraversion (EX), Anxiety (AX), Tough-Mindedness (TM), Independence (IN), and Self-Control (SC).

5.4.2 The 16PF and Paranormal Performance

In the late nineteenth century, the psychical researcher Frederic W. H. Myers wrote: "We have indications of something complex and obscure in the structure of human personality, of something transcending sensory experience in the reserves of human faculty" (Myers, cited in Nicol & Humphrey, 1953, p. 133). Since that time, many studies in parapsychology have been undertaken with the intention of finding personality correlates of paranormal performance (see **5.5**). However, not until the early 1950s did the 16PF come into use in parapsychological experiments—it was still a "new and relatively untried instrument" at that time (Nicol & Humphrey, 1953, p. 144).

Nicol and Humphrey (1953) were amongst the first to use the 16PF in personality-ESP research. Three of nine 16PF factors used in their ESP experiment (a Zener card-

guessing task) correlated significantly with scores on the paranormal task: Factor C (Emotional Stability), which was a positive correlation, and Factor O (Apprehension) and Factor Q_4 (Tension), which were both negative correlations. These factors contribute to Global Factor AX (Anxiety). In fact, high ESP scores were likely to be held by low anxiety participants.

Nicol and Humphrey (1955) attempted to replicate these results, but failed. There were no significant correlations with 16PF factors and no directional trends. When they combined both samples, however, there was a return to the earlier significant results.

Kanthamani and Rao (1971, 1972, 1973) and Rao (1974) used the HSPQ (Cattell's High School Personality Questionnaire), which incorporates factors common to the 16PF. The following results were obtained:

(i) Kanthamani and Rao (1971) used concealed drawings as targets in their ESP task. Factor A (Warmth), Factor E (Dominance), Factor F (Liveliness), and Factor I (Sensitivity) correlated significantly and positively with ESP scores.

(ii) Kanthamani and Rao (1972) used extraversion (E) scores only in an ESP card test. Extraverts' scores were significantly above *MCE*, introverts were significantly below *MCE*, and there was a significant difference between the two types.

(iii) Kanthamani and Rao (1973) tested participants' N (Neuroticism) scores against scores on an ESP card-guessing task. Scores on the ESP task correlated negatively with Neuroticism.

(iv) Rao (1974) noted that (a) belief was related to performance on the ESP task; (b) extraversion and neuroticism were related to psi, and (c) combined personality factors were better predictors of psi performance than individual factors.

Sargent, Barlett, and Moss (1982) used the ganzfeld setting to test participants' performance against Factor EX (Extraversion) on the 16PF. The correlation found was positive, but nonsignificant.

Sudhakar and P. Rao (1986) also tested participants in a ganzfeld-GESP setting. Factor A (Warmth) correlated positively and significantly with ESP scores. This correlation may have been the result of multiple analyses, given that there are 21 factors in the 16PF.

Some of the results of Storm and Thalbourne's (1998-1999) experiment with 16PF factors have already been reported (see *5.3.3*). The following results are also relevant to the 16PF:

(i) Significant correlations of hexagram-hitting with six 16PF factors (see Table 5.3). Hitters were therefore more likely to be 'lively' and 'animated' (Factor F), 'socially bold' and 'venturesome' (Factor H), 'group-oriented' and 'affiliative' (Factor Q_2), 'relaxed' and 'patient' (Factor Q_4), 'extraverted (Factor EX), and 'independent' and 'persuasive' (Factor IN).

Table 5.3
Correlations between Hitting and Six 16PF Factors

Variable	*r*	*p*
Factor F (Liveliness)	0.26	.013
Factor H (Social Boldness)	0.41	< .001
Factor Q_2 (Self-Reliance)	–0.22	.031
Factor Q_4 (Tension)	–0.23	.029
Factor EX (Extraversion)	0.28	.007
Factor IN (Independence)	0.23	.030

Notes: $N = 93$; $df = 91$; p values are two-tailed

(ii) Significant correlations of number of changing lines with five 16PF factors (see Table 5.4). Participants scoring high on changing lines tended to be 'reserved' and 'impersonal' (Factor A), 'reactive' (Factor C), 'imaginative' (Factor M), 'self-reliant' and 'solitary' (Factor Q_2), and 'introverted' (Factor EX).

Table 5.4
Correlations between Number of Changing Lines and 16PF Factors

Variable	*r*	*p*
Factor A (Warmth)	–0.21	.040
Factor C (Emotional Stability)	–0.25	.014
Factor M (Abstractedness)	0.31	.003
Factor Q_2 (Self-Reliance)	0.30	.004
Factor EX (Extraversion)	–0.23	.028

Notes: $N = 93$; $df = 91$; p values are two-tailed

The next section looks specifically at extraversion-ESP research since Factor EX and constituent primary factors of Extraversion (which include Factor A, Factor F, Factor H, and Factor Q_2) featured frequently as correlates in Storm and Thalbourne's (1998-1999) I Ching experiment. Some extraversion-ESP studies using measures other than the 16PF are also reviewed since there is evidence that extraversion generally refers to the same personality factor across a number of personality measures.[47]

5.4.3 Extraversion-ESP Research

Extraversion is a primary personality factor in many of the major personality tests, including the 16PF, the Maudsley Personality Inventory (MPI—later to become the Eysenck Personality Inventory, or EPI. See Eysenck & Eysenck, 1965), the MMPI, and the EPQ. Consequently, the various forms of extraversion (regardless of their slight definitional variation—see n47 below) show consistent relationships with paranormal performance (viz., extraversion often correlates positively with psi hitting).

Åström (1965), for example, used the MPI to separate the extraverts from the introverts. Scoring in an ESP task was significantly above *MCE* in extraverts and significantly below *MCE* in introverts. However, when Green (1966a, 1966b) used the MPI (specifically the 'E' scale) in conjunction with an ESP task, she did not replicate Åström's result.

[47] Eysenck's E and Cattell's 'exvia' probably refer to the same personality variable, since they correlated very highly with a single factor in a study by McKenzie, Tindell, and French (1996, p. 272): Eysenck's E, $r = .91$, and Cattell's Factor EX, $r = .85$. Gentry, Wakefield, and Friedman (1985) constructed a new MMPI Extraversion scale in which MMPI items had to correlate significantly with E on the EPQ and found the new factor overall correlated significantly with Eysenck's E ($r = .79$, $p < .01$).

Nash (1966) found that scores on the Social Introversion Scale of the MMPI correlated significantly and negatively with psi scores in one of eight experiments. In five experiments, the correlations were consistently negative also (though not significant). Perhaps Nash's (1966) eight experiments should have been tested for differences in performance amongst themselves rather than individually tested for significance. These studies preceded Rosenthal and Rubin's (1989) recommendation that an effect size comparison between studies can be more fruitful for the psi hypothesis.

Interestingly, Eysenck (1967, p. 68) considered the introvert likely to be a poor "psi participant," but Palmer's (1977) and Irwin's (1986) reviews gave evidence that the introvert was actually a more likely to be a 'psi misser' (a participant who tends to score significantly below chance).

Palmer (1977, p. 187, Table 1) presented a table of 33 extraversion-ESP experiments, of which 23 (70%) showed positive relationships between extraversion and ESP scoring (see also Sargent, 1981, and the extraversion-ESP meta-analyses in *3.1.8* in Chapter 3). However, Irwin (1999, p. 102) warns that extraversion-ESP correlations may be more frequently significant and positive because high scorers on extraversion scales have the ability to relax and feel comfortable in the test situation. Irwin argues that introverts too might produce significant positive scores if they were able to relax as well in the test situation as they would do in other more familiar settings.

5.5 The *I Ching* Experiment

5.5.1 Description of the I Ching *Experiment*

The experimental component of this study involved the use of the *I Ching* in determining the ability of participants to (a) achieve a designated hexagram outcome, and (b) generate changing lines by throwing three-of-a-kind as often as possible (i.e., up to six times). Participants were first required to select sixteen hexagrams (as potential targets) according to their feelings or thoughts, as selected on the new Hexagram Descriptor Form (see Appendix D). The new Hexagram Descriptor Form used in this replication study does not contain the 64 hexagram symbols, as did the old Hexagram Descriptor Form (Appendix E) used by Storm and Thalbourne (1998-1999). Participants were then required to throw three coins, six times, to generate the changing line(s), if any, and the outcome hexagram. Participants were also required to complete the TLS (Form B) (see Appendix B), and Cattell's 16PF.

The analytical component of this study involved testing for overall hitting, and determining correlates, if any, between paranormal ability, transliminality, and 16PF personality factors, in accordance with the hypotheses given below (see *5.5.2* and *5.5.3*). With the completion of this replication study, there were two samples: Storm and Thalbourne's (1988-1999) 'Sample 1998', and the sample from this replication study, 'Sample 1999'. These two samples will be analyzed further and compared in the next chapter (see Chapter 6).

5.5.2 Parapsychological Hypotheses

The parapsychological and psychological hypotheses proposed in this study are mostly based on Storm and Thalbourne's (1998–1999) significant findings, both predicted and post hoc.

The following parapsychological hypotheses were proposed. (The tests used are given in parentheses with each hypothesis.):

1. Hexagram-hitting, when expressed as a proportion of hits, is at a rate greater than *MCE* (P_{MCE} = 0.25; binomial test). This is the primary analysis. As a secondary analysis, hexagram-hitting, when expressed as an effect size π, is at a greater rate than *MCE* (π_{MCE} = 0.50; Rosenthal & Rubin's, 1989, p. 333, Formula (1)—given in p. 35, n29—is used to calculate a π value, and Rosenthal & Rubin's, 1989, p. 334, Formula (4) is used to calculate a *z* score). This two-part hypothesis is one-tailed, because above-chance scores were observed in the first experiment.

2. Hitting for second-hexagram 'hitters' who were also first hexagram 'hitters' is above *MCE* (binomial test).

3. Hitting for second-hexagram 'hitters' who were first hexagram 'missers' is above *MCE* (binomial test).

4. Hexagram-hitting for Sample 1999 is lower than hexagram-hitting for Sample 1998 (chi-square test—see Appendix N).

5. There is a positive correlation between transliminality and hexagram-hitting (Pearson *r*)[48].

[48] The Pearson *r* tests were one-tailed for all hypotheses that used this test since they were all directional hypotheses.

6. There is a positive correlation between transliminality and number of changing lines (Pearson *r*).

7. Hexagram-hitting correlates positively with Factor F (Liveliness), Factor H (Social Boldness), Factor EX (Extraversion), and Factor IN (Independence) and negatively with Factor Q_2 (Self-Reliance) and Factor Q_4 (Tension) of the 16PF (Pearson *r*).

8. Number of changing lines correlates positively with Factor M (Abstractedness) and Factor Q_2 (Self-Reliance) and negatively with Factor A (Warmth), Factor C (Emotional Stability), and Factor EX (Extraversion) of the 16PF (Pearson *r*).

9. Number of changing lines correlates positively with answers to Question 2: "Do you think it is possible for at least some people to exhibit paranormal effects in this experiment, by predicting the outcome hexagram, or influencing the fall of coins so that the outcome hexagram matches 1 of their 16 choices?" (Pearson *r*).[49]

5.5.3 Psychological Hypotheses

The following psychological hypotheses were proposed. (The tests used are given in parentheses with each hypothesis.):

10. Transliminality correlates positively with answers to Question 2 (the "sheep" question) and with answers to Question 3: "Do you believe in your own abilities to exhibit paranormal effects in this experiment, by predicting the outcome hexagram, or influencing the fall of coins so that the outcome hexagram matches one of your sixteen choices?" (Pearson *r*).

11. Answers to Question 2 ("possibility") and Question 3 ("ability") correlates positively with each other (Pearson *r*).

12. Transliminality correlates positively with Factor A (Warmth) and Factor M (Abstractedness) and negatively with Factor G (Rule-Consciousness), Factor TM (Tough-Mindedness), and Factor SC (Self-Control) of the 16PF (Pearson *r*).

(In this study, the only 16PF factors that were tested were those that yielded significant results in Storm and Thalbourne's, 1998–1999, first study.)

[49] The relationships hypothesized in Hypotheses 9, 10, 11, and 12 were all based on significant post hoc discoveries of Storm and Thalbourne (1998-1999, pp. 110-111; 114-115).

5.6 Method

5.6.1 Participants

A total of 107 participants volunteered for this experiment.[50] Thirty-one (29%) were Adelaide University psychology students from all levels (undergraduate, honours, or postgraduate), and the remaining 76 (71%) were Adelaide University students not enrolled in any psychology course, at any level. These 76 participants were either students from other departments or were found through friends and colleagues by word of mouth. Fifty-four percent of the total sample were women. The mean age was 26 years (*SD* = 8.90).

5.6.2 Measures

Four measures were used in the experiment:

(i) The 'states of mind' scales (see Appendix A). This scale was used in another experiment to test for relationships between perceived simplicity of mechanical set-up(s) of paranormal task(s) and paranormal success, and perceived ease of paranormal task(s) and paranormal success—see Chapter 11.

(ii) Thalbourne's (1998) TLS (Form B), which contains 29 items taken from various scales, 14% of which refer to paranormal phenomena (see Appendix B). The participant answers "true" or "false" to each item, and the total number of "true" answers out of 29 is his or her transliminality score.

(iii) The *I Ching* Hexagram Descriptor Form, which includes a question about previous use of the *I Ching*, and 2 questions about belief in the *I Ching* process (see Appendix C). The form also contains 64 two-word descriptors representing each of the 64 hexagrams (Appendix D), but does not contain the 64 hexagram symbols as used in Storm and Thalbourne (1998-1999, p. 118; see Appendix E).

[50] Note that the odd number of participants (viz., *N* =107) should not raise the suspicion that optional stopping had taken place in the experiment. It was decided before the experiment began to test 107 participants with a view to combining them wherever possible with Sample 1998 (*N* = 93), thus yielding a round total of 200 participants (see Chapter 6 for the analyses of these two samples as a pooled data set).

(iv) Cattell's 16PF, designed to measure and "identify the primary components of personality," including five global factors (Russell & Karol, 1994, p. 7).

5.6.3 Apparatus

Ten sets of material were used in the experiment: (1) invitation to volunteers (Appendix F); (2) information sheet (Appendix G); (3) a consent form (Appendix H); (4) an *I Ching* hexagram file, containing an introductory page, a how-to-score page, and the 64 hexagram readings (one reading per page, totaling 64 pages, Wing, 1982, with the changing line readings on the back of each page, Wing, 1979); (5) three coins (Australian 10-cent pieces—75% copper, 25% nickel), a coin cup (for shaking the coins), and a felt-lined box (as a receptacle for the falling coins); (6) a score record sheet for recording coin throws (Appendix I); (7) a "How to generate an *I Ching* hexagram" sheet which are instructions to the experimenter on how to convert the outcomes of the coin tosses to "yin" and "yang" lines, and whether they were so-called changing lines (Appendix J); (8) an "eight-by-eight (8 x 8) trigram matrix" for calculating hexagrams (Appendix K); (9) a debriefing sheet for "hitters" (see Appendix L); and (10) a debriefing sheet for "missers" (see Appendix M).

5.6.4 Procedure

Once ethics approval was granted from the relevant departmental ethics committee, psychology students were approached to participate in the experiment by way of a written invitation lodged in their pigeonholes in the psychology department. Non-psychology students placed response slips in a 'ballot' box placed in the university library.

At the experimental sessions, participants first read the information sheet and then signed the consent form. The information sheet outlined the general nature of the experiment, describing it in three stages. Participants were instructed to take their time and start when they felt ready because there was no time limit. The three stages were as follows.

Part 1: Participants are required to completie the *I Ching* Hexagram Descriptor Form, which also asks three questions:

1. Have you ever used the *I Ching* before?

2. Do you think it is possible for at least some people to exhibit paranormal effects in this experiment, by predicting the outcome hexagram, or influencing the fall of coins so that the outcome hexagram matches 1 of their 16 choices?

3. Do you believe in your own abilities to exhibit paranormal effects in this experiment, by predicting the outcome hexagram, or influencing the fall of coins so that the outcome hexagram matches 1 of your 16 choices?

Question 1 separates the naïve from the sophisticated participants, and Questions 2 and 3 measure the participants' beliefs about the ostensible paranormal effects involved in the *I Ching* process.

Participants were required to choose 16 descriptor-pairs that they felt to be relevant to their feelings "Lately, or right now . . ." [51] These choices were not ranked. Under the watchful eye of the experimenter and a witness,[52] the participant then threw three coins six times, recording the number of heads and tails of each throw on the score record sheet, from the bottom up, according to the conventions of the *I Ching*.

Each of the six 'heads-and-tails' counts was converted to its respective hexagram line, as shown on the how-to-score page of the hexagram file (see "How to generate an *I Ching* hexagram"—Appendix J). A second hexagram was also generated if changing lines were produced from throws of three-of-a-kind. Hexagrams were decoded using the 8 x 8 trigram matrix (Appendix K). The bottom three lines and the top three lines each form trigrams, which are collated with each other with the aid of the trigram matrix to form the hexagram.

When the participant and the witness were satisfied that the hexagram(s) had been calculated correctly, they signed and dated the bottom of the score sheet. (N.B.: the experimenter was also witness to this whole process from the coin-throwing stage to the signing stage.) A 'hit' was a match of the participant's outcome hexagram with 1 of his or her 16 selections, as marked on the Hexagram Descriptor Form, whereas in the case of a "miss" there was no such match.

Part 2: When the *I Ching* component of the experiment was over, participants completed the TLS (Form B).

Part 3: The 16PF component (Russell & Karol, 1994) was the third and last stage of the experiment. Participants completed this component, and their tasks were thus finished.

---

[51] Traditionally, the *I Ching* process requires a "general question" (Hazel, 1990, p. 7), or a question "preferably of great personal relevance" (Thalbourne et al., 1992-1993, p. 13). Therefore, the use of the *I Ching* in this unorthodox study is itself somewhat unorthodox.

[52] The presence of a witness ensured that accurate recordings of coin throws were made. The witness also certified that no 'unsuccessful' coin throws were neglected (and re-thrown).

During the three stages, no feedback was given to participants as to whether or not they were successful at the paranormal task (i.e., whether or not they generated a hexagram that matched 1 of the 16 chosen on the descriptor form). Neither were participants given performance feedback or results on the TLS or the 16PF. These precautions were taken in order to eliminate the rival hypothesis that some (or all) of the significant personality correlates with hitting were caused by artifactual responses to 16PF questions as a result of knowledge of success or failure at the paranormal task (e.g., a participant's mood may shift from habitual introversion to temporary extraversion if he or she got a hit, or he or she might suddenly feel that there is some truth to the paranormal process after all). The same rival hypothesis would apply to significant correlates of transliminality with hitting and changing lines. (It is noted that lack of performance feedback on the paranormal task might also produce disgruntlement effects on the personality measures.)

Alternatively, it can be argued that normal feedback may be redundant given the possibility that anomalous knowledge of a hit or miss could still have a biasing effect on responses to the test items. It is also conceivable that the choices made on the Hexagram Descriptor Form could affect item responses. There may be no way of avoiding all these problems. As was stated in Storm and Thalbourne's (1998-1999, p. 108) initial study, there are methodological problems regardless of order of administration of materials and paranormal task, and in this study the author decided to avoid just some of the more pertinent problems, for example, the possible motivational obstacle participants might face when spending one hour, on average, on personality testing before the paranormal task.

Some time after testing, once scores were calculated on both the TLS and the 16PF, debriefing sheets (stating "you got a hit" or "you did not get a hit," transliminality scores, and instructions on how to interpret the 16PF results) were issued to all participants. Included with each debriefing sheet was a copy of the participant's consent form and copies of hexagram readings and changing line readings (if any). Participants were thanked for their participation in the experiment.

5.7 Results

5.7.1 Use of, and Attitude toward, the I Ching

Of the total sample of 107 participants, only 16 (15%) had used the *I Ching* prior to the experiment. (Note that none of these previous users had gained that experience as a participant in Storm & Thalbourne's, 1998-1999, initial run of the *I Ching* experiment.)

Eighty participants (a surprising 75% of the sample) believed that it was "possible" for other people in the sample to achieve a hexagram outcome matching 1 of the 16 designated on the descriptor form (i.e., to get a hit) by paranormal means, but only 30 participants (28%) believed in their "own [paranormal] abilities" to get a hit.

5.7.2 Changing Lines

Figure 5.6 shows the distribution of scores on the changing lines measure. As can be deduced from the figure, there was a subgroup of 79 participants (74%) within the total sample of 107 participants who generated second hexagrams, since only those participants threw three-of-a-kind at least once, and thus generated changing lines (see bars '1', '2', '3', and '4' in Figure 5.6). Thus there were 28 participants (26%) who did not generate changing lines (see bar '0' in Figure 5.6).

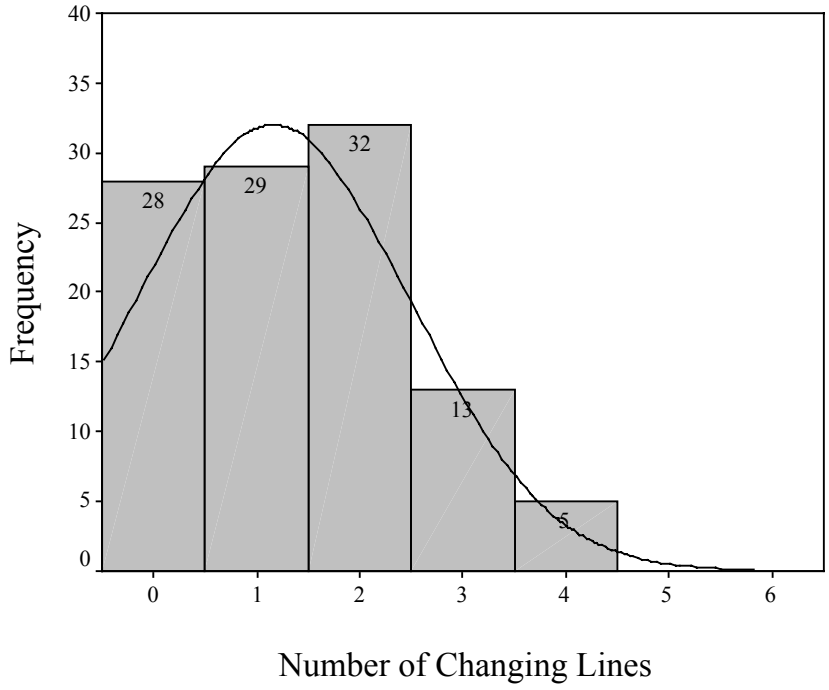

Figure 5.6. Distribution of changing lines (N = 107).

The average number of changing lines was 1.42 (SD = 1.14), where MCE = 1.5. This performance was not significantly below average. There is a nonsignificant positive

skew[53] of 0.393 (*SE* = 0.234), which only suggests that the task of generating more than four changing lines might be difficult.

5.7.3 Transliminality

As was the case in Storm and Thalbourne's (1998-1999) initial study, the TLS was shown to have a satisfactorily high reliability coefficient, expressed as a Cronbach's alpha of .82. This result was only slightly smaller than the previous alpha of .86 reported by Storm and Thalbourne (p. 108).

The theoretical range of the TLS is 0 to 29, but the observed range was 4 to 28 (*N* = 107). The mean transliminality score was 16.96 (*SD* = 5.43), which was not significantly larger than Storm and Thalbourne's (1998-1999, p. 108) result of 16.38 (*SD* = 6.08). The distribution of transliminality scores is shown in Figure 5.7.

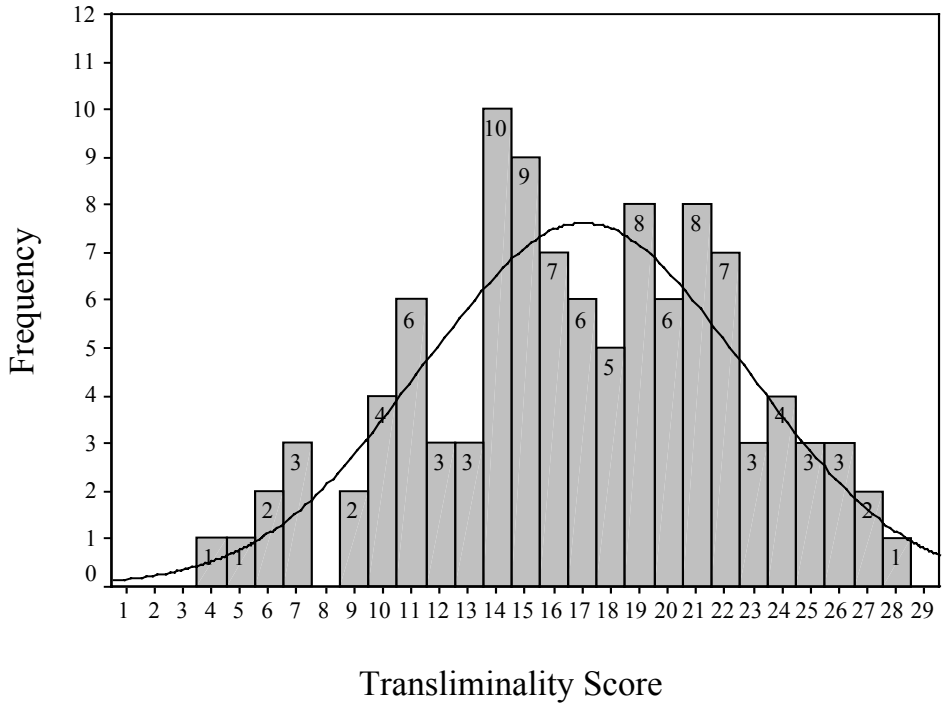

Figure 5.7. Distribution of transliminality scores (*N* = 107).

[53] Dividing the skew value (.393) by its *SE* (.234) yields 1.68 which is less than 1.96, and therefore not significant.

Figure 5.7 shows a slight negative skew of –0.173 (*SE* = 0.234; Kurtosis = –0.467). Storm (1998, p. 29) reported a slight positive skew (0.094, *SE* = 0.250), but neither of these two skews departs significantly from chance. Consequently, any apparent variability of transliminality scores between the two samples can be taken as chance fluctuation.

5.7.4 The 16PF

The STEN scores on the 16PF are standardized scores—the theoretical range of scores on all factors, therefore, is 1 to 10. Seventeen of 21 factors (including global factors) were on or between the standardized range of 4 to 7 (Russell & Karol, 1994, p. 19). The exceptions were Reasoning (Factor B) with a mean STEN score of 8.23; Rule-Consciousness (Factor G) with a mean STEN score of 3.42; Abstractedness (Factor M) with a mean STEN score of 7.80; and Openness to Change (Factor Q_1) with a mean STEN score of 7.61.

Using the single-sample *t*-test, Reasoning (Factor B) scores were significantly higher than the norm extreme (i.e., 7) for that factor, $t(106) = 2.16$, $p = .033$, two-tailed. The reason for this difference may have to do with the fact that the majority of participants were university students with above average IQ scores.

The other three factors—Factor G (Rule-Consciousness), Factor M (Abstractedness), and Factor Q_1 (Openness to Change)—were also significantly different from the relevant norm extremes. Factor G was significantly below the norm extreme (i.e., 4), which suggests that the sample was comprised of participants who tended to be significantly more 'expedient' and 'nonconforming' than can be explained by chance alone, $t(106) = –4.14$, $p < .001$, two-tailed. This significant deviation might be explained by the fact that the sample was over-represented by adolescents.

Factor M was significantly above the norm extreme (i.e., 7), which suggests that participants tended to be significantly more 'abstracted', 'imaginative' and ' idea-oriented' than can be explained by chance alone, $t(106) = 5.13$, $p < .001$, two-tailed. By way of conjecture, this factor may typify the average participant in experiments that require an interest (belief?) in the paranormal. Unconventional belief might require some capacity for abstracted (lateral-minded) and imaginative thought. Note that the 'sheep' question (Question 2) on the Hexagram Descriptor Form (see Appendix C) correlates with Factor M, $r(105) = 0.20$, $p = .043$.

Factor Q_1 was also significantly above the norm extreme (i.e., 7), which suggests that participants tended to be significantly more 'open to change' and 'experimenting', $t(106) = 3.78$, $p < .001$, two-tailed. This factor did not correlate with any belief questions

on the Hexagram Descriptor Form. Note, however, that the typical volunteer for experiments may likely be the 'experimenting' type open to new experiences (i.e., one who is 'open to 'change').

5.7.5 Parapsychological Hypotheses

Hypothesis 1: Hexagram-hitting is at a rate greater than MCE *when expressed as a proportion of hits (*P_{MCE} *= 0.25), and as an effect size* π *(*π_{MCE} *= 0.50).* The observed proportion of hits was 0.35. The binomial test was significant (p = .015). Hypothesis 1 was therefore supported. This proportion translates as an effect size π of 0.61. A direct significance test of π can also be calculated because the sample is large (N = 107). Rosenthal and Rubin's (1989, p. 334) Formula (4) was used to calculate a z score of 2.27 (p = .012, one-tailed). Thus, the effect size π of 0.61 was significantly higher than the *MCE* effect size of 0.50. Hypothesis 1 was again supported. The corresponding findings from Storm and Thalbourne (1998-1999, pp. 108–109) and Storm [and Thalbourne] (2001a) were thus replicated.

Hypothesis 2: Hexagram-hitting for second-hexagram hitters who were also first hexagram hitters is above MCE. Seventy-nine participants generated changing lines and therefore had a second chance of achieving a hit because changing lines change the first hexagram into one of the 63 remaining hexagrams. Participants (n_1 = 27) who threw changing lines, and got a hit on their first hexagram, have 15 chances out of 63 of getting a hit on their second hexagram (P_{MCE} = .238).

The proportion of second hexagram hits in n_1 was only 0.22; a 22% hit rate, p = .513. Thus (as expected), the Null hypothesis failed to be rejected (see *5.1.2* for reasons for this expected outcome). This result is consistent with the hypothesis that a pro attitude for a specific target is necessary to bring about a paranormal effect.

Hypothesis 3: Hexagram-hitting for second-hexagram hitters who were first hexagram missers is above MCE. Participants (n_2 = 52) who also threw changing lines, but did not get a hit on their first hexagram, have 16 chances out of 63 of getting a hit on the second hexagram (P_{MCE} = 0.254). The proportion of second hexagram hits in n_2 was only 0.29; a 29% hit rate, p = .340. Again (as expected), the Null hypothesis failed to be rejected (see *5.1.2* for reasons for this expected outcome). This result also is consistent with the hypothesis that a pro attitude for a specific target is necessary to bring about a paranormal effect.

Hypothesis 4: Hexagram-hitting for Sample 1999 is lower than hexagram-hitting for Sample 1998. Sample 1999 actually scored higher on average (35% hit-rate) than Sample 1998 (32% hit-rate), which is not in the direction hypothesized. The difference was not significant, $\chi^2(1, N = 200) = 0.120$, $p = .729$; see Appendix N for calculations). Thus, evidence was not found that psi is a mediational process that depends on information processing (see *5.1.3*).

Hypothesis 5: There is a positive correlation between transliminality and hexagram-hitting. Storm and Thalbourne (1998-1999, p. 108) found a positive and significant relation between these two variables, $r(91) = 0.27$, $p = .010$. In the 1999 study the correlation was positive, but it was not significant, $r(105) = 0.003$, $p = .489$. Hypothesis 5 was not supported.

Hypothesis 6: There is a positive correlation between transliminality and number of changing lines. Storm and Thalbourne (1998-1999, p. 110) found a positive relation between these two variables, $r(91) = 0.19$, $p = .062$. This result was not replicated in the 1999 study: An even weaker, negative, and nonsignificant correlation was found, $r(105) = –0.01$, $p = .476$. Hypothesis 6 was not supported.

Hypothesis 7: Hexagram-hitting correlates positively with Factor F (Liveliness), Factor H (Social Boldness), Factor EX (Extraversion), and Factor IN (Independence), and negatively with Factor Q_2 (Self-Reliance) and Factor Q_4 (Tension) of the 16PF. There were no significant correlations between hitting and any of the six factors on the 16PF. Thus, Hypothesis 7 was neither generally nor specifically supported. Storm and Thalbourne's (1998-1999, p. 110) previous findings, therefore, were not replicated.

Hypothesis 8: Number of changing lines correlates positively with Factor M (Abstractedness) and Factor Q_2 (Self-Reliance), and negatively with Factor A (Warmth), Factor C (Emotional Stability), and Factor EX (Extraversion) of the 16PF. Number of changing lines did not correlate significantly with any of the five hypothesized 16PF factors. Thus, Hypothesis 8 was not supported. Therefore, Storm & Thalbourne's (1998-1999, p. 111) previous findings did not replicate.

Hypothesis 9: Number of changing lines correlates positively with answers to Question 2 ("Possibility"). A significant correlation was found, $r(105) = 0.16$, $p = .05$, one-tailed. This result compares favourably with Storm and Thalbourne's (1998-1999, p. 111) statistically significant result, $r(93) = 0.21$, $p = .047$. Thus, participants who believed in the

possibility that paranormal phenomena might be involved in the *I Ching* process tended to generate significantly more changing lines than those who did not believe. (Note that this measure of paranormal belief does not conform to Schmeidler's, 1945, original paranormal belief question, which strictly refers to a generic belief in psi.)

5.7.6 Psychological Hypotheses

Hypothesis 10: Transliminality correlates positively with answers to Question 2 (the "sheep" question) and with answers to Question 3 (the "super-sheep" question). Transliminality did correlate positively with Question 2, $r(105) = 0.41$, $p < .001$. Storm and Thalbourne's (1998-1999, pp. 110-111) post hoc finding was thus replicated: Highly transliminal participants tended to believe in the possibility that a paranormal process might take place during the *I Ching* process.

Transliminality also correlated positively with Question 3, $r(105) = 0.35$, $p < .001$. Again, Storm and Thalbourne's (1998-1999, pp. 110-111) post hoc finding was replicated. Highly transliminal participants tended to believe that their own paranormal abilities might contribute to, or be responsible for, success in the paranormal component of the *I Ching* process.

Hypothesis 11: Answers to Question 2 (the sheep question) and Question 3 (the super-sheep question) correlate positively with each other. Answers to Question 2 (the sheep question) and Question 3 (the super-sheep question) correlated significantly and positively with each other, $r(105) = 0.32$, $p < .001$. Hypothesis 11 was therefore supported, and Storm and Thalbourne's (1998-1999, pp. 114-115) post hoc finding was replicated. Participants who believed in their own psi ability believed it was possible for other participants in the sample to have psi ability.

Hypothesis 12: Transliminality scores correlate positively with Factor A (Warmth) and Factor M (Abstractedness) and negatively with Factor G (Rule-Consciousness), Factor TM (Tough-Mindedness), and Factor SC (Self-Control) of the 16PF. Table 5.5 shows the results for these correlations. Transliminality did not correlate significantly or positively with Factor A, but did correlate significantly and positively with Factor M, suggesting that highly transliminal participants tended to be idea-oriented. Transliminality also correlated significantly and negatively with Factor G and Factor TM. Highly transliminal participants, therefore, tended not to be rule-conscious (Factor G) or tough-minded (Factor TM).

As can also be seen from Table 5.5, this last correlation of transliminality with Factor SC just missed significance. Nevertheless, it can be regarded as "tell[ing] about the

same story" (Rosenthal & Rubin, 1979, p. 1165) as Storm and Thalbourne's (1998-1999, p. 111) statistically significant result (see Table 5.2), since Rosenthal and Rubin (1979) would say: "Both results are in the same direction and the studies are of similar size" (p. 1165). Thus, highly transliminal participants tended to be lacking in self-control (Factor SC).

Table 5.5
Correlations between Transliminality and 16PF Factors

Variable	*r*	*p*
Factor M (Abstractedness)	0.23	.008
Factor G (Rule-Consciousness)	–0.24	.006
Factor TM (Tough-Mindedness)	–0.17	.042
Factor SC (Self-Control)	–0.15	.062

Notes: $N = 107$; $df = 105$; p values are one-tailed

Given the failure of the transliminality/Factor A correlation to replicate, it might be claimed that Hypothesis 12 has been only partially confirmed, and therefore that Storm and Thalbourne's (1998-1999, pp. 110-111) findings were only partially replicated. However, they (Storm & Thalbourne, p. 115) had already acknowledged that multiple analyses could produce chance results. The failed transliminality/Factor A correlation may be such a case, and recent research supports this conclusion: Lange, Thalbourne, Houran, and Storm (2000) performed a top-down purification procedure on the Transliminality Scale, thus producing the Revised Transliminality Scale. This procedure identifies and removes any "differential item functioning" (i.e., item bias). They rejected 12 of 29 questions that were biased in relation to sex and age.

When a Pearson *r* test was run again on Sample 1998 ($N = 93$) using the Revised Transliminality Scale, Factor A dropped out. It is highly likely that the transliminality-Factor A correlate of the 1998 sample was the one significant correlation that could be expected by chance alone. (See Storm & Thalbourne, 1998-1999, p. 115, where there would be an expected 1.1 significant correlations by chance alone given that the Null hypothesis is true, and that the 16PF variables are independent.)

5.7.7 Belief in the Possibility of Psi as a Necessary Condition

Pursuant to the protocol for testing for necessary conditions, the significant result found in testing Hypothesis 9 warrants a further (i.e., post hoc) test on the belief variable 'Possibility'. Those who answered 'Yes' to this question (n = 80), were tested separately on number of changing lines, but the mean number of changing lines, although above *MCE* (1.53, where *MCE* = 1.50), was not significantly above chance. Therefore answering 'Yes' to this 'sheep' question was not sufficient or necessary in bringing about paranormal performance on changing lines generation.

5.7.8 Success rates

In the planned analyses for this study, 27 specific statistical tests were performed to test 12 hypotheses. Of these tests, 9 produced significant results, and 2 were not significant as expected (i.e., 41% of all tests were 'successful' in the sense that were supportive of the theory of psychopraxia). More specifically, for the parapsychological hypotheses, there were 5 'successful' tests out of 19 (26%), and for the psychological hypotheses, there were 7 successful tests out of 8 (88%). All these percentages are greater than 5% and therefore are unlikely to be explained by chance alone. (Note that these percentages may be inflated due to non-independent correlations.)

5.8 Discussion

The *I Ching* is an ancient Chinese form of divination based on the principle of duality, the yin and the yang. From a simple yin/yang polarity are derived the 64 hexagrams. Hexagram symbols carry with them certain meanings, which come in the form of readings. The traditional view is that in the uncertain flux of world events is the possibility of knowing that in the present are the seeds for solutions in the future, and that these solutions can be derived through divination. Replicable evidence was found in this study that suggested that an anomalous process underlies the *I Ching* process. Given a significant 35% hit-rate (where *MCE* = 25%), the readings do not appear to be generated purely by chance. Thus there might be some validity in the traditional view.

5.8.1 Summary of Results

The *I Ching* experiment revealed itself to be complex in procedure, and a challenge to participants. There were no less than two parapsychological tasks to be performed concurrently (hexagram-hitting and generation of changing lines), as well as three

questionnaires, including the time-consuming 16PF. Nevertheless, most aspects of each of the five aims (see **5.1**) were supported. The aims of this study will be given next in terms of supported hypotheses.

5.8.1.1 Replication: Results relevant to this aim are as follows:

(a) Hypothesis 1: Hexagram-hitting was achieved by 35% ($\pi = 0.62$) of the sample where only 25% would be expected by chance ($\pi_{MCE} = 0.50$).

(b) Hypothesis 9: Number of changing lines correlated positively and significantly with Question 2 (the 'sheep' question).

(c) Hypothesis 10 (Part I): Transliminality correlated positively and significantly with Question 2.

(d) Hypothesis 10 (Part II): Transliminality correlated positively and significantly with Question 3.

(e) Hypothesis 11: The sheep question (Question 2) and the super-sheep question (Question 3) correlated positively and significantly.

(f) Hypothesis 12: Transliminality correlated positively and significantly with 4 out of 5 16PF factors.

Hypotheses not confirmed were as follows:

(a) Hypothesis 5: Transliminality did not correlate positively and significantly with hexagram-hitting.

(b) Hypothesis 6: Transliminality did not correlate positively and significantly with changing lines.

(c) Hypothesis 7: Hexagram-hitting did not correlate positively and significantly with 6 factors on the 16PF.

(d) Hypothesis 8: Number of changing lines did not correlate positively and significantly with 5 16PF factors.

Six of ten hypotheses were confirmed. Generally, Aim One was achieved, where it was proposed that Storm and Thalbourne's (1998-1999) findings would be replicated.

However, four hypotheses were not confirmed so that replication was demonstrated in the 1999 study only to a limited degree. Furthermore, the four failed hypotheses happen to be parapsychological, so that these failures confirm, yet again, the well-recognised fact that psi effects are difficult to replicate, although the successfully replicated psi effects (see Hypotheses 1 and 9) help to quell some of this doubt.

5.8.1.2 The ESP-PK dichotomy: That the traditional ESP-PK division can be unworkable in practical parapsychology has been discussed in Chapter 4 (see also *5.1.1*). Given that a psi process took place in the *I Ching* experiment, the paranormal task would be conventionally described as involving a physical process (coin-throwing) and/or a mental process (predicting or precognising the hexagram outcome). Thus, the traditional view is that ESP *or* PK may have taken place, or indeed, both ESP *and* PK may have taken place simultaneously.

Theorists might like to argue the point over whether the coins are influenced by the user. If that is the case, the theorist might decide in favour of PK, or blind-PK. The blind-PK explanation is proposed because naïve participants in Sample 1999 performed better than non-naïve participants: Naïve (n = 91), P = 35%, p = .017; non-naïve (n = 16), P = 31%, p = .370). However, results so far do not indicate that blind-PK is necessarily involved in the *I Ching* process (see Storm & Thalbourne, 1998-1999, p. 111, where non-naïve participants performed significantly above chance, but naïve participants did not).

Thus, a middle ground position offers the theorist 'real time'-ESP (clairvoyance) in conjunction with PK (causing the coins to fall in the required way), and this process might involve blind-PK. The blind-PK conclusion raises the possibility of psi mediation through information acquisition, which implies extra steps for novices (i.e., the non-naïve participant may already know how to score, whereas the naïve participant does not).

Alternatively, the theorist might decide in favour of ESP (viz., precognition). No doubt, the debate will continue. Nevertheless, the argument that the ESP-PK division can be unworkable in practice was supported by the experimental results of the present study.

5.8.1.3 The pro attitude: The results of three hypotheses relevant to this aim are presented:

(a) Hypothesis 1: Significant hexagram-hitting was achieved ($P_{obs.}$ = 0.35, where P_{test} = 0.25).

(b) Hypothesis 2: Significant hexagram-hitting was not achieved for first hexagram hitters who were also second-hexagram hitters.

(c) Hypothesis 3: Significant hexagram-hitting was not achieved for first hexagram missers who were second-hexagram hitters.

Hypotheses 1, 2 and 3, were 'successful' in terms of the theory of psychopraxia. Thus, the 'pro attitudinal' objective of Aim Three was achieved. These results support Thalbourne's (2000a) thesis that exo-psychopraxia is a paranormal goal that follows an appropriate pro attitude (given the necessary and sufficient conditions). A sufficient number of participants selected and achieved *only* their designated targets, and nothing other than their designated targets. Thalbourne's (2000a, pp. 65-66) claim that paranormal phenomena may be better described as the fulfillment of a goal as the result of a pro attitude was shown to be feasible in this study.

These results need further analysis. Two general statements were presented above (see *5.1.2*): "*Given sufficient other conditions, if there is significant hitting on first hexagrams, then a sufficient number of pro attitudes were present; therefore, if there is no significant hitting on second hexagrams, then a sufficient number of pro attitudes were not present*" (p. 77). Of course, the second statement could be false if there *were* a sufficient number of pro attitudes for second hexagrams, but insufficient *other* conditions. Unlike a repeated measures procedure, however, where conditions can change between (say) two runs, the coin throwing procedure generates second hexagrams *at the same time* the first hexagrams are generated—a change in conditions would be unlikely.

At this stage, however, it should be asked: "Whose pro attitude is being served, that of the participants, or the principal experimenter, or other interested parties, such as supervisor(s) or other parapsychologists"? On the one hand, it is acknowledged that the experimenter (L.S.) and at least one of his supervisors (M.A.T.) both score highly on the Transliminality Scale and measures of belief in the paranormal, and Thalbourne would argue that scoring highly on these scales are conditions conducive to exo-psychopraxia.

On the other hand, at least one theory (i.e., Jung's, 1960, theory of synchronicity) suggests that the participants in the experiment may have the dominating pro attitudes because the physical outcomes are meaningfully related to the mental (cognitive/emotional) states of those participants. Recall that each participant's choices of possible outcome hexagrams were limited to how he/she *felt* "lately, or right now," which could be "an all-pervading mood, feeling, emotion, image, or thought which has dominated [his/her] awareness for some weeks, or only today, or only in the last few moments" (from

the front page of the *I Ching* Hexagram Descriptor Form). (See Storm, 1999, who points to similarities between the paranormal and synchronicity.)

5.8.1.4 Mediational psychopraxia: The result of Hypothesis 4 is relevant to this aim: Hexagram-hitting in Sample 1999 ($P = 0.35$) was not lower than hitting in Sample 1998 ($P = 0.32$). The objective of Aim Four was not achieved—the paranormal *I Ching* effect was not shown to be a mediational process (see *5.1.3*), since a significantly lower hexagram hit-rate was not found for Sample 1999. Like conventional noncybernetic theories of goal-oriented psi (e.g., synchronicity and PMIR; see *4.2.3* and *4.2.4*, respectively), which dispense with the idea that information is needed to assist the psi process, Thalbourne claims that the theory of psychopraxia makes allowances for the fact that information processing may still take place (see *5.1.3*), though there is no evidence for this claim insofar as exo-psychopraxia is concerned. In fact, the hit-rate was higher, which is suggestive of goal-orientation, though there is no evidence for that either.

5.8.1.5 Necessary conditions: Four of five hypotheses were not supported (Hypotheses 5, 6, 7 & 8). Generally speaking, Aim Four was not achieved. However, at this early stage, it may be too soon to conclude that these failed hypotheses indicate that transliminality and some 16PF factors were not necessary for, or at least conducive to, exo-psychopraxia. (The failures of these hypotheses to be confirmed, and indeed the appropriateness of these hypotheses to this study—at least in their current form—will be taken up in Chapter 6.)

Hypothesis 9 was supported. There is a relationship between number of changing lines and the belief ('sheep-goat') measure used in the descriptor form (viz., 'possibility'—i.e., belief in the possibility of getting a hit through the psi process). This correlation is also a replication. However, a post hoc analysis (see *5.7.7*) revealed that this form of belief in the paranormal process was not necessary or sufficient in bringing about exo-psychopraxia.

5.9 Conclusion

The overall results of this study send out a few mixed signals about exo-psychopraxia. Definite support for the theory came in the way of highlighting the ESP-PK dichotomy and the pro attitude (since psi effects were replicated). But this study failed in

its aim to replicate the 1998 results that transliminality[54] correlates with hexagram-hitting, and may therefore be a condition necessary for, or at least conducive to, exo-psychopraxia. One replicated correlation—changing lines with 'possibility' (Hypothesis 9)—gave good evidence that at least one form of belief might have been necessary for exo-psychopraxia, but further testing failed to show it. However, generally, it must be conceded that although psychopraxia may be a more parsimonious description of the events involved in the present experiment, the study considered as a single research project throws little or no light on the necessary mediating conditions that must theoretically be involved. Specifically, knowledge of the level of transliminality and of the 16PF factors was of little use in predicting hitting or number of changing lines.

The next chapter contains the results of various analyses that were performed on the sample used in this study (N = 107), as well as the 1998 sample (N = 93). Post hoc analyses were undertaken as necessary steps towards explaining the failures of Hypotheses 5 and 6 to be confirmed, and the appropriateness of Hypotheses 7 and 8 in their current form to this study. Further analyses were performed in order to unify two disparate data sets, thus allowing the calculation of better estimates of the relevant population parameters. Multiple regression analyses and median-split analyses were also performed in the attempt to replicate some additional post hoc findings in the 1998 study (see Storm & Thalbourne, 1998-1999, pp. 112-114).

[54] Note that replication of the transliminality-belief and transliminality-16PF correlates confirmed the relevance of transliminality for psychology (see Hypotheses 8, 9 and 10 in *5.7.6*)

CHAPTER 6
THE *I CHING* EXPERIMENTS (ADDITIONAL ANALYSES)[55]

One must have a far-reaching psychological understanding in order to enjoy the I Ching *with advantage.*
C. G. Jung (1934, cited in Adler, 1973, p. 139)

Vast numbers of coincidences arise from hidden causes that are never discovered.
Persi Diaconis and Frederick Mosteller (1989, p. 859)

6.1 Reasons for Additional Analyses

The *I Ching* experiment featured in Chapter 5 was an attempt to replicate the significant results of the Storm and Thalbourne (1998-1999) study. Thus, there are now two samples ('Sample 1998' and 'Sample 1999'). In this chapter, the results of additional analyses of these two samples are presented with the objective of achieving the following aims:

1. To determine possible reasons for the failures of four hypotheses (i.e., Hypotheses 5, 6, 7, & 8—see *5.7.5*; bivariate correlation analyses were performed).

2. To pool the two samples, where feasible, thus yielding the combined sample (N = 200), from which more accurate estimates of the relevant population parameters can be calculated.

3. To replicate the regression results in Storm and Thalbourne's (1998-1999, p. 112) study using multiple regression analysis (MRA) on Sample 1999 data. (An MRA was also performed on the combined samples.)

4. To ascertain the success rates of hexagram hitting for 'high' scoring participants on the transliminality scale in Sample 1999, and the combined samples, as originally performed on Sample 1998 by Storm and Thalbourne (1998-1999, p. 114) using

[55] This chapter was published as a refereed article (see Appendix AF).

median-split analyses (MSA). Transliminality was investigated as a possible necessary condition for bringing about paranormal effects. The 16PF Factors F, H, Q_4, EX and IN were also investigated as possible necessary conditions.

6.2 Bivariate Correlation Analyses

In this section, the results of bivariate correlation analyses are presented, the aim being to determine possible reasons why Hypotheses 5, 6, 7 and 8 failed to confirm. For convenience, the results of three performance comparisons between Sample 1998 and Sample 1999 on transliminality, hitting, and number of changing lines are given in advance of the Performance Comparison section (see **6.3**), so that the combined sample can also be subjected to a similar analysis in the following section.

6.2.1 Transliminality and Hitting

In Sample 1999, and in contrast with Sample 1998, a significant correlation was not found between hitting and transliminality. Although the correlation in Sample 1998 was significant, Storm and Thalbourne (1998-1999, p. 114) tested (post hoc) for a decline effect in this correlation across the sample because there seemed to be evidence, to the naked eye, of a decline in effect size over time (viz., high transliminals were not hitting as often as might be expected). Sample 1998 was subdivided into four groups of 20 and one residual group of 13. These groups were then ranked first to fifth (first being the earliest tested group and fifth being the last). In the event, using Spearman's test, a nonsignificant decline over the course of the experiment in size and direction of the five transliminality/hitting correlations was found, $r_s(3) = -0.50$, $p = .196$, one-tailed. The failure of this decline to reach significance may have come about as a result of the small number of groups (viz., five).

If there was a similar decline in Sample 1999, then this decline might also have severe effects on the size and significance of the transliminality/hitting correlation, even to the degree that the failure of the correlation to replicate might be explained by that decline. Thus, Sample 1999 was broken down into four groups of 20 participants and one residual group of 27 participants. The resultant five correlations between hitting and transliminality were also ranked chronologically (first through fifth). A weak and nonsignificant decline was indicated, $r_s(3) = -0.10$, $p = .436$, one-tailed.

The combined sample ($N = 200$) was also investigated for decline effects. The combined sample was formed only after performance comparisons between Sample 1998

and Sample 1999 were made on the relevant variables. Transliminality scores did not differ significantly between samples, $t(198) = -0.71$, $p = .479$, two-tailed ($\omega^2 = 0$). Note that a significant t value implies the existence of an association, but the *estimate* of effect size 'omega-squared' tells how strong that association is. Hays (1963) recommends that omega-squared (ω^2) accompany the result of an independent-samples t test. When $t \leq 1$, the estimate of $\omega^2 = 0$; when $t > 1$, estimated $\omega^2 = (t^2 - 1)/(t^2 + N_1 + N_2 - 1)$, where N_j is the size of each sample. Where omega-squared is less than 0.09, the association is regarded as "functionally" unimportant because there is little predictive power in that association. Consequently, the samples will be combined on the basis that they may be regarded as coming from the same population.

Hexagram hit rates did not differ significantly between samples either, $\chi^2(1, N = 200) = 0.12$, $p = .729$, two-sided ($\phi = 0.03$). A phi-squared value (i.e., the coefficient of determination equivalent to r^2) that falls below 0.09 will also normally be regarded as unimportant, but this criterion does not apply to the categories of sex and academic affiliation and, therefore, is not given for the relevant chi-square tests that follow later (see *6.2.3* & *6.2.4*).

When the individual group correlations of the combined sample are plotted (see Figure 6.1), a decline in the transliminality/hitting correlation was sought at the 8th group before the 9th and 10th groups were formed, which turned out to be significant, $r_s(6) = -0.88$, $p = .004$, one-tailed. The decline was expected to continue, but the remaining two groups when collected went against the linear trend, to produce in fact an overall significant quadratic trend, $r_s(8) = 0.57$, $p = .044$, one-tailed. A U-shaped trend might appear if the experimenter was having an effect on the correlations and enthusiasm and motivation picked up toward the end of the experiment after "burn out" or "heightened anxiety to keep up the promising results" (see Broughton & Alexander, 1997, p. 223). (Note that the two significant correlations in Figure 6.1 are probably nonsignificant when adjusted for multiple analyses.)

In conclusion, neither Sample 1998, nor Sample 1999, nor the combined sample produced significant declines. However, the overall transliminality/hitting correlation was significant for the combined sample, $r(198) = 0.12$, $p = .040$, one-tailed, but this very weak correlation might well have been stronger had it not been for the position effect.[56]

[56] One commentator (S. Ertel, personal communication, September 9, 2000) advised the author to consider the possibility that the ostensible change in Pearson r values across groups may have been attributable to declines in the variance on the transliminality measures for those groups, but this rival hypothesis was not

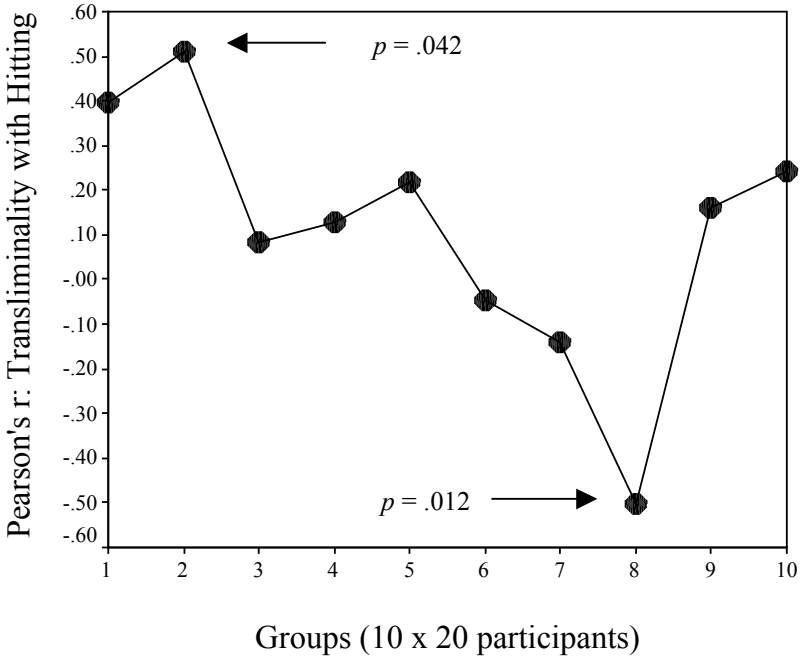

Figure 6.1. The combined sample ($N = 200$): Pearson correlations between transliminality and hitting for each of 10 consecutive groups of 20 participants.

6.2.2 The Transliminality/Number of Changing Lines Correlation

The transliminality/changing lines correlation failed to reach significance in the 1999 sample. Suspecting that decline effects may have played a part in this failure to replicate, the same procedure outlined in the section above was followed. Decline effects were again hypothesized for Sample 1998, Sample 1999, and the combined sample.

For Sample 1998, there was a nonsignificant *incline*, $r_s(3) = 0.70$, $p = .094$, one-tailed. For Sample 1999, the rank-order correlation was also not significant, $r_s(3) = 0.00$, $p = .500$, one-tailed.

Nor was there a significant decline for the combined sample,[57] $r_s(8) = -0.176$, $p = .313$, one-tailed. A test for quadratic trend yielded entirely null results. Thus, no evidence was found that an ostensible position effect might account for the failure to replicate the transliminality/changing lines correlation in Sample 1999.

---

confirmed. There were no significant declines in variances for transliminality across groups, nor were there for hitting and number of changing lines.

[57] Number of changing lines did not differ significantly between samples, $t(198) = 0.56$, $p = .579$, two-tailed ($\omega^2 = 0$). The samples were therefore combined for this analysis (see *6.3.5* for the result of the transliminality/changing lines correlation for the combined sample).

6.2.3 The Hitting/16PF Correlations

As reported in Chapter 5, not one of the six hitting-16PF correlations that were significant in Sample 1998 was significant in Sample 1999 (Hypothesis 7, see *5.7.5*). It was considered highly likely that the two samples were not homogeneous on the following six relevant 16PF factors: Factor F (Liveliness), Factor H (Social Boldness), Factor Q_2 (Self-Reliance), Factor Q_4 (Tension), Factor EX (Extraversion), and Factor IN (Independence). This assumption came from an earlier observation (see Storm & Thalbourne, 1998-1999, p. 116) that there was a preponderance of female and psychology students in Sample 1998. Conducting tests on one sample with a view to replicating the results of an earlier sample is unwarranted if there happen to be fundamental differences between the two samples due to biases of one form or another.[58] Therefore, both samples were investigated for biases in sex composition (because sampling procedures in the 1998 study were different from those of the 1999 study[59]), and for the same reasons, it was decided that the number of participants studying psychology should also be investigated.

There was a drop in female participants from 72% to 54% across the two samples (i.e., 9 fewer female participants in Sample 1999), which proved to be significant using the Pearson chi-square test, $\chi^2(1, N = 200) = 6.75$, $p = .009$ (two-sided). Another bias exists in the form of a significant drop in the number of psychology students: 50.5% in 1998 to 29% in 1999 (i.e., 16 fewer psychology students in Sample 1999), $\chi^2(1, N = 200) = 9.73$, $p = .002$ (two-sided). One final chi-square test was conducted, this time comparing four nonoverlapping groups (female psychology students, female nonpsychology participants, male psychology students, and male nonpsychology participants). The Fisher Exact Test was significant ($p < .001$, two-sided). It was then hypothesized that the original hitting-16PF correlations may have been specific to groups only and not the whole of Sample 1998. Pearson *r* tests were conducted on each of the four nonoverlapping groups in Sample 1998 between hitting and the six relevant 16PF factors.

---

[58] The literature on sex differences on paranormal belief is quite substantial, but for studies on sex differences in ESP performance, see Palmer and Johnson (1991), Peretti (1971), and Rao and Kanthamani (1981, 1983). See Parker, Frederiksen, and Johansson (1997) and Parker, Grams, and Pettersson (1998) for studies on ESP performance differences between psychology students and participants recruited from New Age groups or groups with paranormal experiences.

[59] In the 1999 study, volunteers were sought from all disciplines of Adelaide University, whereas the 1998 study comprised mainly volunteers from the Departments of Psychology, Asian Studies, Architecture, and Computer Science.

Prior to running the Pearson *r* tests, five one-way analyses of variance (ANOVAs) were run to determine possible differences between groups on mean scores for the five 16PF factors. Only the mean scores on Factor EX scores were heterogeneous across groups, but a post hoc Tukey's test showed that there was only a marginally significant difference on Factor EX scores between male nonpsychology participants and female psychology students (p = .087). Consequently, all the relevant Pearson *r* tests were run for the four groups. Only one of the four groups produced significant correlations, and that was the female psychology students (n = 39; see Table 6.1). They produced five of the six significant correlations originally found for the whole of Sample 1998; the hitting/Factor IN (Independence) correlation was not significant.

Table 6.1
Correlations Between Hitting and 16PF Factors for Female Psychology Students 1998

Factor	*r*	*p*
F (Liveliness)	0.58	< .001
H (Social Boldness)	0.50	.001
Q_2 (Self-Reliance)	–0.40	.012
Q_4 (Tension)	–0.48	.002
EX (Extraversion)	0.48	.002

Notes: p values are one-tailed; n = 39; df = 37

It may be possible that female psychology students disproportionately contributed to those same five correlations originally reported for Sample 1998. In support of this possibility, the three remaining groups, when recombined (n = 54; which excludes the female psychology students), produced only one significant correlation: hitting-Factor H (Social Boldness), $r(52) = 0.28$, p = .043, two-tailed.

To test for replication of the five significant hitting/16PF correlations given in Table 6.1 for female psychology students only, five Pearson *r* tests were performed on Sample 1999 (female psychology students only; n = 24). (Note that, prior to the Pearson *r* tests, the two groups of female psychology students, Sample 1998 and Sample 1999, were compared using independent-samples *t* tests. There were no significant differences between respective groups on Factors F, H, Q_2, Q_4, and EX.) Only the hitting/Factor H correlation among female psychology students replicated, $r(22) = 0.44$, p = .016, one-tailed. For the

remaining participants (n = 83), which excludes the female psychology students, the hitting-Factor H correlation was not significant, $r(81) = -0.08$, $p = .224$, one-tailed.

Given these results, sex and academic affiliation alone do not explain the lack of replication of four correlations that were found in 1998 (viz., hitting with Factors F, Q_2, Q_4, and EX). Nevertheless, one correlation (hitting/Factor H) replicated for female psychology students, so that, once again, those participants appeared to have disproportionately contributed to that correlation.

6.2.4 Correlations of Number of Changing Lines with 16PF Factors

Given the above rationale regarding sex and academic affiliation, five one-way ANOVAs were run to determine possible differences between the same nonoverlapping groups on mean scores for the five 16PF factors. Factor M scores were heterogeneous across groups. A post hoc Tukey's test showed that there was a significant difference on Factor M scores between the female psychology students and the male nonpsychology participants (p = .040; note that Factor EX scores were tested in the previous subsection above). Consequently, Pearson r tests were run for the four groups using only four factors: Factors A, C, Q_2, and EX. For Sample 1998, the four different groups produced a number of significant correlations (see Table 6.2).

Table 6.2
Correlations Between Changing Lines and 16PF Factors: Four Groups (Sample 1998)

Factor	Male psych. (1998: n = 8)	Male nonpsych. (1998: n = 18)	Female psych. (1998: n = 39)	Female nonpsych. (1998: n = 28)
A (Warmth)	–0.67*	–0.19	–0.19	–0.22
C (Emotional Stability)	0.34	–0.20	–0.30*	–0.34*
Q_2 (Self-Reliance)	0.32	0.38	0.33*	0.22
EX (Extraversion)	–0.48	–0.45*	–0.12	–0.32

Notes: psych. = psychology students
* $p < .05$ (two-tailed)

On the basis of five significant correlations (see Table 6.2), five respective Pearson r tests were run to see if these correlations might replicate in the Sample 1999 data. (Note again that prior to the Pearson r tests, all four groups in Sample 1998 and Sample 1999 were compared using independent-samples t tests. There were no significant differences between respective groups on the relevant factors.) Only one correlation was significant for

male nonpsychology participants on Factor EX, but it was not in the direction hypothesized.

6.3 Performance Comparisons: Sample 1998 vs. Sample 1999

By conducting performance comparisons (e.g., tests of homogeneity) between Sample 1998 and Sample 1999, it can be determined whether the two samples are drawn from the same population. Where these tests fail to show significant (and relatively substantial) differences between the two samples, the two samples are then combined. Thereby, better estimates of the corresponding population parameters can be calculated. New descriptive statistics for the combined sample are presented here, along with inferential statistics.

In cases where significant differences exist between samples, a critical level of variance in the dependent variable explained by the grouping variable was set in advance at 0.09 for the reasons given above (see *6.2.1*).

6.3.1 Hexagram Hitting Rates

It was shown earlier that hexagram hitting did not vary significantly between the two samples (see *6.2.1*). When the binomial test was performed for the combined sample, the hit-rate was $P_{\text{obs.}} = 0.34$ (i.e., 34%; $p = .004$), where $P_{MCE} = .25$. The effect size π is 0.60 ($\pi_{MCE} = 0.50$). Of 200 participants, 67 obtained a hit, as opposed to a chance expectation of 50. The 95% confidence interval for the proportion of hits obtained is 0.27 to 0.40. Those 200 participants were drawn from a population with a hit rate between 27% and 40%. These rates correspond to π values for the population of between 0.53 and 0.67, which does not include $\pi_{MCE} = 0.50$.

6.3.2 Number of Changing Lines

The *t*-test showed no significant difference between the two samples on the number of changing lines (see p. 119, n57). Therefore, the two samples were combined. The mean number of changing lines was 1.46 ($SD = 1.07$), where $MCE = 1.5$. The 95% confidence limit for number of changing lines for the population from which the two samples were drawn rests somewhere between 1.3 and 1.6 changing lines, which includes the chance level. (See Figure 6.2 for the distribution of Changing Line scores.)

Using the single-sample *t* test, it was found that the mean number of changing lines (1.46) was not significantly below average, $t(199) = –0.527$, $p = .599$, two-tailed. It is

suggested that some unrecognised factor (e.g., relative task difficulty) may account for the failure of participants to throw five or more changing lines, as shown in Figure 6.2. Specifically, the task of generating *at least* four changing lines is up to 10 times more difficult (P_{MCE} = .037) than the relatively easier task of matching a hexagram (P_{MCE} = .250). It is possible that participants were intuitively aware of this increased difficulty and, as a consequence, may have been intimidated by the difficulty of the task (see Chapter 10 for a discussion of the hypothesized effects of perceived difficulty).

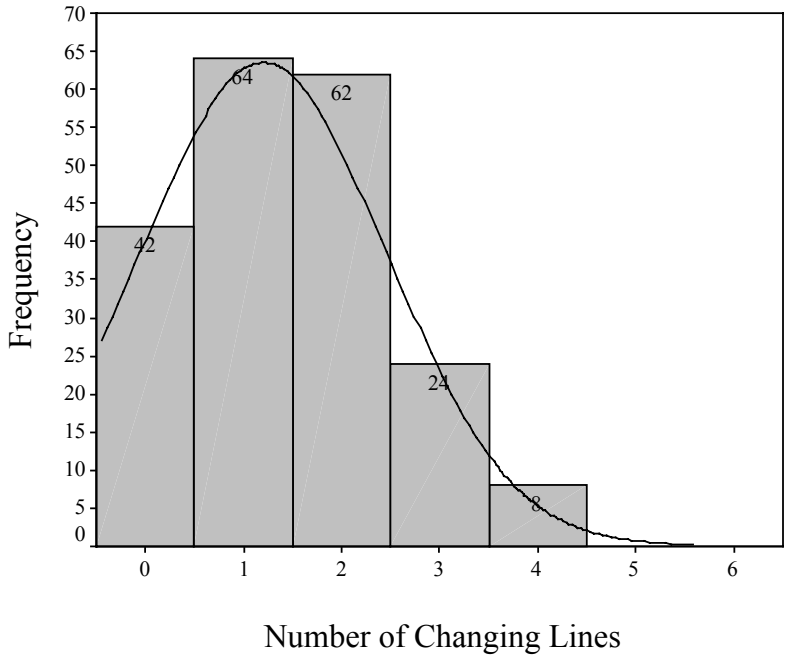

Figure 6.2. Distribution of changing lines (N = 200).

Added to the difficulty of throwing three-of-a-kind is the fact that participants were required to perform two paranormal tasks simultaneously. Expectation of significant performances in both tasks is perhaps analogous to expecting someone to demonstrate a full comprehension of a text or television program while simultaneously holding a coherent and uninterrupted conversation with a friend. Future *I Ching* experiments might yield significant performances on changing line generation if hexagram targeting is irrelevant.

6.3.3 *Transliminality*

The *t* test on the two samples showed no significant difference on transliminality scores (see *6.2.1*). The two samples were therefore combined. A mean transliminality score of 16.69 (*SD* = 5.73) was found, and with 95% confidence, the population from which these two samples were drawn would produce a mean transliminality score between 15.89 and 17.49.

The distribution of transliminality scores for the combined samples is shown in Figure 6.3, which shows an extremely small and nonsignificant negative skew of –0.044 (*SE* = 0.172). The distribution is comparatively normal, given the larger *N*, and certainly more evenly distributed than the scores for Sample 1999 (1999 kurtosis = –0.649. Cf. Figure 5.7, where the 1998 kurtosis = –0.467).

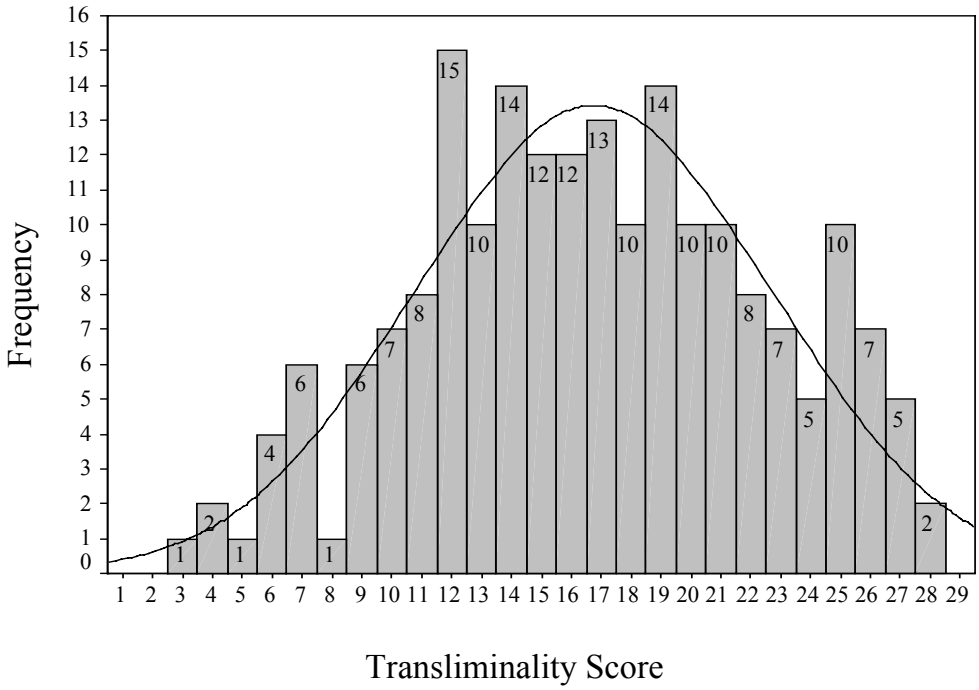

Figure 6.3. Distribution of transliminality scores (*N* = 200).

6.3.4 *The Transliminality/Hitting Correlation*

The transliminality/hitting correlation did not replicate in Sample 1999 (see *5.7.5*; Hypothesis 5). Nevertheless, given the fact that transliminality scores for the two samples are not significantly different from each other (see *6.2.1*), and the same applies to the effect

sizes for the hexagram hitting task for the two samples (again, see *6.2.1*), the transliminality/hitting correlation can be re-calculated for the combined samples.

A significant correlation was found, $r(198) = 0.12$, $p = .040$, one-tailed. Although weak, the non-chance relationship between the two variables suggests that highly transliminal participants tended to be more successful at hexagram hitting than other participants. Therefore, transliminality was shown to be a condition conducive to an expected paranormal effect, even though this relationship was not shown in Sample 1999.

6.3.5 The Transliminality/Changing Lines Correlation

Although Hypothesis 6 was not confirmed in Chapter 5 (see *5.7.5*), and, therefore, did not replicate Storm and Thalbourne's (1998-1999, p. 110) result, the correlation between these two variables can still be calculated for the combined samples. The correlation was not significant, $r(198) = 0.08$, $p = .120$, one-tailed. On this occasion, transliminality was not shown to be a condition conducive to the expected paranormal effect.

6.3.6 The Hitting/16PF Correlations

Independent-samples *t* tests were performed on the two samples to test the mean differences on the six relevant personality factors (Factors F, H, Q_2, Q_4, EX, and IN), each of which correlated with hitting in Sample 1998, but not in Sample 1999 (see *5.7.5*, Hypothesis 7). The results were not significant. Therefore, the two samples were combined, and only one out of six correlations failed to reproduce (hitting/Factor Q_2), although it was in the right direction, i.e., negative (see Table 6.3).

Table 6.3
Correlations between Hitting and 16PF Factors

Variable	*r*	*p*
Factor F (Liveliness)	0.16	.011
Factor H (Social Boldness)	0.21	< .001
Factor Q_2 (Self-Reliance)	–0.08	.128
Factor Q_4 (Tension)	–0.13	.044
Factor EX (Extraversion)	0.16	.012
Factor IN (Independence)	0.12	.042

Notes: $N = 200$; $df = 198$; *p* values are one-tailed

These five significant correlations reproduce the earlier significant findings of Storm and Thalbourne (1998-1999, p. 110), suggesting that participants who were successful at hexagram hitting tended to be lively (Factor F), socially bold (Factor H), free of tension (Factor Q_4), extraverted (Factor EX), and independent (Factor IN).

Given the results of *6.2.3* above, the hitting/Factor H correlation was calculated for female psychology students only in the combined samples (n = 63), and the result was found to be significant, $r(61) = 0.47$, $p < .001$, one-tailed.

6.3.7 The Changing Lines/16PF Correlations

Independent-samples t tests were performed on the two samples for each of the five relevant factors (Factors A, C, M, Q_2, and EX), each of which correlated with changing lines in Sample 1998, but not in Sample 1999 (see *5.7.5*, Hypothesis 8). No significant differences were found. The two samples were therefore combined. However, none of the original correlations reproduced for the combined sample (Storm & Thalbourne, 1998-1999, p. 111).

6.3.8 The Changing Lines/Possibility Correlation

In Sample 1999, the number of changing lines correlated with answers to the sheep-goat question concerning the "possibility" of a paranormal effect among participants (see *5.7.5*, Hypothesis 9). For the combined sample, the correlation was significant, $r(198) = 0.18$, $p = .006$, one-tailed. Therefore, participants who believed that it was possible for other participants to exhibit paranormal effects in the *I Ching* experiment tended themselves to generate significantly more changing lines than those who did not believe in the possibility of paranormal effects.

6.3.9 The Transliminality/Possibility Correlation

Sample 1999 replicated the significant correlation between transliminality and answers to the possibility question (see *5.7.6*, Hypothesis 10). For the combined sample, the correlation was reproduced, $r(198) = 0.37$, $p < .001$, one-tailed. High scorers on transliminality tended to believe in the possibility that participants other than themselves could achieve paranormal effects.

6.3.10 The Transliminality/Ability Correlation

Sample 1999 replicated the significant correlation between transliminality and answers to the ability question (see *5.7.6*, Hypothesis 10). For the combined sample, the

transliminality/Ability correlation was significant, $r(198) = 0.30$, $p < .001$, one-tailed. High scorers on transliminality tended to believe that their own paranormal abilities would directly contribute to a successful outcome on the paranormal task.

6.3.11 The Possibility/Ability Correlation

In Sample 1999, a significant correlation between possibility and ability was found (see *5.7.6*, Hypothesis 11). The correlation was also significant for the combined sample, $r(198) = 0.37$, $p < .001$, one-tailed. Those participants who believed that other participants had paranormal abilities tended to believe that they too had paranormal abilities.

6.3.12 The Transliminality/16PF Correlations

In Sample 1999, significant correlations between transliminality and each of four factors on the 16PF were replicated (see *5.7.6*, Hypothesis 12). When the two samples were compared on Factors A, G, M, TM, and SC using the independent-samples *t* test, it was found that there were significant differences on all but two factors: Factor A and Factor G. However, the omega-squared values calculated from the *t* values for the other three factors (M, TM, and SC) did not reach the critical level (all three values were less than 4%). Therefore, the two samples were pooled and the results of the correlations are presented in Table 6.4.

Table 6.4
Correlations between Transliminality and 16PF Factors

Variable	*r*	*p*
Factor A (Warmth)	0.13	.031
Factor G (Rule-Consciousness)	–0.25	< .001
Factor M (Abstractedness)	0.33	< .001
Factor TM (Tough-Mindedness)	–0.26	< .001
Factor SC (Self-Control)	–0.23	.001

Notes: $N = 200$; $df = 198$. *p* values are one-tailed

Note that it may still be *possible*, as stated in Chapter 5 (see *5.7.6*, Hypothesis 12), that the transliminality/Factor A correlation is an artifact of multiple analyses since *at least* one significant correlation, which is 5% of 21 correlations (there are 21 16PF factors, and therefore 21 Pearson *r* tests were originally performed on Sample 1998 data) could be the

result of chance alone. Nevertheless, the combined results indicate that highly transliminal participants in the combined sample tended to be warm and outgoing (Factor A) and idea-oriented (Factor M). They also tended not to follow rules (Factor G), were not tough-minded (Factor TM), and were lacking in self-control (Factor SC).

6.4 Multiple Regression Analyses (MRA)

In this section, the results of an MRA on Sample 1999 using the 'forward' method are presented in an attempt to replicate the regression results in Storm and Thalbourne's (1998-1999, p. 112) initial study. An MRA using the 'forward' method is also performed on the combined samples (this method is most favoured at Adelaide University). The aim of this section is to find predictors of transliminality and hitting.

6.4.1 Belief and 16PF Factors as Predictors of Transliminality

Storm and Thalbourne (1998-1999, p. 112) found that Factor M (Abstractedness), Factor A (Warmth), and answers to the question "Do you believe *in your own ability* to cast coins for a hexagram which matches one of your sixteen choices?" were significant predictors of transliminality, $R = 0.52$ (adjusted $R^2 = 0.25$). These three factors together explained 25% of the variance in transliminality scores. An MRA of Sample 1999 using the forward selection method was run to test whether ability and Factor M (but not Factor A, which failed to correlate significantly) might replicate as predictors of transliminality.

Factor M and ability both entered the model summary (Sample 1999). Ability made a moderate contribution as a predictor, $R = 0.35$ ($R^2 = 0.12$, adjusted $R^2 = 0.11$), followed by Factor M, which raised the R value to a moderate 0.41 ($R^2 = 0.17$, adjusted $R^2 = 0.15$). Thus, 15% of the variance of the dependent variable (transliminality) was explained by the two predictor variables. Two of our three findings were thereby replicated (Storm & Thalbourne, 1998-1999, p. 112).

The multiple R was significant, $F(2, 104) = 10.41$, $p < .001$. The standardized beta coefficients were as follows: ability, $\beta = 0.34$; and Factor M, $\beta = 0.21$. Ability made the greater contribution as a predictor, but only a little more than Factor M.

Given these results, it was deemed a possibility that the same predictors might replicate for the combined sample. A chi-square test on the two samples found that answers to the ability question were significantly different, $\chi^2(1, N = 200) = 11.62$, $p = .001$, two-sided ($\phi = 0.24$; $\phi^2 = 0.06$), with 52% of Sample 1998 saying yes to the ability question but only 28% of Sample 1999 saying yes to the same question. Nevertheless, the

two samples were regarded as homogeneous because the phi-squared value was less than the critical 9% (see *6.2.1*).

A *t* test was performed on Factor M scores for the two samples, and the result was also significant, $t(198) = -2.69$, $p = .008$, two-tailed ($\omega^2 = 0.03$). However, the omega-squared value also did not reach the critical level. The differences between the two samples on ability and Factor M scores were deemed not important, so they were combined. The two variables (ability and Factor M) were both entered into an MRA for the combined sample.

Ability and Factor M both entered the model, with Factor M entering first and making a moderate contribution as a predictor, $R = 0.33$ ($R^2 = 0.11$, adjusted $R^2 = 0.10$). Ability followed, raising the *R* value to a moderate 0.42 ($R^2 = 0.18$, adjusted $R^2 = 0.17$). Thus 17% of the variance of the dependent variable (transliminality) was explained by the two predictor variables. Once again, two of our three findings were replicated (Storm & Thalbourne, 1998-1999, p. 112). The multiple *R* was significant, $F(2, 197) = 21.20$, $p < .001$. The beta coefficients were as follows: ability, $\beta = 0.27$, and Factor M, $\beta = 0.30$. Factor M made the greater contribution as a predictor, but only a little more than ability.

6.4.2 Transliminality as a Predictor of Hitting

Transliminality was found to be a predictor of hitting in the 1998 study (Storm & Thalbourne, 1998-1999, p. 112). However, conducting an MRA on these two variables for Sample 1999 was academic because transliminality failed to correlate significantly with hitting in Sample 1999. Transliminality also failed to enter an MRA for the combined sample.

6.5 Median-Split Analyses (MSA)

In this section, the results of MSAs on Sample 1999 and the combined samples are presented. The first aim is to determine the success rates of hexagram hitting for 'low' and 'high' scoring participants on (i) the TLS ($N = 98$; $N = 107$ and $N = 200$), (ii) Factor H—female psychology students only (Sample 1998: $n = 39$; Sample 1999; $n = 24$; and the combined samples: $n = 63$), and (iii) Factors F, H, Q_4, EX, and IN ($N = 200$).

The second aim is to determine whether high scores on transliminality, low scores on Factor Q_4, and high scores on Factor H (for female psychology students only), and Factors F, H, EX and IN were necessary conditions that brought about significant

hexagram hit-rates. The binomial test (two-tailed) was used to calculate hexagram hit-rates as proportions of hits P, where $P_{MCE} = 0.25$.

6.5.1 Transliminality as a Necessary Condition

Storm and Thalbourne (1998-1999) performed MSAs on transliminality scores to "ascertain the locus and form of the psi effect" (p. 114). They found that high scorers on transliminality (scores greater than 17) produced a significant hit-rate, but low scores did not. Thus, there was evidence that high scoring on transliminality was a necessary condition that brought about a paranormal effect.

Sample 1999 has a median transliminality score of 17. There were 50 high-scorers, but the hit-rate was not significant, $P_{obs.} = 0.32$ ($p = .164$). It was, however, comparable to the hit-rate for Sample 1998, the proportion of which was also 0.32. This proportion translates as an effect size $\pi = 0.59$, which has a Z of 1.23, $p = .109$ (using Rosenthal & Rubin's, 1989, p. 334, formulae). There was no evidence that high scoring on transliminality was a necessary condition. There were 51 low-scorers. The hit-rate for 'low' scorers approached significance, $P_{obs.} = 0.35$ ($p = .062$). This proportion translates as an effect size $\pi = 0.62$, which has a Z of 1.72, $p = .043$. A reversal of effect took place, where low scoring on transliminality was a necessary condition that brought about a paranormal effect.

The combined samples ($N = 200$) have a median transliminality score of 17. There were 88 high-scorers, who yielded a significant result, $P_{obs.} = 0.36$ ($p = .010$), but the low scorers had a hit-rate of only $P_{obs.} = 0.30$ ($p = .135$). Thus, there was a return to a significant hit-rate for highly transliminal participants. High scoring on transliminality was a necessary condition for bringing about hexagram hitting.

6.5.2 Factor H (Social Boldness) as a Necessary Condition

It is of interest to report here that the hexagram hit-rate for all female psychology students in Sample 1998 ($n = 44$) was significant, $P_{obs.} = 0.50$ ($p < .001$; Storm & Thalbourne, 1998-1999, p. 114). However, hexagram hitting for 'high' scorers on Factor H (female psychology students only) were calculated, since we are looking at Factor H as a possible necessary condition. There was evidence that high scoring on Factor H was a necessary condition in Sample 1998.

Female psychology students in Sample 1998 have a median Factor H score of 6. There were only 16 female psychology students with 'high' scores (> 6), but the hit-rate was again very high and significant, $P_{obs.} = 0.50$ ($p = .027$). There were only 15 female

psychology students with 'low' scores (< 6), and the hit-rate was not significant, $P_{obs.}$ = 0.07 (p = .580). Thus there was evidence that high scoring on Factor H was a necessary condition, but in this case, only for female psychology students.

It is of interest to note that the hit-rate for all female psychology students in Sample 1999 (n = 24) was also significant, $P_{obs.}$ = 0.42 (p = .05). However, once again, we are looking at Factor H as a possible necessary condition. Female psychology students in Sample 1999 have a median Factor H score of 6. There were only 8 female psychology students with 'high' scores (> 6), but the hit-rate was very high, $P_{obs.}$ = 0.63 (p = .027). There were only 10 female psychology students with 'low' scores (< 6), and the hit-rate was not significant, $P_{obs.}$ = 0.30 (p = .474). Thus there was evidence that high scoring on Factor H was a necessary condition so long as participants were female psychology students.

For the combined samples, female psychology students (n = 63) again have a median Factor H score of 6. There were 24 female psychology students with 'high' scores (> 6), and the hit-rate was again high, $P_{obs.}$ = 0.54 (p = .002). There were only 25 female psychology students with 'low' scores (< 6), and the hit-rate was not significant, $P_{obs.}$ = 0.16 (p = .214). Thus again, so long as participants were female psychology students, there was evidence that high scoring on Factor H was a necessary condition.

6.5.3 Factors F, H, Q_4, EX, and IN as Necessary Conditions

It should be pointed out that the hexagram hit-rate for the combined sample is significant, $P_{obs.}$ = 0.34 (p = .004; see *6.3.1*), but once again, we are looking for necessary conditions that brought about paranormal effects. Thus, Factors F, H, Q_4, EX and IN are presented in Table 6.5 with their respective median scores, $P_{obs.}$ values and p values. For the positive correlations between hitting and Factor F, EX and IN, participants' median scores are greater than 6, and for the negative hitting/Factor Q_4 correlation, participants' median scores are lower than 6. For the positive correlation between hitting and Factor H, participants' median scores are greater than 5.

Table 6.6 lists the same variables as Table 6.5, but the positive correlations between hitting and Factor F, EX and IN, participants' median scores are smaller than 6, and for the negative hitting/Factor Q_4 correlation, participants' median scores are greater than 6. For the positive correlation between hitting and Factor H, participants' median scores are lower than 5.

Table 6.5
Hexagram Hit-Rates for 'High' Scorers on 16PF Factors F, H, Q_4, EX, and IN

Variable	*n*	Median Score	$P_{obs.}$	*p*
Factor F (Liveliness)	86	6	0.41	.001
Factor H (Social Boldness)	94	5	0.40	< .001
Factor Q_4 (Tension)	92[a]	6	0.37	.006
Factor EX (Extraversion)	65	6	0.42	.002
Factor IN (Independence)	88	6	0.40	.001

Note: $N = 200$;
[a] Low scorers

Table 6.6
Hexagram Hit-Rates for 'Low' Scorers on 16PF Factors F, H, Q_4, EX, and IN

Variable	*n*	Median Score	$P_{obs.}$	*p*
Factor F (Liveliness)	67	6	0.31	.145
Factor H (Social Boldness)	67	5	0.21	.263
Factor Q_4 (Tension)	48[a]	6	0.21	.309
Factor EX (Extraversion)	95	6	0.29	.187
Factor IN (Independence)	65	6	0.32	.112

Note: $N = 200$
[a] High scorers

In Table 6.5, it can be seen that participants with high scores (or low scores in the case of Factor Q_4) on Factors F, H, EX and IN produced significant hit-rates. Table 6.6, however, shows no hit-rates that were significant. Thus it is evident that high scoring on any, or all, Factors F, H, EX and IN were necessary conditions, and low scoring on Factor Q_4 was also a necessary condition.

6.6 Success Rates

In the Bivariate Correlation Analyses section above (see **6.2**), four nonoverlapping groups were individually tested for significant correlations between 16PF factors and the two paranormal tasks (hitting and changing lines), but only 1 test for replication was successful out of a total of 10 relevant tests: That is, the hitting/Factor H (Social Boldness) correlation for female psychology students in Sample 1999 was significant.

In the Performance Comparisons section above (see **6.3**), 24 specific statistical tests were performed to test 11 hypotheses. Of these tests, 16 produced significant results that confirmed or partially confirmed 8 of those 11 hypotheses (i.e., 67% of all tests were successful, resulting in the confirmation of 73% of all hypotheses tested).

More specifically, for the parapsychological hypotheses, there were 8 successful tests out of 16 (50%), and for the psychological hypotheses, there were 8 successful tests out of 8 (100%). All these percentages are greater than 5% and therefore are unlikely to be explained by chance alone. (Note that these percentages may be inflated due to non-independent correlations.)

Finally, in the Median-Split Analyses section (see **6.5**), all but one of 14 binomial tests (for Sample 1999 and the combined samples only) produced significant hit-rates on hexagram hitting. These tests showed that high scoring on Factors F, H, EX and IN were necessary conditions that brought about a specific paranormal effect (i.e., hexagram hitting). Low scoring on Factor Q_4 was also a necessary condition. It was necessary for scoring on transliminality to be low in Sample 1999, but high in the combined samples, leaving only one thing clear—extreme scoring on transliminality was a necessary condition that brought about hexagram hitting.

6.7 Discussion

In an initial study with the *I Ching*, Storm and Thalbourne (1998-1999) sought to find evidence in the *I Ching* process of ostensible paranormal effects. In the tradition of ESP-personality research, they also attempted to find relationships between (a) 16PF factors and success at two *I Ching* tasks (i.e., hexagram hitting and number of changing lines) and (b) transliminality and success at the two *I Ching* tasks. There were a number of significant findings (see Storm & Thalbourne, 1998-1999, pp. 108-112).

In a follow-up *I Ching* study (see Chapter 5) it was shown that the ESP-PK dichotomy was in this case unworkable in practice because the main paranormal effect (i.e., hexagram hitting) could not be categorized exclusively as ESP or PK. This effect could be explained as an example of exo-psychopraxia. However, in the same study, little light was thrown on the conducive mediating conditions that must theoretically be involved in the paranormal process which Thalbourne refers to as psychopraxia, although answering yes to a paranormal belief question ("Do you think it is possible for at least some people to exhibit paranormal effects?") was conducive to a paranormal effect (specifically, the generation of changing lines).

In Chapter 5, transliminality and nine factors on the 16PF (Factors A, C, F, H, M, Q_2, Q_4, EX, and IN), each of which previously correlated with paranormal performance of one form or another (see Storm & Thalbourne, 1998-1999), failed to replicate as conducive conditions of the same paranormal effect.

The analyses in the present chapter were run to determine possible reasons why so many previously significant findings in Storm and Thalbourne's (1998-1999) initial study were not found to be significant in the follow-up study featured in Chapter 5. In the present chapter, performance comparisons were made in order to combine, where possible, the relevant variables in the two samples.

The binomial test yielded significant hitting on designated hexagrams—the official pro attitude asked of participants (a disposition toward hitting one of these 16 pre-selected hexagrams) was fulfilled to a significant extent for Sample 1998, Sample 1999, and the combined sample.

For the combined samples, the fluctuations of the transliminality/hitting correlations for 10 groups of 20 participants were first shown to conform at the 8th group mark to a decline effect, which was hypothesized at that point, but later, when all the relevant data were collected, the correlations of the 10 groups conformed to a U-shaped trend. Thus, high scores on transliminality were associated with hitting at the beginning and the end of the experiment, whereas low scores on transliminality were associated with hitting at the middle of the experiment. The fact that the decline did not continue may be related to the experimenters focusing on that effect or to extraneous influences coming into play (see Broughton & Alexander, 1997, p. 223). Nevertheless, transliminality was repeatedly confirmed (at least suggestively) as being a condition that might help bring about exo-psychopraxia.

The alternative paranormal measure, number of changing lines, did not replicate in Sample 1999 as a significant correlation with transliminality, and no evidence of a decline effect or U-shape trend in 10 transliminality/changing lines correlations across 10 groups was present that might help explain the failure to replicate. Nor did changing lines correlate with five relevant 16PF factors in Sample 1999 or the combined sample.

For Sample 1999, six factors on the 16PF did not replicate as significant correlates with hitting, contrary to previous findings (Storm & Thalbourne, 1998-1999, p. 109), although a replicated hitting/Factor H (Social Boldness) correlation for female psychology students was found (see Table 6.1 and *6.2.3*). For the combined sample, however, five of the six significant hitting-16PF correlations reappeared (see Table 6.3). Likewise for Sample 1999, five factors on the 16PF did not replicate as significant correlates with

changing lines, contrary to previous findings (Storm & Thalbourne, 1998-1999, p. 111). For the combined sample, there were no significant changing lines/16PF correlations.

All five 16PF correlates of transliminality (four of which replicated in 1999) were significant in the combined sample (see Table 6.4).

Finally, some necessary and sufficient conditions were identified in Sample 1999 and the combined samples. These were high or low scoring on transliminality, high scoring on Factors F, H, EX, and IN, and low scoring on Factor Q_4.

Despite the poor level of replication of correlations in Sample 1999 (especially of psi-personality correlates), reproduction of many significant correlations was obtained in the combined sample for (a) parapsychological effects, particularly, the overall hexagram hitting rates and hitting/16PF correlations, and (b) psychological effects (i.e., the 16PF correlates of transliminality and various belief correlations). Thus, regarding (a), the effects of hitting and its correlates are at least consistent with the notion that exo-psychopraxia was operating and that necessary conditions may apply (i.e., some personality types tend to be more able than others to produce paranormal effects). Further studies with the *I Ching* may put the researcher on firmer ground in regard to establishing the form psychopraxia and its correlates may take.

Returning to the hexagram-hitting task, it was noted that evidence for one other pro attitude was not present (at least in sufficient strength)—second-hexagram-hitting was at chance (see again, *5.8.1.3*). Had there been psi-hitting on second hexagrams, the experimenter might be forced to conclude that there was good evidence that participants 'ran amok' by holding pro attitudes not intended by the experimenter. In accordance with these results, it is clear that another issue needs to be raised, and this issue, the problem of compliance, is introduced in the next chapter. In this chapter, a forced-choice experiment is reported, which was conducted in the attempt to clarify some of the difficulties raised in respect of the concept of the pro attitude and the problem of compliance. The so-called 'Gambling' experiment is also an attempt to discover other conditions that might be conducive to, and possibly necessary in, bringing about paranormal effects.