JP JOURNAL OF PARAPSYCHOLOGY

Volume 79 / Number 1 Spring 2015

BOOK REVIEWS

We would like to thank the following persons for their work in translating abstracts for this issue of the Journal*: Renaud Evrard (French), Eberhard Bauer (German), and Etzel Cardeña, (Spanish)*

Journal of Parapsychology, 79(1), 5–19

INVITED ARTICLE

THE MULTIPHASIC MODEL OF PRECOGNITION: THE RATIONALE[1]

By Sonali Bhatt Marwaha* and Edwin C. May**

ABSTRACT: Precognition is defined as an atypical perceptual ability that allows the acquisition of noninferential information arising from a future point in space-time. Despite the controversies, there is sufficient empirical evidence for the validity of the phenomenon. The *multiphasic model of precognition* (MMPC) is capable of addressing the experimental data. The MMPC identifies two distinct phases: The physics domain (PD) addresses the question, "How is it possible for information to traverse from one space-time point to another?" We suggest that the solution might be found within entropic considerations. The acquisition and interpretation of retrocausal signals from a future point in space-time is via three stages in the neuroscience domain (ND): Stage 1, perception of signals from an information carrier, which is based upon psychophysical variability in a putative signal transducer; Stage 2, cortical processing of the signals mediated by a cortical hyper-associative mechanism; and Stage 3, cognition, which is mediated by normal cognitive processes that lead to a precognitive response. The model is comprehensive, brain-based, and provides a new direction for research, requiring multidisciplinary expertise. In this article, the authors present the MMPC and discuss the rationale for the hypotheses put forth for the PD and the ND.

Keywords: precognition, retrocausation, entropy, signal-based model, cortical hyperassociative mechanism

Operationally, precognition (PC) is specified when a stimulus is generated relative to a response. In traditional human-centered research, a response is always elicited *after* some stimulus has been administered. In precognition studies, however, researchers reverse that order—a randomly determined stimulus is generated *after* a response has been registered. At the end of this article, we provide a more formal definition, based on the multiphasic model of precognition (MMPC).

The observables in informational-psi phenomena are: (a) information originating at some distant space-time point and (b) the information eventually reported as some form of cognition. Both informational and person perspectives are necessary for the holistic understanding of the phenomenon.

The problem of precognition (a person-centric perspective) needs to incorporate the external world. However, researchers trained in a person-centric perspective apply the principles of their field in interpreting physical theories. The absence of fidelity to the domain of a theory based in another discipline lends to utter confusion in its interpretation, leading to extending the theoretical constructs of a particular theory way beyond its intended applications, limitations, and domain specificity. Examples include quantum theory, which addresses the domain of the micro-world of matter; information theory, which addresses the domain of "pure information" without cognitive content; and phenomenological theories, which address the subjective experience of the external physical reality.

[1] This is the written version of an Invited Address delivered at the 57th Annual Convention of the Parapsychological Association, Concord, CA, USA, August 14–17, 2014.

Defining the Problem Space

The multiphasic model of precognition (MMPC) is a signal-based,[2] process-oriented model designed to determine the causal mechanisms leading to the experience of precognition. As a starting point, we have formally defined the problem space of precognition (PC) by considering two phases.

1. Phase I falls exclusively within the physics domain (PD). It addresses the question, How is it possible that information can go between two space-time points and be used, especially if the two points are separated such that their can be no causal relations between the two points? It is related to how information is carried from an external source, which is distant in time and space, to the percipient. Or in simpler terms, the PD considers everything about PC that occurs outside the head. We refer to psi in the PD as retrocausation (RC) and the carrier of such information coming from a future point in space-time as a retrocausal signal (RC signal). In doing so, we provide a distinction between the PD aspect of the process and use PC for the ND, to refer to the internal and subjective part of the process.
2. Phase II falls entirely within the neuroscience domain (ND)—internal to the human percipient. It addresses the experiential part of the problem, that is, how the information is acquired by a putative sensory system, how this information is processed in the brain, and how it is expressed. We define the acquiring mechanism as a transition between the PD and ND.

In our view, RC information arises from the external world (i.e., the PD), and is perceived and processed in the internal world (i.e., the ND). This formal bifurcation clearly lays out the domain of activity for researchers from various disciplines. A physicist interested in the question of how information can travel from there/then to here/now need not be concerned with how the information is perceived. A geneticist searching for a possible genetic basis for brain structures that permit the interpretation of a putative RC signal need not concern himself with the questions faced by a physicist. While advances in both the domains will lead to an understanding of the *process* of PC, in the intermittent phase, fidelity to one's own discipline without encroaching on the constructs of the other domain will result in less confusion and generalization of terms to domains where they are not applicable.

The MMPC addresses both the PD and the ND by considering the well-established laws of the physical world and what we currently know—and will know—about brain–behavior relationships. Thus, the MMPC is a coherent assimilation of existing concepts that we believe can lead to understanding the *process* of PC—from the point of information origin to cognition. The structure of the MMPC permits a "specialization" in PC research based on the domain that a researcher chooses to focus on, based on his area of expertise. In our view, the two domains (PD and ND) and the three stages within the ND seem to be immutable, with scope for development of more hypotheses within each domain and stage.

In this article, we discuss the rationale for the hypotheses put forth in the two domains of the MMPC. Following this, we discuss the formal definition of precognition. As with advances in any field, new developments are based on the historical past and current thinking. At the outset, we acknowledge the foundations—across disciplines—on which this model is based.

The Rationale

In this section, we discuss the rationale behind the model. To do this, we first examine what we *do know* and *do not know* about informational psi and briefly discuss the development of the MMPC. Follow-

[2] Signal detection theory "provides a general framework to describe and study decisions that are made in uncertain or ambiguous situations. It is most widely applied in psychophysics—the domain of study that investigates the relationship between a physical stimulus and its subjective or psychological effect—but the theory has implications about how any type of decision under uncertainty is made. It is among the most successful of the quantitative investigations of human performance, with both theoretical and practical applications. . . . The decisions depend on many particulars: what the decider knows, his or her expectations and beliefs, how later information affects the interpretation of the original observations, and the like. An understanding of much domain-specific knowledge is needed [to extract meaning from the physical stimulus]." (Wickens, 2002, pp. 3–4).

ing this, we discuss the considerations on which the PD and ND are based, and in the following section, present the model. Further details of the MMPC can be found in Marwaha and May (2015a, 2015b) and May and Depp (2015).

What We *Do* Know About Informational Psi

The consolidated informational psi data suggest that it is a robust statistical phenomenon and provides robust qualitative information for model development as well. Examples of quantitative data are found in the entire body of psi literature. Meta-analyses of these data present a summary picture (Bem, Palmer, & Broughton, 2001; Honorton, 1993; Honorton & Ferrari, 1989; Honorton, Ferrari, & Bem, 1998; May, Utts, Trask, Luke, Frivold, & Humphrey, 1989; Mossbridge, Tressoldi, & Utts, 2012; Steinkamp, Milton, & Morris, 1998; Storm, Tressoldi, & Di Risio, 2010, 2012).

Qualitative information for a model is found in the entire database of remote viewing studies; examples can be found in May and Marwaha (2015, p. 345), McMoneagle (2015), and Puthoff, Targ, and May (1977). As these examples suggest, the remote viewing responses have been used in operational situations along with data from other traditional sources of intelligence. The research further indicates that one cannot be trained to develop a psi ability; it is an innate ability much like musical giftedness and other aptitudes that we have (May et al., 1989). For instance, there are people who are tone deaf, many who can appreciate music, and gifted musicians who create the music. The PC information is obtained from forward in time and appears independent of distance. The channel capacity appears to be low; that is, over any reasonable time for a session, the total amount of formal information is limited. For example, it would be like trying to listen to your favorite concert with an old-fashioned telephone. This implies that the information received is diffused and scant. PC may be a nonstationary stochastic system; that is, its statistical properties are not constant. Although this uncertainty could be in the source, transmission, or detector (brain) systems, it most likely arises in the detection system, like the vagaries of perception for other sensory inputs.

There are two classes of analyses that we have for any study: (a) quantitative analysis, which helps determine the robustness of the data, and (b) qualitative information for model building, which usually is not contained in the numbers. There is much value in the qualitative information that one can acquire by examining the raw data, especially when we are undertaking the task of understanding the process of PC. While the statistical data are critical for determining what is valid, it does not help much to determine the process. An example such as Joe McMoneagle's remote viewing of a building in Severodvinsk in the northern part of the Soviet Union, conducted over several days and sessions, is best analyzed qualitatively, because like in most operations no quantitative data are available. Analysis of the interview process during response generation, transcript recording, and response drawings yield answers to questions such as the point in space-time from which the information arises, the process of data acquisition, the amount of data acquired, the correspondence of remote viewed data to the target site, insight into cognitive overlays during the process of response generation, and so forth. In describing the remote viewing of this site, McMoneagle (2015, p. 286) states, "At the time, our side did not know what was going on inside, and I was tasked with finding out via remote viewing. It was learned later that the building was being used as a Soviet submarine base to construct an unknown (to the United States) first Typhoon class submarine (as it was called by the North Atlantic Treaty Organization [NATO]). The accuracy of this viewing was later confirmed." Examples such as this provide indicators to the underlying mechanism. As a prototype of PC, this example cannot be explained via quantum theory based models or models using a QM metaphor. Examples such as this, and the numerous spontaneous experiences, need a different explanation, and at the same time they provide useful information with regard to model building.

Existing data in the psi literature, such as personality, neurophysiology and so forth, appear not to be predictive nor are they explanatory. Extraversion-introversion was one of the most widely explored dimensions of personality factors in relation to ESP. However, extraversion has not been predictive of psi ability and its apparent correlation with psi appears to be an artifact of the data collection procedure (Honorton, Ferrari, & Bem, 1998). Because we are examining the process of PC, the question as to how exactly

does extraversion contribute to the acquiring of RC information, whether we use a brain-based, a dualist, or a quantum-based model, does not yield satisfactory answers. If extraversion were a critical factor in the PC process, and not an artifact of the test situation, then, considering that extraverts are widely distributed in the population, psi too should have a wider distribution. However, this is not the case. Moreover, even if extraversion were a critical factor, we would still need to look at the neurological correlates of extraversion to understand how PC occurs.

What We Do *Not* Know About Informational Psi

One of the biggest problems that psi researchers face is that we do not know when, where, or for how long psi occurs (i.e., we are unable to pin down the specific moment when RC information is obtained from the external world). Further, we do not know the nature of the RC signal, nor what the information carrier is. We do not know from where the apparent stochastic nature of the RC signals arises. We do not know its genuine transmission rate (bits/symbol). We do not know who has psi ability and why we cannot train for it—although we now have a model (MMPC) that addresses this point and provides testable hypotheses. Confidence calling is problematic, in that for the most part it appears impossible to determine via some a priori method whether or not a given response is correct. One hint that this problem may be tractable comes from May (2007/2014a), in which he demonstrates 10 correct calls in 12 attempts in a one-in-three setup. We do not see *stable* CNS correlates. Further, we do not know *many* other things, such as whether the remote viewer sees actual or probable futures. What about free will? Can we bilk the future? And other such interesting questions that deal with the nature of time and information.

Development of the MMPC

Considering that there is statistical evidence for an information transfer anomaly that we currently do not understand, our interest was in understanding that anomaly—that is, the "how" of PC. The initial name for our model was "multiphasic model of *anomalous* cognition." Despite the many blank spaces and a need for experimental verification for the hypotheses put forth, at the end of the exercise of model development, we realized that we did have a possible answer for the process of how psi may occur; that is, the process of the PC experience may no longer be an anomaly.

An extremely important realization was the primacy of PC. An important point for PC as the *only* form of psi—both experimentally and theoretically—is that if we are to accept the possible existence of PC, then, we cannot ignore the PD (Marwaha & May, in press). Extensive details about the entropy hypothesis of the PD can be found in the section on "Entropy: A Fundamental Model of Anomalous Cognition" in May and Marwaha (2014), and May and Depp (2015). Another approach has been to say that psi is simply a form of correlation (von Lucadou, 1995, 2015) that abandons the concept of signals; however, that cannot be if the information obtained can be used. The psi literature, including the anecdotes of the Star Gate spying program, describes examples in which psi has been successfully used.

While the physics domain is especially challenging, nonetheless, the observable that psi can be used demands a signal based model.

Considerations for the neuroscience domain. Based on precognition protocols and qualitative analysis, we know that information is coming from a future point in space-time. For all our normal perceptions, the sensory organs are not extending *out* to the object of perception (as was thought by ancient theorists). Rather, information from the object is coming *to* the sensory system. For example, the eyes are not reaching out to the sun; rather, 8 minutes and 19 seconds later, photons from the sun are impinging on the retina, from where they are absorbed, biochemical reactions take place, and information is sent to the visual cortex. In our view, there is no apparent reason why precognition should be any different.

Moreover, once information from a distant space-time point (i.e., the future), traveling via the processes suggested by, and yet to be discovered in, the PD, comes to the percipient, at that moment it is no longer "pre"-cognitive. Rather, the percipient is perceiving the information in the "here and now." Thus, from the point of view of the ND, that is, the person-centric perspective, we do not have to be concerned

about the distant in space-time nature of the information. *The information—that may be from the future—is available in the present*.

In that case, as far as the perception of retrocausal signals is concerned, in a process-oriented model that addresses the person-centric perspective, we need to take into account (a) the nonuniversality of psi ability, (b) the stochastic nature of the signal, (c) the absence of an identified transducer for a putative RC signal, and (d) the probably different nature of the RC signal.

Thus, we have to ask the question *how* the human brain can perceive such RC informational signals.

The normal process of acquiring information from the external world occurs in the following three stages: perception (through a transducer), cortical processing of the information, and the cognitive processing that interprets the information. Thus, a transducer and cortical connections are required for the conversion of any external signals to biochemical actions that can be processed by the brain, so that "pure" information can be interpreted in a subjectively meaningful way.

Much of the information that we have on the "normality" of any aspect of human ability is based on averaged data. While this gives us population norms, and we can identify persons who are at the lower end of the normality curve (for developing treatment and rehabilitation programs, and for research if they present as savant syndromes), we rarely pay attention to and examine those on the higher ends, particularly the outliers. Similarly even within the ranges of what is termed normal human perceptual limits, there are outliers on both sides of the curve, whereby those individuals may have the ability to perceive ranges that are beyond the normative ranges but *within* the limitations of the species-specific abilities.

Thus, in order to address these factors, we need a model that can account for how information travels from there/then to here/now, a feature that can serve as a link between the external and internal world and process a different type of signal. The multiphasic model of precognition addresses this process across two phases: Phase I—the PD, and Phase II—the ND. It occurs in three stages: Stage 1, the perception of RC signals; Stage 2, cortical processing of RC signals; and Stage 3, cognition. In the following section, we illustrate the PD and ND of the MMPC.

Phase I: The Physics Domain

Phase I is within the PD. As stated, it addresses the question, how is it possible for information to propagate between two noncausally related points and be acted upon? It is related to how information is carried from an external source, which is distant in space-time, to the percipient. We suggest that the solution might be found within entropic considerations. Figure 1 illustrates the PD.

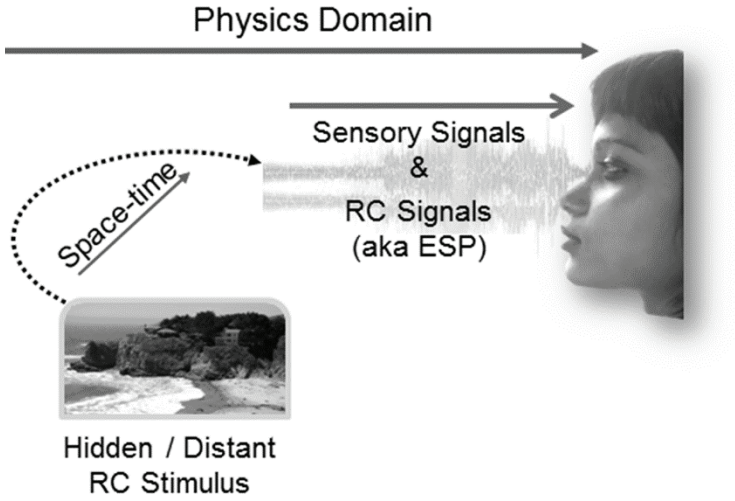


Figure 1. Physics domain. How the information traverses space-time and arrives in the near vicinity of the participant is completely independent of the parameters of that participant. He or she could be an atheist, a Buddhist monk, a stockbroker, or a skeptic.

What Might the Concept of Entropy Hold for Psi Research?

Building upon the pioneering work of Leó Szilárd (1964; Szilárd & Feld, 1972a, 1972b), Shannon and Weaver (1949) developed what is now called information theory. This theory formalizes the intuitive idea of information that there is more "information" in rare events, such as a train accident, than in common ones, such as taking a breath. Shannon defined the entropy for a given system as the weighted average of the probability of occurrence of all possible events in the system. Entropy, used in this sense, is defined as a measure of our uncertainty, or lack of information, about a system. Consider, for example, a raw egg sitting on the table. It is well organized as a small spheroidal shaped object. Now knock it to the floor and this once organized (i.e., low entropy object) becomes highly disorganized (i.e., higher entropy state). This tendency in nature for systems to move from low to higher entropic states is called the second law of thermodynamics. The details of how this relates to understanding psi can be found in May and Depp (2015). It is well known that this law provides us with the observable that the flow of time at the human level is unidirectional. For example, the skin of a newborn is smooth and unblemished (i.e., low entropy) and the skin of an old person is wrinkled and blemished (i.e., higher entropy); this apparent unidirectional flow of time is at odds with time at the micro-level.

The second law of thermodynamics relates the bidirectional flow of time in the micro world of atoms and molecules to the unidirectional flow of time at the human level. As we discuss next, entropy considerations lie at the heart of psi data.

The Data

In the Star Gate program, we noticed that a class of anomalous cognition "spying" missions *never* failed. The targets for these remote viewing sessions represented a large change of thermodynamic entropy at various target sites (May & Lantz, 2010/2014). Further research in seven laboratory studies has shown a persistent correlation with changes of entropy of the target stimuli; that is, $r = .211$, 95% confidence interval [0.084, 0.332], $p = 6.4 \times 10^{-4}$ (May, 2011/2014b). Because entropy and its changes connect the micro- to the macro-flow of time, and because PC appears to violate that flow at the macro-level, these data strongly provide a clue on how to move forward in the PD.

Criteria for the PD

There are no examples of information transfer with which one can do work (i.e., use) that do not have a concomitant energy carrier. Examples include vision information as carried upon EM waves and sound carried upon a compression wave in some medium. Therefore, it is unlikely that PC will be the first counterexample. Thus, we require an energy carrier. One implication of this requirement is that the putative energy carrier must be capable of propagating from the future to the present.

Conclusions for the Physics Domain

We have noted that there is considerable experimental evidence that suggests entropic gradients may be important in understanding the PD. Clearly, more work is needed. Further details on the PD can be found in Marwaha and May (2015a, 2015b) and May and Depp (2015).

Although the concepts of entropy, arrow of time, and information are all connected, these notions are all highly speculative at this moment as candidates for the PD part of the model. What the carrier of RC information is comprised of is even more speculative. We speculated in the PD in the MMPC that perhaps information can propagate backward in time through wormholes in hyperdimensional space.

Once the information is somehow in present time, how it gets into the CNS remains a major mystery, and the requirements for the transducer (i.e., psychic retina so to speak) will depend upon the energy details of the transfer mechanism.

The ND is far better developed, with immediately testable elements.

Phase II: The Neuroscience Domain

Phase II of the model refers to the processes that occur once the signals from any external source, including RC signals, have reached the percipient's CNS, and the processes that occur from perception to cognition of that data. This phase is primarily an implicit process. The MMPC deconstructs this domain into three discrete but fluid stages: (a) Stage 1—perception of RC signals from an energy carrier, (b) Stage 2—cortical processing of RC signals, and (c) Stage 3—cognition. One aspect of our model is that Stages 1 and 2 are critically different from normal perception in PC, following which, in Stage 3, normal processing occurs as it does for any other sensory input. Figure 2 illustrates the process of PC.

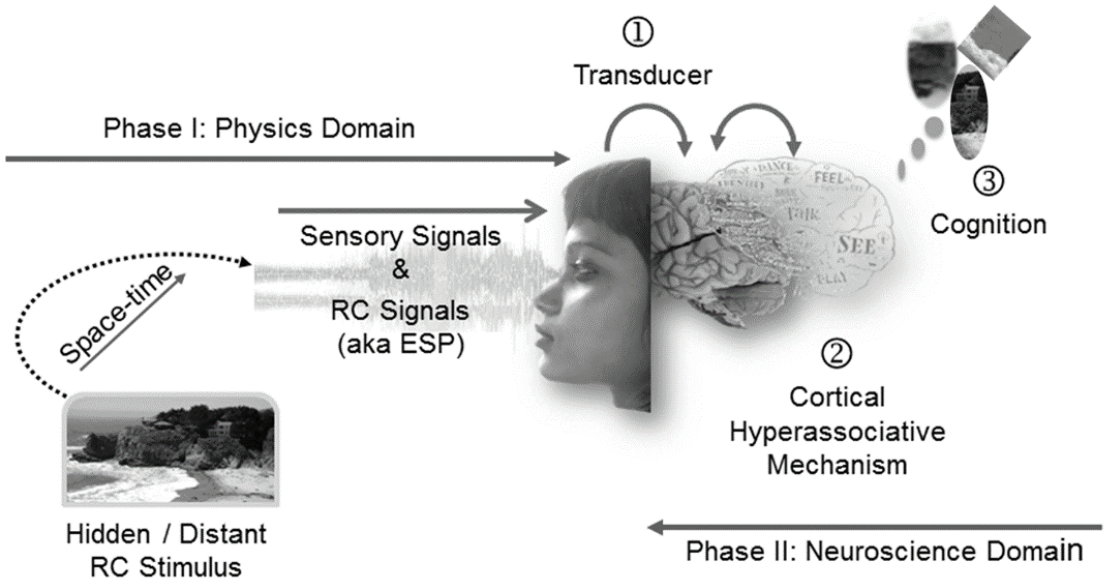


Figure 2. The process of precognition according to the multiphasic model of precognition. Although difficult to see, the cognition is scrambled compared to the stimulus. The house on the right of the stimulus is perceived accurately, but the rock on the left is missing, and the houses in between are blurred. The cognitive fragments are jumbled as well.

Stage 1: Perception of Retrocausal Signals

An important element in the process of PC is the presence of a signal transducer to serve as an interface between the incoming signals and the processing of that data. For example, the visual receptors in the retina are the transducers for signals from the visible EM spectrum. We can keep the nature of a putative PC channel an open question until we have a better understanding of the PD. Nevertheless, we can work under the assumption that PC information is received and processed internally in the same manner as are signals impinging on other sensory systems. As the nature of the putative RC signal is presently unknown, we have to assume that it is different from the normal thresholds perceived by us. This requires us to consider a possible variation in the transducer and the processing mechanisms. We have labeled the following hypotheses 1.1 and 2.1 to indicate the scope for additional hypotheses within each stage.

Hypothesis of psychophysical variability in a signal transducer (Hypothesis 1.1). This hypothesis states that psychophysical variability, in both CNS extent and function, can account for variation in the reception of RC signals; that is, some individuals demonstrate PC ability, whereas others do not. Approximately 1% of the general population possesses a natural remote viewing ability (May et al., 1989). As we

have seen over 40 years of experimental work, PC ability is seen in varying levels of proficiency across the population, much like the varying levels of music ability, for example.

This hypothesis suggests that individuals with PC ability are different from those without it at the level of sensory input, as follows: The RC signal is acquired by those with PC ability due to a possible variation in a putative biological transducer, which is the point where external signals are transformed into biochemical/bioelectrical sensory signals and transmitted to the cortical structures where cognition occurs. These signals are connected and processed across multiple cortical areas. These steps are part of Stage 2 of the MMPC as described in the following section. The question of what the transducer is may be answered by our understanding of the PD, and vice versa. (We note that this is the only part of the ND that is dependent on the PD. From Stage 2 onward, the ND is independent of any PD considerations.)

The *hypothesis of psychophysical variability in a signal transducer* is the first stage wherein retrocausal signals emerging from a distant space-time point but existing in the "now" of the percipient may be acquired from the near, external environment.

Stage 2: Cortical Processing of RC Signals

Stage 2 of the model involves the processing of RC signals received as hypothesized in Stage 1.

Hypothesis of cortical hyperassociative mechanism (Hypothesis 2.1). Considering the possible variation in the nature of an RC signal, we may assume that it has characteristics that are different from known signals, and thus we propose that RC signals are processed via a crossmodal mechanism leading to a PC experience. We consider this notion by formulating the hypothesis of cortical hyperassociative mechanism (Hypothesis 2.1).

Cortical hyperconnectivity has been associated with atypical perceptual abilities such as synesthesia (Hänggi, Beeli, Oechslin, & Jäncke, 2008; Ramachandran & Hubbard, 2001; Rouw & Scholte, 2007; Simner & Hubbard, 2013) and savant skills in autism (Wallace, Happé, & Giedd, 2009). Hence, we borrow the term *hyperassociative mechanism* as used by Simner (2012) for describing the possible underlying mechanisms for synesthesia. As Simner states:

> . . .[there may be] *one of any number of* [original emphasis] possible neurological processes that might give rise to the "open channel" between different brain regions, which allows sound to be interpreted as colour, taste as touch, touch as smell, and so on. In fact, this neutral term should cover not one of several possible mechanisms, but rather, one or more of these possibilities. . . . Whether a functional connection is established by hyperconnectivity, by disinhibited pathways, by other means, or indeed, a combination of these, the outcome is the opening of a communication between regions that would otherwise not directly interact to produce a conscious experience in the average person. (p. 25)

Considering the possibility of a similarity in underlying mechanisms between synesthesia and the proposed internal PC mechanism, in our view the synesthete population may be a good point to start for examining this hypothesis. A crucial distinction between synesthesia and PC experiences is that synesthetic experiences apparently may occur in the *absence* of an external signal. As Sean Day (personal communication, January 3, 2014) stated, "Synesthesia does not convey information of any reality; it adds one or more nonreal perceptions to an initial inducing perception which may (or may not) be based upon reality." PC experiences, as we propose here, arise due to the *presence* of external signals. Others have also claimed that synesthesia may actually underpin ESP experiences (e.g., Alvarado, 1994; Myers, 1903; Simmonds-Moore, 2010, 2014).

In summary of Stages 1 and 2, RC signals may be acquired by an individual with PC ability, due to psychophysical variability in the signal transducer (Hypothesis 1.1); these signals are then processed in accordance with Stage 2. The cortical hyperassociative mechanism (Hypothesis 2.1) takes advantage of a possible increase in spectral range of EM and/or the perception of the RC signals that the brain can process.

Stage 3: Cognition

Following Stage 2, we propose that in Stage 3 cognition of the information occurs through the normal process, as it does for other sensory inputs, wherein information is stored, retrieved, influenced by subjective experiences, and manifested in the form of a response such as ideas/writing/drawing/narration/ in a dream state.

We strive to keep what is known in physics, psychology, and neuroscience as intact as possible. Therefore, it seems reasonable to assume that the processes involved in cognition of signals from normal sensory modalities will also be involved in the cognition of RC information. We have indicated in Stages 1 and 2 of the ND of the model that high variability of sensorial systems connected via a hyperassociative mechanism to other CNS structures may be the pathway for RC signals into the CNS. Once there, the cognitive correlates of RC signals may be indistinguishable from those of normal sensory signals. This would suggest that it might not be possible to observe CNS correlates of PC inputs simply because they would be indistinguishable from other CNS correlates. This is particularly so because, so far, we are unable to determine what signals to look for and unsure about *when* the RC signals were received and processed by a transducer. For instance, the percipient may have received them before being connected to the EEG gear or placed into an fMRI scanner, or the day before, or in the parking lot before entering the laboratory; or perhaps the signals have such a short duration that they are missed by the hardware.

As stated earlier, RC signals are generally not robust and are difficult to detect. Moreover, they also appear to be statistically nonstationary; that is, statistical properties vary with regard to when they are measured. Where that apparent nonstationary aspect arises is unknown; however, as stated, there are only three possibilities—at the source, in the transmission channel, or in the detection mechanism. Therefore, the cognition resulting from PC is unreliable. Normal psychological influences, such as memory formation, emotional overlays, lack of attention or intention, ill health, effects of medications, and so on, will interfere with PC response formation as they do with other forms of cognitive activities (McMoneagle & May, 2004/2014).

As PC is an unconscious process, we can assume that a PC-abled person is privy to a wider spectrum of information that he or she incorporates in his or her daily life. In our view, the most crucial aspect in the PC process lies in the PD, and PC involves Stages 1 and 2 of the ND. The MMPC proposes that once the information (PC stimulus—new information) has been implicitly received, it is stored in the memory in the same way as information from the other sensory systems. It is retrieved when the need for it arises or in dreams.

Factors such as attention, emotions, beliefs, memory, creativity, uncontrolled random thoughts, intellectual decisions, linguistic influences, and so on (May & Trask, 1988) may interfere with PC responses as they do for the other senses. In the parlance of PC protocols, they are referred to as "overlays." Thus, in a normal protocol in a PC session at our laboratory, a participant is asked to first note-down/illustrate the thoughts/images that are on the top of the participant's mind before the session begins. In this manner, the cognitive overlays—the personal preoccupations—are brought to the conscious level, and PC cognitions can be recognized by the percipient as being distinct. In an experimental situation, an experienced PC-abled individual can distinguish information that is emerging from his own frame of reference from newly acquired information.

As stated, the crucial distinction in PC—or the perception of RC-signals—occurs in Stages 1 and 2 of the ND. We further propose that examining these domains will lead to clues to the nature of the RC-signal and probably to the PD. In Marwaha and May (2015a, 2015b) and May and Depp (2015), we have provided supporting evidence from other areas of human perception and cognition that serve as a basis for the hypotheses put forth here.

The hypotheses put forth in the MMPC open up a new dimension in psi research. Figure 3 summarizes the testability feature of the ND based on a standard remote viewing test, for the identification of persons with PC ability.

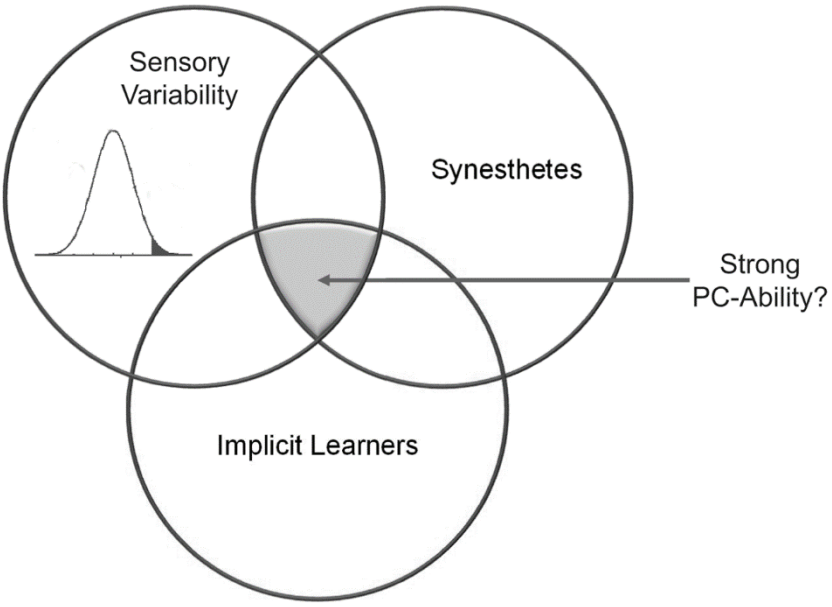


Figure 3. Identifying people with a precognition ability.

Defining Precognition

Based on the MMPC, we define precognition as *an atypical perceptual ability that allows the acquisition of noninferential information arising from a future point in space-time.*

In an earlier version of the paper, we had used the term "space-like separated point" instead of "future point." We are indebted to Stephan Baumgart for pointing out (personal communication, April 29, 2015) that the light cone shows that most all PC target stimuli are not space-like separated from the participant. Figure 4 shows the Minkowski light cones that illustrate this point.

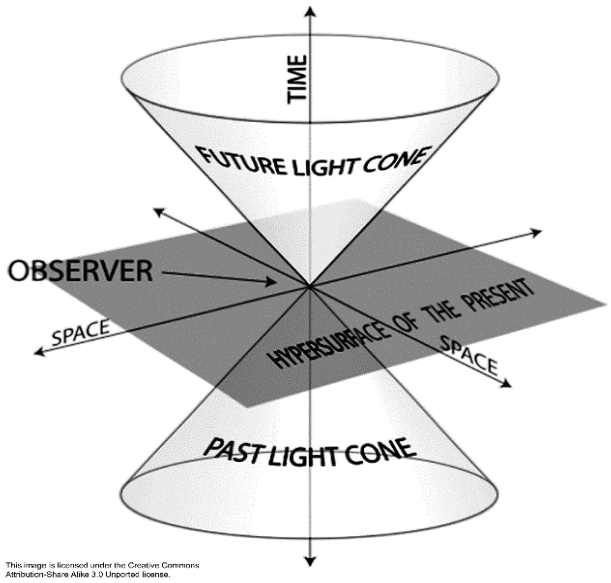


Figure 4. The standard depiction of the Minkowski light cones. The vertical axis represent time—times greater than "now" (i.e., time of the observer) represent the future. The past is represented by times that have gone by relative to the observer and are indicated by the "past light cone." The x- and y-axes represent space.

Any points in space that are outside the upper cone in Figure 4, are called space-like separated from the observer. The fact that the edges of the cone represent how long it would take light to reach the observer given its distance, means that space-like separated points cannot be causally connected to the observer. For example, the circumference of the Earth is approximately 40,000 km, so the farthest away from the observer a psi stimulus could be is half that or 20,000 km. Because the speed of light is 300,000 km/s, it takes only 67 ms for light to reach the observer. That means a PC target, at best, is space-like separated from the observer by times shorter than that. Thus, for the most part, all laboratory-based PC experiments reside inside the upper cone yet remain noncausally related to the observer.

As a laboratory anecdote, we conducted a number of sessions wherein the target stimulus was the half-lit moon Io of Jupiter as it was eclipsed by the planet. At the time of the series, Jupiter was ~40 light-minutes away from earth, so in this case the target stimulus was space-like separated.

Conclusion

From the point of view of a psi experience, there are a few take-home messages from this model:

1. In our view, the bifurcation into two distinct domains collapses the problem space for experts within each domain to address separately.

2. The PD is concerned with the nature, carrier, and transmission of a putative RC-signal, emerging from a distant point in space-time. This is purely an information-centric perspective.

3. The "pre" aspect of the precognition experience is occurring in "real-time." Just as we are not concerned with the nature, carrier, and transmission of photons from the sun to our retina, as percipients of putative RC signals, we are not concerned with from what distant space-time point the RC signal emerged. As far as experience is concerned, by intention and/or attention a percipient has focused or accessed information—that is "floating" in the "here and now" vicinity of the transducer. Thus, from an experiential point of view (i.e., a person-centric view), the question of the "logical possibility of PC" does not arise.

4. Thus, what we now know as "precognition" becomes just another type of perception, albeit an atypical one.

5. The ND of the MMPC, particularly Stages 1 and 2, provide a bottom up approach to investigating the larger questions of the nature of RC signals. Examining the ND may yield information on the probable bandwidth on which RC signals may be carried and the sensory modalities that are involved. This may provide data for the PD to explore to determine the nature of a possible RC signal carrier.

6. This opens the door to understanding the fundamental questions that the experience of PC has raised—the nature of time, causality, and information.

7. The MMPC thus naturalizes the supernatural and the spiritual.

8. In the PD and ND there is scope for developing additional hypotheses.

9. Each aspect of the model, particularly in the ND, is eminently testable, requiring multidisciplinary expertise.

10. The two domains and the three stages provide us a language with which to analyze any psi problem.

11. The MMPC has the potential to fulfil the six criteria for the evaluation of a theory, as dis-

cussed by Cramer (2013): comprehensiveness, precision and testability, parsimony, empirical validity, and both heuristic and applied value.[3]

We strongly believe that as we shift focus from a person-centric perspective to a signal-based information-centric perspective, the seemingly difficult problems of the PC experience become relatively easy to explore. To examine the suggested hypotheses we need a truly interdisciplinary team. For all we know, answers for many elements of our questions may already be there in other disciplines and superspecialties. If we want to solve this problem, we need mainstream scientists to view psi as an atypical ability, rather than a spiritual, supernatural, or paranormal ability. The final theatre of this experience rests in the information-centric perspective, that is, in the PD for which the ND can provide clues.

References

Alvarado, C. S. (1994). Synesthesia and claims of psychic experiences: An exploratory study. *Proceedings of Presented Papers: The Parapsychological Association 37th Annual Convention,* 11–17.

Bem, D. J., Palmer, J., & Broughton, R. S. (2001). Updating the ganzfeld database: A victim of its own success? *Journal of Parapsychology, 65*, 207–218.

Cramer, K. M. (2013). Six criteria of a viable theory: Putting reversal theory to the test. *Journal of Motivation, Emotion, and Personality*, *1*, 9–16.

Hänggi, J., Beeli, G., Oechslin, M. S., & Jäncke, L. (2008). The multiple synaesthete ES— Neuroanatomical basis of interval-taste and tone-colour synaesthesia. *Neuroimage*, *43*, 192–203.

Honorton, C. (1993). Impact of the sender in ganzfeld communication: Meta-analysis and power estimates. In E. C. May, W. L. W. Luke, & N. D. Lantz (Eds.), *Phenomenological research and analysis* (Final Report: 6.2 and 6.3). Menlo Park, CA: SAIC.

Honorton, C., & Ferrari, D. C. (1989). "Future telling": A meta-analysis of forced-choice precognition experiments, 1935–1987. *Journal of Parapsychology, 53*, 281–301.

Honorton, C., Ferrari, D. C., & Bem, D. J. (1998). Extraversion and ESP performance: A meta-analysis and a new confirmation. *Journal of Parapsychology, 62*, 255–276.

Marwaha, S. B., & May, E. C. (in press). Precognition: The only form of psi? *Journal of Consciousness Studies*.

Marwaha, S. B., & May, E. C. (2015a). Rethinking extrasensory perception: Towards a multiphasic model of precognition. *SAGE Open*, January-March, 1–17. Retrieved from http://sgo.sagepub.com/content/5/1/2158244015576056

Marwaha, S. B., & May, E. C. (2015b). Multiphasic model of precognition. In E. C. May & S. B. Marwaha (Eds.), *Extrasensory perception: Support, skepticism, and science: Vol. 2. Theories of psi* (pp. 145–170). Santa Barbara, CA: Praeger.

May, E. C. (2014a). Advances in anomalous cognition analysis: A judge-free and accurate confidence-calling technique. In E. C. May & S. B. Marwaha (Eds.), *Anomalous cognition: Remote viewing research and theory* (pp. 80–88). Jefferson, NC: McFarland. (Original work published 2007)

May, E. C. (2014b). Toward a classical thermodynamic model for retro-cognition. In E. C. May & S. B. Marwaha (Eds.), *Anomalous cognition: Remote viewing research and theory* (pp. 327–338). Jefferson, NC: McFarland. (Original work published 2011)

May, E. C., & Depp, J. (2015). Entropy and precognition: The physics domain of the multiphasic model of precognition. In E. C. May & S. B. Marwaha (Eds.), *Extrasensory perception: Support, skepticism, and science: Vol. 2. Theories of psi* (pp. 125–144). Santa Barbara, CA: Praeger.

May, E. C., & Lantz, N. D. (2014). Anomalous cognition technical trials: Inspiration for the target entropy concept. In E. C. May & S. B. Marwaha (Eds.), *Anomalous cognition: Remote viewing research and theory (*pp. 280–298).

[3]*Comprehensiveness* refers to the range and diversity of phenomena encompassed by a theoretical perspective—the more comprehensive a perspective, the more ground it covers. *Precision and testability* demands that a good theory consist of constructs that are clearly defined, tightly interrelated, and readily open to reliable and valid measurement through falsifiable hypotheses (Popper, 1963). The criterion of *parsimony* highlights the notion that the preferred theoretical account of events is the one requiring the fewest number of concepts—the fewer the concepts the more parsimonious the theory. The idea of parsimony also includes simplicity—the simpler the theoretical account, the more parsimonious the theory. *Empirical validity* reflects the extent to which a theory manages disconfirming evidence, because studies with negative results carry more weight than those with positive results. A theory's *heuristic value* involves its ability to generate unique thoughts and perspectives and directions in other fields. A theory's *applied* value can be measured by the extent to which it offers effective solutions to problems.

Jefferson, NC: McFarland. (Original work published 2010)
May, E. C. & Marwaha, S. B. (2015). Next step—Process-oriented research: Guidelines for experimenters. In E. C. May & S. B. Marwaha (Eds.), *Extrasensory perception: Support, skepticism, and science: Vol. 2. Theories of psi* (pp. 329–354). Santa Barbara, CA: Praeger.
May, E. C., & Trask, V. V. (1988). *Forced-choice remote viewing* (Project No. 1291). Menlo Park, CA: SRI International.
May, E. C., Utts, J. M., Trask, V. V., Luke, W. W., Frivold, T. J., & Humphrey, B. S. (1989). *Review of the psychoenergetic research conducted at SRI International (1973-1988)* (Final Report, Task No. 6.0.1, Project No. 1291). Menlo Park, CA: SRI International.
McMoneagle, J. W. (2015). Evidence for precognition from applied remote viewing. In E. C. May & S. B. Marwaha (Eds.), *Extrasensory perception: Support, skepticism, and science: Vol. 1. History, controversy, & research* (pp. 285–316). Santa Barbara, CA: Praeger.
McMoneagle, J. W., & May, E. C. (2014). The possible role of intention, attention and expectation in remote viewing. In E. C. May & S. B. Marwaha (Eds.), *Anomalous cognition: Remote viewing research and theory* (pp. 368–376). Jefferson, NC: McFarland. (Original work published 2004)
Mossbridge, J., Tressoldi, P., & Utts, J. (2012). Predictive physiological anticipation preceding seemingly unpredictable stimuli: A meta-analysis. *Frontiers in Psychology, 3*, Article 390. Retrieved from http://dx.doi.org/10.3389/fpsyg.2012.00390
Myers, F. W. H. (1903). *Human personality and its survival of bodily death*. London, England: Longmans, Green.
Popper, K. (1963). *Conjectures and refutations*. New York, NY: Basic Books.
Puthoff, H. E., Targ, R., & May, E. C. (1977). *Advanced threat technique assessment* (Final Report, Project No. 7403). Menlo Park, CA: SRI International.
Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies*, *8*, 3–34.
Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience*, *10*, 792–797.
Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.
Simmonds-Moore, C. A. (2010). Personality variables in spontaneous psi research: Contextualising the boundary construct in its relationship to spontaneous psi phenomena. In C. A. Roe, W. Kramer, & L. Coly (Eds.), *Proceedings of an International Conference Utrecht II: Charting the future of parapsychology* (pp.151–215). New York, NY: Parapsychology Foundation.
Simmonds-Moore, C. A. (2014). Do synesthetes perform better at a precognition task than a group of matched controls? *Abstracts of Presented Papers: The Parapsychological Association 57th Annual Convention,* 16–17.
Simner, J. (2012). Defining synaesthesia: A response to two excellent commentaries. *British Journal of Psychology*, *103*, 24–27.
Simner, J., & Hubbard, E. M. (2013). *The Oxford handbook of synesthesia.* New York, NY: Oxford University Press.
Steinkamp, F., Milton, J., & Morris, R. L. (1998). A meta-analysis of forced-choice experiments comparing clairvoyance and precognition. *Journal of Parapsychology, 62*, 193–218.
Storm, L., Tressoldi, P. E., & Di Risio, L. (2010). Meta-analyses of free-response studies 1992–2008: Assessing the noise reduction model in parapsychology. *Psychological Bulletin, 136*, 471–485. doi:10.1037/a0019457.
Storm, L., Tressoldi, P. E., & Di Risio, L. (2012). Meta-analysis of ESP studies, 1987–2010: Assessing the success of the forced-choice design in parapsychology. *Journal of Parapsychology, 76*, 243–273.
Szilárd, L. (1964). On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Behavioral Science*, *9*, 301–310.
Szilárd, L., & Feld, B. T. (1972a). On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *The collected works of Leó Szilárd: Scientific papers* (pp. 103–129). Boston, MA: MIT Press.
Szilárd, L., & Feld, B. T. (1972b). On the extension of phenomenological thermodynamics to fluctuation phenomena. *The collected works of Leó Szilárd: Scientific papers* (pp. 34–102). Boston, MA: MIT Press.
von Lucadou, W. (1995). The model of pragmatic information (MPI). *European Journal of Parapsychology*, *11*, 58–75.
von Lucadou, W. (2015). The model of pragmatic information. In E. C. May & S. B. Marwaha (Eds.), *Extrasensory perception: Support, skepticism, and science: Vol. 2. Theories of psi* (pp. 221–242). Santa Barbara, CA: Praeger.
Wallace, G. L., Happé, F., & Giedd, J. N. (2009). A case study of a multiply talented savant with an autism spectrum disorder: Neuropsychological functioning and brain morphometry. *Philosophical Transactions of the Royal So-*

ciety B: Biological Sciences, *364*, 1425–1432.
Wickens, T. D. (2002). *Elementary signal detection theory*. Oxford, England: Oxford University Press.

**Visakhapatnam*
Andhra Pradesh, India
sonalibm@gmail.com

***Laboratories for Fundamental Research*
Palo Alto, CA, USA

Abstracts in Other Languages

French

LE MODELE MULTIPHASIQUE DE LA PRECOGNITION : L'ARGUMENTAIRE

RÉSUMÉ: La precognition est définie comme une habileté perceptuelle atypique qui permet l'acquisition d'information non-inférentielle en provenance d'un point futur de l'espace-temps. En dépit des controverses, il y a suffisamment de preuves empiriques en faveur de la validité du phénomène. Le modèle multiphasique de la précognition (MMPC) est apte à rendre compte des données expérimentales. Le MMPC identifie deux phases distinctes : le domaine physique (PD) répond à la question « Comment est-il possible qu'une information traverse l'espace-temps d'un point à un autre ? Nous suggérons que la solution pourrait se trouver dans une réflexion sur l'entropie. L'acquisition et l'interprétation de signaux rétro-causaux d'un point futur de l'espace-temps passe par trois étapes dans le domaine neuroscientifique (ND) : l'étape 1 où les signaux sont perçus au moyen d'un porteur d'information, qui est basé sur la variabilité psychophysique dans un supposé transducteur de signal ; l'étape 2 avec le traitement cortical des signaux médiatisés par un mécanisme d'hyper-associativité corticale ; et l'étape 3 de la cognition, médiatisée par les processus cognitifs normaux qui aboutissent à la réaction précognitive. Ce modèle est global, basé sur le fonctionnement du cerveau et fournit de nouvelles directions de recherche qui nécessitent une expertise multidisciplinaire. Dans cet article, les auteurs présentent le MMPIC et discutent de la raison d'être de ses hypothèses concernant tant le PD que le ND.

German

DAS MULTIPHASISCHE MODELL DER PRÄKOGNITION: DER GRUNDGEDANKE

ZUSAMMENFASSUNG: Präkognition wird hier definiert als ein atypisches Wahrnehmungsvermögen, das den Erwerb von Information gestattet, die auf keinen Schlussfolgerungen beruht und einem zukünftigen Punkt in der Raumzeit entspringt. Ungeachtet der Kontroversen gibt es ausreichende empirische Evidenz für die Echtheit des Phänomens. Mit Hilfe des multiphasischen Modells der Präkognition (MMPK) können die experimentellen Befunde eingeordnet werden. Das MMPK unterscheidet zwei Phasen: Der Physikbereich (PB) behandelt die Frage, wie es möglich ist, dass Information von einem Raumzeitpunkt zu einem anderen gelangt? Wir schlagen vor, dass die Lösung mit Überlegungen zur Entropie zusammenhängt. Der Erwerb und die Interpretation retrokausaler Signale von einem zukünftigen Punkt in der Raumzeit geschieht mittels dreier Stadien im neurowissenschaftlichen Bereich (NB): Stadium 1: Wahrnehmung von Signalen eines Informationsträgers, der auf der psychophysischen Variabilität eines angenommenen Signalwandlers beruht; Stadium 2: die kortikale Verarbeitung der Signale, die durch einen kortikalen hyper-assoziativen Mechanismus vermittelt werden, und Stadium 3, die Kognition, die durch normale kognitive Prozesse vermittelt wird und zu einer präkognitiven Reaktion führt. Das Modell ist umfassend, basiert auf dem Gehirn und eröffnet eine neue Forschungsrichtung, die multidisziplinäre Kenntnisse erfordert. Die Autoren stellen in diesem Artikel das MMPK vor und diskutieren die Grundgedanken für die Hypothesen, die den PB und NB betreffen.

Spanish

EL MODELO MULTIFÁSICO DE PRECOGNICIÓN: LA JUSTIFICACIÓN

RESUMEN: Se define a la precognición como la capacidad perceptiva atípica que permite la adquisición de información no inferencial resultante de un punto futuro en el espacio-tiempo. A pesar de las controversias, existe suficiente evidencia empírica de la validez del fenómeno. El modelo multifásico de la precognición (MMPC) es capaz de hacer frente a los datos experimentales. El MMPC identifica dos fases distintas: El dominio de la física (PD) se refiere a la pregunta, ¿Cómo es posible que la información pueda viajar de un punto del espacio-tiempo a otro? Sugerimos que la solución podría encontrarse dentro de las consideraciones entrópicas. La adquisición e interpretación de señales retrocausales desde un punto en el espacio-tiempo futuro es a través de tres etapas en el dominio de la neurociencia (ND): Escenario 1, la percepción de señales de un agente de información, basada en la variabilidad psicofísica en un transductor putativo de señales. Etapa 2, el procesamiento cortical de las señales mediadas por un mecanismo cortical hiper-asociativo. Y Etapa 3, la cognición, mediada por procesos cognitivos normales que conducen a una respuesta precognitiva. El modelo es comprehensivo, basado en el cerebro, y proporciona una nueva dirección para la investigación que requiere conocimientos multidisciplinarios. En este artículo, los autores presentan el MMPC y discuten los fundamentos de la hipótesis planteados por el PD y el ND.

Journal of Parapsychology, 79(1), 20–36

THE JUNGLE OF HYPNOTIC PSI: PART 1. RESEARCH ON HYPNOSIS RELEVANT TO PSI

By Adrian Parker

ABSTRACT: Contemporary research efforts have attempted to establish a psi-conducive hypnotic state primarily in terms of hypnotizability and dissociation. Much of this appears to be founded on the early historical association between hypnotic and psi phenomena, a misunderstood meta-analysis, and a number of shared variables. In reality, the state concept of hypnosis is still highly disputed and a review of the 60 years of intensive research on hypnosis indicates hypnosis is best seen as a complex combination of various factors: a "jungle of variables" rather than a unitary state. This implies the search for a specific psi-conducive state may be an illusory venture. However, some of the efforts to reduce hypnosis to unconsciously elicited expectancies or "automaticity" appear to reflect the current trend to denigrate the role of conscious processes. Paradoxically, the more dramatic phenomena of hypnosis may instead illustrate the potency of altered belief systems in effecting major changes in psychological processes.

Keywords: hypnotic state, psi, consciousness, placebo effect

In a previous paper (Parker & Millar, 2014) the crisis in psi research was related to the failure to deal with the psi-based experimenter effect. The myth here was that by studying psi like any other ability, as a normally distributed variable, progress would be made. The current two-part paper is a companion to this and deals with another apparent myth that exacerbates the crisis: the excessive belief in the historic strength of the association of hypnosis with psi. It will be argued here that research has headed off in wrong directions with the result that it has now lost itself in a forest of findings. Like experimenter effects, a critical review of the evidence suggests there are nevertheless some promising ways forward.

Research on what I call "hypnotic psi" is one of the few consistently active areas of research remaining in contemporary parapsychology Arguably, the revival of interest in using hypnosis to facilitate psi is appealing in psi research circles because it promises a refuge for what appears to be the robust historical association between the two areas.

It will be argued here that poorly financed and ill-equipped ventures of this nature are potentially perilous given the complexity and demands of the issues encountered. Most of the contemporary publications on the topic of hypnotic psi (to be discussed fully in Part 2) have implicitly assumed—actually on little or no evidence—that hypnosis has been established not only as an altered state but also as a psi-conducive one. These current reports of the investigations of hypnotic psi show unfortunately little overt awareness of the complexity of the issues surrounding the nature of hypnosis that have evolved during recent years. Moreover, it will be later argued that in the face of this complexity, the need for maintaining the rigorous research tradition in parapsychology has been neglected. The risk is that the application of a less than state-of-the-art methodology to a heavily disputed area can actually worsen rather than improve the credibility of parapsychology.

The fact of the matter is that the existence or nonexistence of a hypnotic state has been and still is the subject of 60 years of intensive research and debate in which there is little agreement (Accardi, Cleere, Lynn, & Kirsch, 2013; Kallio & Revonsuo, 2003; Kirsch, 2005; McConkey, 2008). What is more disconcerting is that many of the above studies based their claims for hypnosis being a psi-conducive state on the meta-analysis by Stanford and Stein (1994) as having established that such a state probably exists. Stanford

and Stein were, however, very cautious, especially given the number of methodological flaws that they discovered (along with numerous other problems in the database), to avoid any simple and firm conclusions about the effect of hypnosis above and beyond control groups.

Stanford and Stein were well aware that any procedure that appeared powerful to the participants could thereby have an enabling effect similar to that of a physician administering a placebo or indeed a magician possessing showmanship. The fact that Mesmer was both a physician and a showman did in fact set the stage for future practitioners of the art and prepared the ground for the ensuing controversies over the true nature of hypnosis (Buranelli, 1975). This is not to say that some of these later hypnotists were not reputable and successful practitioners. One of these, the Swedish physician Axel Munthe (best known in parapsychology for being the attending physician at F. W. Myers' death in Rome) lived on until 1949 and left us with a vibrant autobiographical link with the heyday of hypnosis. After an intense dispute with neurologist-hypnotist Jean-Martin Charcot over Charcot's dubious form of hypnosis and exploitation of women patients, Munthe became himself a celebrated high society physician while at the same time maintaining a benevolent practice for treating the destitute. He openly admitted that many of his successes in high society were due to using a mixture of psychological expectancies, placebo effects, and hypnotic-type suggestions (Munthe, 1929).

Munthe's success illustrates how the issue of deciding what is genuine hypnosis is clearly a complex one. The present writer received a memorable reminder of the difficulties inherent in dealing with this complexity following the presentation of his undergraduate research project on hypnosis at the Foundation for Research on the Nature of Man (FRNM). The presentation elicited a redoubtable response from the FRNM's director, J. B. Rhine, that "Using hypnosis to facilitate psi is returning to the jungle of experimentation." So as far Rhine was concerned, hypnosis was not just becoming lost in a forest of findings but in a jungle of variables.

Although these comments may have seemed at the time to be slightly ungracious, especially to a then 19-year-old, it can be said with hindsight that they contained an element of truth. It was this element of truth, together with a pragmatic sense for survival (universities might tolerate one controversial area but two controversies easily become "two too many"), that became the impetus for my involvement in the development of the ganzfeld technique. I regarded the ganzfeld as means of incorporating some of the active ingredients in hypnosis without any of the inherent complexity and controversy surrounding the use of hypnosis (Parker, 2003b).

Now, seen with respect to the recent hypnotic revival, some salient questions rightly reassert themselves: Has something been missed in trying to lift out the active features of hypnosis and redact it as the ganzfeld? Do the above recent efforts, despite difficulties, offer the possibility of finding, albeit not a royal road, but at least a promising path through the forest of various findings?

The problem is that it seems more appropriate to rename the forest as "Rhine's jungle," albeit not of vines but of variables. Table 1 lists just some of the major more contemporary ones that we will need to critically examine in some detail before any attempt can be made at finding a way forward. Many of the variables are obviously intimately interrelated. An overview is made easier if we regroup some of these variables under the major alliances, which are those of dissociation theory (encompassing neodissociation, dissociated control, and neurocognitive function) and social-neurocognitive theory (encompassing compliance and expectancy). Seen as such, dissociation theory is the natural heir to the trance and altered-state theory of hypnosis, whereas social-neurocognitive theory has a more reductionist-skeptical orientation. These divisions represent the main contestants involved in what has become at times a full-scale fight among the claimants for the hypnosis territory.

Certainly, there can be no doubt from this list that the current history of hypnosis is highly relevant to research efforts in parapsychology. Many of the variables being dealt with, such as absorption and dissociation, are of mutual concern to psi research and hypnosis. Moreover hypnosis and hypnosis-psi research share not only a common history, but also parallel controversies. These controversies are so perennial and persistent that the cynic might be tempted to say that only the actors change while the issues remain written into the script along with the roles for acting these out. Finally, it should be pointed out that it is my

view that psi does not exist in a vacuum, nor in rational states, but rather in altered states of consciousness (Parker, 2003b).

Table 1
A List of Some Current Variables and Authorities Endorsing the Explanatory Power of Variables for Understanding What Hypnosis Is

Some Variables and Theories of Hypnosis	Some Principal Investigators With Representative Publications
Neodissociation	Hilgard, 1992; Kihlstrom, 2005, 2008
Dissociated control	Bowers, 1992
Fixation of attention plus neurocognitive factors	Rainville & Price, 2003; Gruzelier, 2006
Social role play	Barber, 1969
Suggestibility	Raz, 2007; Halligan & Oakley, 2014
Absorption	Roche & McConkey, 1990
Compliance	Spanos & Coe, 1992; Wagstaff & Cole, 2005
Fantasy proneness	Wilson & Barber, 1983; Lynn & Rhue, 1986
Expectancy and nondeceptive placebo	Kirsch, 1994; Kirsch & Lynn, 1995; Accardi et al., 2013
Mind-wandering influence	McGeown, Mazonni, Vennen, & Kirsch, 2009
Phenomenological state (empowered with causal effects)	Kallio & Revonsuo**,** 2003, 2005

As far as the historical association between psi and hypnosis is concerned, it is true that many of the early classical mesmerists such as Puységur, Elliotson, and Janet believed strongly in the occurrence of the "higher phenomena of mesmerism" (community of sensations and travelling clairvoyance) and they attributed these to the somnambulistic trance stage of hypnosis (Dingwall, 1967; Gauld, 1992). By contrast, Bernheim, the leader of the Nancy school of hypnosis, which was to become the ancestor of the social-neurocognitive theory of hypnosis, declared that he had never witnessed a single genuine case of paranormal experience (Dingwall, 1967; Forrest, 1999). The same dismissal of hypnotic psi was also true of Braid, who coined the term "hypnotism" and whose efforts in relating it to a sleep state established some scientific respectability for hypnosis (Braid, 1852, p. 118).

Little seems to have changed in the polarized attitudes to the subject since then. There are still leading hypnosis researchers who dismiss the claims of hypnotic psi out-of-hand as magical thinking (Nadon & Kihlstrom, 1987) while others give a degree of credibility to the evidence (Uneståhl, personal communication, August, 2003; Wickramasekera, 1989). My own survey suggests that the phenomena, although not particularly welcomed nor sought after, are still observed (Parker, 2003a).

That the issue of whether or not hypnosis facilitates psi still remains largely unsettled is often attributed to the abandonment of the use of hypnosis in the early days of Rhine's laboratory at Duke University. This decision was, however, not based on a dilettante effort from Rhine's side. Rhine was well aware of the dramatic historical anecdotes of psi associated with hypnosis, and he believed in their suggestive evidential value (Rhine, 1952). Moreover, Rhine had the supreme advantage of a hypnotically versed colleague and professor at Duke, Helge Lundholm, who had emigrated from Sweden, where hypnosis had for a brief period achieved a degree of respectability amongst physicians such as Alfred Backman, Poul Bjerre,

and John Björkhem. Later, Rhine even invited Björkhem to Duke because of his outstanding track record with hypnotic psi, but with the discouraging outcome that Björkhem was unable to reproduce this success during his visit (Pope, 1952; Stolt & Björkhem-Bergan, 2004). There appear to be several possible reasons for this: the language and cultural differences and that Rhine and Björkhem failed to develop a positive working relationship (Johnson, 1998).

There is perhaps a rather ironic explanation of Rhine's abandonment of hypnosis that is revealed in his detailed descriptions of the almost trance-like appearance of high-scoring subjects when they were performing as such in the laboratory. These descriptions suggest that it could well be the case that most of the star ESP performers were already during testing in an induced "state" of high absorption or self-hypnosis. There are striking descriptions of them being fully absorbed in inner experiences and detached from outer stimulation (for a full description see Parker, 2003b, p. 69). If this is so, then it might be the case that formal hypnosis would have become superfluous. This of course begs the question that is of central concern here: Is hypnosis a specifically identifiable altered state of consciousness that facilitates specific abilities?

Is Hypnosis an Altered State?

In further reviewing the hypnosis research, parapsychologists who are despondent over the failure to make headway in the debate over psi can take some comfort in the fact that the 60 years of intensive research into the existence or nonexistence of a trance state has not led to any unequivocal resolution of the issue. It is of course true that there is some tacit agreement as far as recognizing that hypnosis does produce valid experiential and phenomenological changes in the hypnotized person (Kirsch, 2011). Although some authorities might regard this agreement as progress, the cynic might wonder why it took 60 years to reach an apparently self-evident conclusion. Moreover, move beyond this point and an impasse is reached between the so-called "credulous view," in which a participant's testimony is given credibility, and the "skeptical view," in which this testimony is a mere product of the suggestions and the context (McConkey, 2008; compare Kihlstrom, 2003; Kirsch, Mazzoni, & Montgomery, 2007).

The contemporary heir to the traditional trance theory of hypnosis is the theory of neodissociation. The theory maintains that the bizarreness of hypnosis is simply due to the executive decision-making function of the individual being to some degree taken over and directed by the hypnotist—almost as a mild form of possession. Indeed, Hilgard, the originator of this theory, described the residual part of the person metaphorically as the "hidden observer" (Hilgard, 1992). The theory supposes that there is an amnesic barrier for the hypnotic process, which means that actions carried out during it become dissociated from conscious awareness. As a result, these actions are no longer regarded as part of the person's own sense of agency and are experienced as involuntary and alien ones. In order to explain away these actions, the posthypnotically re-established executive self will begin to "invent" or confabulate delusory beliefs in order to explain what happened. It is for this reason that hypnosis provides the ideal means of studying how various disturbances such as alien abductions and dissociated identity disorder arise (Kihlstrom, 2013).

Several variations of the neodissociation theory are now on the market, depending on how all-encompassing the "possession" is deemed to be. One of these, dissociated control theory, focuses on how the lower systems of routine actions, which are relatively easy to elicit by suggestions, surreptitiously bypass executive control (Bowers, 1992). Seen this way, dissociation is an integral part of normal life. Hypnosis becomes then less remarkable, because many actions in normal life are delegated to routine habits and multitasking and may even result in memory recall difficulties that are similar to hypnotic effects.

There is a further contemporary reinterpretation of dissociation theory that is influential and worthy of note. This is in terms of neuropsychology, and forms a program of research by Gruzelier. According to Gruzelier, the fixation of attention and gaze that was featured in classical hypnosis and Hollywood movies works because it reduces the activity of the thalamo-limbic connections to the frontal lobes and inhibits the activity of the anterior cingulate cortex. The anterior cingulate serves as an inner regulator for the focusing of attention, monitoring for errors, and possibly even the production of self-awareness. The hypnotic instructions typically promote a passive form of imagery in highly susceptible individuals, which leads to the activation of the right temporal lobe (Gruzelier, 2005, 2006). This resonates with the historical belief of

hypnosis being linked to the sleep-onset period, now described by Gruzelier in terms of alpha waves being gradually replaced by the theta waves of the sleep-onset period. However, hypnosis is evidently not identical with stage one sleep because none of the other indices, such as k-complexes, are present.

The opposing skeptical position to that of the dissociation theories derives from the well-known work of T. X. Barber describing hypnosis as a form of role-playing. Many, if not all, of the traditional dramatic effects of hypnosis were replicated by Barber with motivated and imagination control groups who had not gone through any formal hypnotic induction. The outcome of the studies of Barber and co-workers is succinctly and poignantly summarized by saying these "have compared hypnotic and control groups in their willingness, amongst other things, to make slanderous remarks, plunge their hands in a beaker containing acid and throw acid at the experimenter, mutilate the bible, cut up the national flag, make homosexual advances, steal, and deal in heroin, and no convincing evidence has emerged that hypnotic subjects are more likely to obey such instructions than control subjects who are subjected to the same demands and pressures" (Heap, 2008, p. 748).

Despite this assertion, there might be some exceptions to the unlimited success of simulators and control groups. One striking example is that given by Evans and Orne (1971), who used the scenario of "the disappearing hypnotist." When the hypnotist was unexpectedly called away, participants who were simulating hypnosis immediately gave up their act, whereas the hypnotized participants continued considerably longer in what seemed to be an oblivious hypnotic state.

Two more general exceptions to the above relate to the involuntary aspect of hypnosis. Both these exceptions taken at face value suggest that hypnosis enables processes to occur that are dissociated from normal consciousness. One of these processes is so-called "trance logic," which means that highly hypnotized participants go along with illogical suggestions such as seeing a person in two places at the same time (McConkey, 2008). The second of these, the Stroop effect, is the automatic interference of color words when they are written in a color contrasting with the word for color, for example the word "red" written in the color blue. Normally it is a difficult task to say the word due to the contradictory meaning of the word reaching our executive functions, but during hypnosis the executive interference can apparently be held in abeyance or enhanced by hypnotic suggestions as a form of cognitive inhibition or disinhibition (Lifshitz, Bonn, Fischer, Kashem, & Raz, 2013). This approximates the standard hypnotic demonstration in which the hypnotized person is successfully instructed to forget a certain number on returning to the normal waking state. Nevertheless, the claim that these effects are exclusive to a hypnotic state has been contested, because waking suggestions can apparently also produce the effect (Raz, Kirsch, & Nitkin-Kaner, 2006). Certainly, there are less dramatic examples of such hypnotic effects that are not hard to find in ordinary life. Most of us have experienced not being able to recall a name and found that the harder we tried to do so the more difficult it became, only to experience later that when "disinhibition" was established, the name came spontaneously into consciousness.

Barber's theory forms a cornerstone to the broad category of the social-neurocognitive theories of hypnosis. The most influential of these is compliance theory, developed by Spanos, who viewed hypnotic role-play as more pervasive than it is in Barber's theory. Spanos believed that the hypnotized person consistently, without formal awareness, takes cues from the social situation as to how to behave. Following "hypnosis," the executive self then disowns the behavior that is irreconcilable with the normal self-image and thus misattributes this as being due to the hypnotic trance (Spanos & Coe, 1992; Wagstaff & Cole, 2005). A fairly strong relationship between compliance and hypnotic suggestibility has been found (Polcyzk & Pasek, 2006).

Compliance theory derives from the well-known work of Milgram, who found that as many as 65% of ordinary individuals would under pressure from an authority figure administer potentially lethal shocks to stooges pretending to be participants in experiments on learning. Although this as a theory of hypnosis naturally gives no credibility whatsoever to the existence of a hypnotic state, it is worth noting that Milgram himself thought that his extraordinary results were due to the "agentic state," by which he meant that the compliant individual performs an automaticity of actions in an altered state of consciousness (Milgram, 1974, p. 134).

Common sense dictates that it would be necessary to do violence to the limits of such concepts as compliance and role-playing alone in order to get them to explain the extraordinary surgical successes during the mid-1800s of Esdaile in India and witnessed by Elliotson in London. Major surgical amputations were carried out on patients using hypnotically induced anesthesia (reviewed in Gauld, 1992; Playfair, 1987/2011). Even if cultural and historical factors are considered to play a role here, there are contemporary studies indicating that presurgical hypnosis treatment leads to a reduction in the amount of anesthetic needed and promotes more rapid healing (reviewed in Enqvist, 1996; Wobst, 2007). During the 1980s, Finer used hypnotic suggestions prior to surgery instead of an anesthetic. He described clinical hypnosis as a means of creating dramatic placebo effects via the encouragement and training of relaxation and detachment from the pain (Finer, 1988).

It is here, concerning the generality of placebo effects, that one of the most influential social-neurocognitive theories of hypnosis has come to play a decisive role in the research literature. The major contributor in this context is Kirsch, who became internationally known for his work indicating that the effects of antidepressives are largely due to placebo effects (Kirsch, 2009). Working often with Lynn, Kirsch is also one of the most prolific publishers of research papers indicating that hypnosis exploits the same means of making expectancies work as placebos do. The difference is that in the case of hypnosis, the participants are not being deceived but are told directly what they are to believe. This is not to say that the expectancies are consciously formulated, but that they achieve their goals by a form of what Kirsch calls "automaticity"—a form of nonconscious conditioning in fulfilling the demands of the hypnotic situation without the person needing to think (Accardi et al., 2013; Kirsch & Lynn, 1995, 1999; Lynn, Vanderhoff, Shindler, & Stafford, 2002).

As with the perennial debate over psi, it is disillusioning for those holding a belief in the power of science to resolve issues, that 60 years of hypnosis research has failed to settle the basic dispute over the existence or nonexistence of a trance state (Accardi et al., 2013; Kallio & Revonsuo, 2003; Kirsch, 2005; McConkey, 2008). During these 60 years, three major handbooks in the research literature have been published bringing together different theorists who often disagree with one another (Fromm & Nash, 1992; Fromm & Shor, 1972; Nash & Barnier, 2008)**.** As is often the case in psi research, middle positions are seldom tolerated, so that Fromm and Nash in reviewing the research wrote, "The polarization of the field has so tainted the methodologies and the statistical interpretations of these experiments that it impossible to sit comfortably in either of the two camps" (Fromm & Nash, 1992, p. 42). Worse, like the "psi wars," the debate about the hypnotic state has become so intense as to depart from rules of détente. Whereas the critics of parapsychology have been described as "wolverines" for their aggressiveness (Cardeña, 2011), the "skeptics" in the area of hypnosis who dismiss the existence of a hypnotic state have been called "hedgehogs" for their limited perception and unfairness (Kihlstrom, 2003). Also reminiscent of the psi wars, skeptics in the debate readily cite the dictum "extraordinary phenomena require extraordinary proof" (Lynn, Fassler, & Knox, 2005).

Mapping Out the Jungle

Even in the absence of an agreed-upon theory of hypnosis, there is at least a tacit agreement on the gold standards for the instruments needed for mapping out "hypnotizability." These standards take the form of the Stanford Hypnotic Susceptibility Scale (SHSS) and a similar scale developed for group testing, the Harvard Group Scale of Hypnotical Susceptibility (HGSHS). Nevertheless, the influence of the different theories on the instrument of choice is very apparent. In accordance with state theories, a brief hypnotic induction is part of the administration of the SHSS and HGSHS, but other scales exist, such as the Barber Suggestibility Scale, which, in accordance with the belief that trance states do not really exist, skips entirely the hypnotic induction.

One of the major weaknesses of the hypnotic state concept is the tautology of defining hypnosis as a special state in which unusual behavior occurs and then explaining the unusual behavior by means of the hypnotic state. It is for this reason that considerable effort has been put into finding strong correlates of hypnotizablity with known psychological measures.

The most obvious candidate for a strong correlate is suggestibility. Waking suggestibility explains about 64% of the variance in tests of hypnotic suggestibility (Kirsch, 1997, p. 214), but is hypnotic suggestibility the same as hypnotizability? It has been argued that because this value approximates the reliability of hypnotizability tests, waking suggestibility does not contribute to the further understanding of hypnotizability (Kirsch, 1997; Kirsch et al., 2007; Raz, 2007). Moreover, it has been known for some time that there are several diverse types of suggestibility, and it still remains unclear as to how they relate to one another (Halligan & Oakley, 2013, 2014; Tasso & Péerez, 2008). Despite all these misgivings "waking suggestibility" and its responses were agreed upon as the only aspects of the definition of hypnosis that the American Psychological Association and researchers could reach a consensus on (Green, Barabasz, Barrett, & Montgomery; 2005; Kirsch et al., 2011). The underlying problem with any definition of hypnosis is, of course, the ineffability present in defining altered states, as there is no generally accepted definition of normal consciousness.

A further prime candidate for getting a handle on hypnotizability is *absorption* as measured by Tellegen's Absorption Scale (Tellegen & Atkinson, 1974). Absorption is defined rather vaguely as "openness to experience emotional and cognitive alterations across a variety of situations" (Roche & McConkey, 1990, p. 567). The lack of precision is due to absorption being regarded as a personality trait related to openness to inner experience, feelings, and fantasy (Glisky, Tataryn, Tobias, Kihlstrom, & McConkey, 1991; Wild, Kuiken, & Schopflocher, 1995), as well as being a cognitive skill relating to vividness of imagery and attention (Roche & McConkey, 1990). Absorption is estimated to account for about 10% of the variance in hypnotic suggestibility, but even this may be an overestimate because the context in which the tests are given influences such estimates. If participants are expecting hypnosis to follow, they then score higher on the absorption scale (Council, Kirsch, & Hafner, 1986). When the tests are given on separate occasions, the common variance is reported to shrink to about 1% (Dienes et al., 2009). Yet there seems to be an experimenter effect here: Not all researchers have found this to be the case (Nadon, Hoyt, Register, & Kihlstrom, 1991).

A more reliable correlate of hypnotizability than absorption seems to be that provided by a measure derived from Barber and Wilson's concept of *fantasy proneness* (Lynn & Rhue, 1986; Wilson & Barber, 1983). Fantasy proneness offers a moderate relationship accounting for about 9% of the variance. What makes the relationship seem more reliable is that it is apparently maintained even after context effects and similar items common to both scales are eliminated (Green & Lynn, 2008; Lynn, Kirsch, & Hallquist, 2008).

Turning now to Kihlstrom's dissociation theory of hypnosis, dissociation itself should be an obvious contender for a close relationship to hypnotizability. Dissociation is defined as "one or more parallel paths or systems operating outside of awareness and influencing cognition, affect, or behavior" (Kirsch & Council, 1992, p. 275) and is usually measured by a self-report scale, the Dissociation Experiences Scale (DES). The DES was developed as a clinical instrument to identify pathological dissociation, with the result that when it is used on a normal population the data become extremely skewed or even bimodal. Contrary to what might be expected from the altered state theories, the correlations between the DES and measures of hypnotizability have consistently been nonsignificant or small (Silvia & Kirsch, 1992). Worse, this may merely reflect an underlying relationship with absorption, because absorption itself relates independently to both and dissociation and hypnotizability (Whalen & Nash, 1996).

Magnetic Resonance as a Way Out?

As dream research remained in a similar position to hypnosis research until the discovery of neurophysiological correlates, there was much hope that the same thing might revolutionize and legitimatize the hypnosis field. There was already in the 80s and 90s some pioneer work that found relative shifts in activity amongst high hypnotizables from the left to right parieto-occipital areas, along with an increase in theta (Crawford, Clark, & Kitner-Triolo, 1988; Uneståhl & Blundzen, 1996). Unfortunately, the lack of adequate control groups meant that we could never know if the same neural underpinnings could have been found with groups using imagination and relaxation.

There was much anticipation that the application of neuroimaging to hypnosis would provide a more precise form of compass to localize the hypnotic state and transform the area into respected neuroscience (Kihlstrom, 2008, 2013). A landmark in this quest is said to have occurred in 2003 at the Tennessee Conference on Brain Imaging and Hypnosis, where the now much acclaimed findings of Rainville and colleagues were presented. The authors reported that specific changes occurred in cerebral blood flow during deep hypnosis, affecting activity in the anterior cingulate cortex, the thalamus, and the brain stem links with the right occipital cortex. Because these areas involved the noradrenergic system, the changes were interpreted as representing the relaxation and absorption components of the hypnotic state (Rainville, Hofbauer, Bushnel, Duncan, & Price, 2002; Rainville & Price, 2003). The hypnotic state could at last be clearly defined as a state of absorption and relaxation along with a loss of the executive functions of the self, these being held in abeyance until directed by the hypnotist's task suggestions. Several authorities now have the confidence to claim that there is now a solid case for the existence of hypnosis as an altered state (Barabasz & Barabasz, 2007; Kihlstrom, 2013).

The neurocognitive explanation of hypnosis came at a politically optimum time, as it fits with the current collegial interest in "mind-wandering," regarded as the default mode network (DMN) of the brain (Mason et al., 2007). Moreover, a recent fMRI study carried out under the aegis of Kirsch did indeed find that high hypnotizables showed a shift during hypnosis towards deactivation of some parts of the DMN (McGeown et al., 2009). Although the results are open to several interpretations, the favored one is that the high-hypnotizable individuals could *both* eliminate distracting stimuli during relaxation and increase their concentration during the period when specific hypnotic instructions were given (McGeown et al., 2009).

Yet, the question which the endorsers of the viewpoint of this MRI "proof" of hypnosis avoided asking is the one which is recurrent throughout this review: Would the same changes have occurred if hypnotic induction was replaced by induced relaxation and positive expectancies (Oakley, 2008)? If so, this would minimize, if not trivialize, the hypnotic "trance" as being a rather nonspecial state. The above neuropsychological correlates provide little help dealing with this issue. It has been known for many decades that the anterior cingulate cortex is involved in the focusing of attention on tasks, so this aspect is hardly startling or proof of hypnosis being a special state.

It might be thought then that a control group would resolve the issue but there are some inherent difficulties with this solution. Some members of the nonhypnotic control group may have the ability to spontaneously enter or "drift into" an altered state during any relaxation procedure. Although the claim that such spontaneous trances occur has been strongly disputed on empirical grounds (Kirsch et al., 2008; Mazonni, et al., 2009), if these do occur then the use of between-subjects designs as separate control groups would be a potentially misleading way of deciding the issue, because members of the control group might spontaneously drift into trance, thereby more or less nullifying the comparison. Whatever the case, the claim has been sufficiently influential to lead hypnotic state researchers and also hypnotic psi researchers to abandon the usual control groups in favor of designs where participants perform as their own controls, or to compare groups preselected on low versus high hypnotizability scores.

A further issue of concern is that the results of neuroimaging procedures such as the fMRI and PET, unless they are scored blindly, become the modern day equivalent of a Rorschach inkblot test. Several experts on using these techniques in hypnosis research have expressed their reservations about the claims being made from brain-imaging techniques on the grounds that there are often innumerable data points to work with (Mazzoni et al., 2012; Ray & Oathes, 2003; Raz, 2011).

More Illusory Exits?

Many readers may have paused by now and wondered if the above dichotomy of opinion over the existence of a hypnotic state is a false one. Why can the issue not be resolved by giving the "prize" to all the contestants through defining the hypnotic state as changes encompassing relaxation, expectancies, compliance, and imaginal abilities? This is certainly the viewpoint of some state supporters (Kihlstrom, 2005, 2008; McConkey, 2008). Yet for many critics, in particular Lynn and Kirsch, it would be glossing over the disagreement to avoid asking a deeper pertinent question: Is this "phenomenological state" a causal one

specific to hypnosis with its own influences on other psychological processes? In other words is this state merely an incidental epiphenomenon produced by the effects of compliance and suggestion itself, or does it actually have the causal power to steer and facilitate hypnotic behavior?

This was the starting point for what seemed to be a relentless attempt at resolving this deeper issue that came to occupy the journal *Contemporary Hypnosis* between the years 2003 and 2005 and which led to important conceptual developments for future hypnotic psi research. The lead article by Kallio and Revonsuo argued that the historical record indicates that the true altered state is actually a rare phenomenon limited to the so-called "somnambulists" and hardly equivalent to today's students recruited by psychology departments. The authors coined the term "virtuoso" in the search for the contemporary equivalent of the somnambulist (Kallio & Revonsuo, 2003). For the virtuoso, the difference is between hallucinating an experience and merely imagining it, as most participants might do (Kallio & Revonsuo, 2003, p. 130). It was said that if expectations are as powerful as the skeptics claim, when normal people misplace their keys they would see hallucinations of the keys at the place the keys were expected to be. This failed to impress Lynn and his co-workers, who argued that a powerful *context* is first required before we can engage our hallucinatory powers (Lynn et al., 2005). Thus, people occasionally report hallucinations in a religious context but rarely in their own house.

Nevertheless, the rejuvenation of the somnambulists as virtuosos has gained support from an unexpected posthumous source. Theodore Barber in one of his last papers redefined his position by proposing that there are three types of individuals susceptible to hypnotic induction: those prone to expectancy, those prone to fantasy, and those prone to dissociation (Barber, 1999). Some support for this typology was found in a study of highly hypnotizable people who could conveniently be divided into fantasy-prone and dissociation-prone. Whereas the fantasizers appeared to be on the whole a healthy group, many of the high dissociaters had a history of trauma and abuse (Barrett, 2010). In an ironic turn of events, it is now Barber's earlier skeptical position concerning the evidence for hypnosis as an altered state that is being used against this viewpoint: State-skeptics continue to demand comparisons of the performances of virtuosos with those of other individuals who are instructed to succeed by using their imagination (Wagstaff & Cole, 2005).

Before the reader despairs of any progress being possible in the hypnosis field, it now seems appropriate to reveal that in 2011 some progress was actually achieved. This followed an fMRI study of the changes that accompanied suggestions for alterations in color perception that Kirsch and his colleagues carried out. During this procedure highly suggestible participants gave state reports of when they went in and out of hypnosis. The reports coincided with changes in the fMRI that were consistent with alterations in the default or mind-wandering mode of the brain. The observation that these changes were significant only for those high suggestible participants who reported experiencing hypnosis caused Kirsch to admit openly that he had become more "agnostic" as to the existence of the hypnotic state. It was indeed a shift of position but only a small one (Kirsch, 2011). Kirsch and his co-workers continue to regard hypnosis as a form of waking suggestibility and liken hypnosis to a "direct placebo effect" (Accardi et al., 2013).

Is the 60 Years War Over the Hypnotic State an Experimenter Effect?

The reader may recall that the work of Kirsch and his co-workers originally supported the view of hypnosis as being explained solely on the basis of expectancies and the nondeceptive placebo effect (Accardi et al., 2013; Kirsch, 1994). The use of "expectancy" might be thought to have raised a central issue concerning experimenter expectancies as it has done in parapsychology, but instead the issue has been strangely muted in the 60 years of hypnosis research. There is a near absence of any discussion about the possibility that the expectancies of the experimenters themselves can influence the performance of participants to produce phenomena in accordance with the experimenter's theories (Schor, 1972, pp. 39–40 is a singular exception). This is truly remarkable, if not entirely enigmatic, given that demand characteristics and expectancies are a firm part of the hypnosis lore. In view of the proven potency of such experimenter effects and demand characteristics, it seems then that we should not be shy about raising the question of whether an experimenter effect could sometimes be at work.

There is a similar study to that of Kirsch and his co-workers carried out by other researchers who instead of expectancy and placebo effects favored the altered state theory of hypnosis. This study was highly compatible with the placebo study (in which the word placebo was essentially substituted for hypnosis) but here the comparison between two otherwise identical procedures involved one being labeled "relaxation" and the other "hypnosis." On this occasion there was a marked significant difference favoring "hypnosis" (Ghandi & Oakley, 2005).

There is an additional striking example of the apparent contradictory findings, perhaps due to experimenter effects, concerning the phenomenology of hypnosis. This can be crucial, because hypnosis researchers only agree on this one aspect, that participants report a subjective experience of being hypnotized.

How then do highly hypnotized persons describe the hypnotic state? Kirsch and co-workers found that their expert judges could not reliably distinguish the descriptions given by the hypnotic induction group from those given by participants performing imagination and relaxation exercises (Kirsch, Mobayed, Council, & Kenny, 1992). Yet a very similar study by Cardeña using the same Phenomenology of Consciousness Inventory found that "hypnotic virtuosos experience consistent and significant alterations of consciousness, even after controlling for relaxation" (Cardeña, 2005, p. 51). This clearly underscores the need for cooperation between critics and supporters in hypnosis research (as well as in psi research).

Let's return to using the "Rhineland jungle" analogy. When lost in the forest or jungle, the traditional advice is to try three strategies: (a) to attempt to gain a treetop overview of the situation, (b) to follow the outward flow of a stream, and (c) to retrace steps to the starting point and look for missed cues. These symbolic solutions are readily applicable here.

In the case of the first strategy, gaining an overview of the issues can reveal some of the gratuitous assumptions guiding research which turn out to be misleading. Here the overview can reveal the philosophical assumptions influencing the outcome of experiments.

It is also instructive to examine the direction of flow of the "main stream" of research. When this is done in parapsychology, it seems undeniable that the flow of the main "stream," rather than leading out from the forest, has become stagnated by correlational research. The illusion is created where mainstream publications are constantly fed by the flow of correlations that are measuring the interactive force of psi with surrounding factors, whereas on closer examinations, most correlations represent only temporary maelstroms in guessing behavior.

Finally, as was earlier raised in introduction to this topic, some vital clues might be found by retracing our steps seeing what focusing on the ganzfeld and hypnosis exclusively with ordinary psychology students may have lost.

An Overview of Issues

In changing now the perspective to that of how the hypnotic state controversy fits into psychology and neuroscience, one common hindrance to progress stands out. This is the well-recognized difficulty in defining consciousness and its associated altered states in a precise, noncircular, and objective manner (Farthing, 1992). Despite the fact that a chapter on consciousness and its altered states is today regarded as obligatory for any marketable book on psychology, the above difficulty promotes a shyness in psychology for using the term "consciousness" in any explanatory fashion or attributing any phenomena to it in more than a perfunctory manner (Parker, 2014).

Remarkably, the same reluctance does not apply to the use of the term "cognitive unconscious," a concept that Kihlstrom is a major architect of and which recently became popularized by psychology's contemporary Nobel prize winner (Kahneman, 2011), as "system 1"—as opposed to "system 2" which requires conscious steering in order to function**.** One reason for the popularity of the cognitive unconscious is that it is consistent with reductionist neuropsychology.

Today's reductionist neuropsychology owes much of its persuasive power to the discovery that "Bereitschaftspotential" (readiness potential) occurs in the brain prior to conscious decision-making. Data taken from the medial frontal cortex using the fMRI appear to anticipate the conscious decision to make a movement up to 10 s before that decision (Haynes et al., 2007). In neuroscience circles this has been

enthusiastically and perhaps uncritically received, given that the effect is only 10–20% above chance. Nevertheless, taken at face value the claim would mean that unless you believe in backward causation, the belief in conscious decision-making is illusory and largely redundant as a causal psychological explanation. Of course, there are explanations of Haynes' findings that emphasize the imprecision in the process of decision-making. If findings from psi research are taken seriously then the concepts of presentiment and dissociation may also add to the imprecision in identifying the exact timing of mental events.

Whatever the case, the emergence of this neuroscience reductionist ethos may explain some of the diehard resistance towards giving credibility to hypnotic states. The ethos is also consistent with the contemporary effort to redefine hypnosis as a "nondeceptive placebo." The nondeceptive placebo effect was part of a thought-provoking hypnosis study that compared the effect of two induction procedures: One procedure involved a standard hypnotic induction while the other essentially substituted the word "placebo" for the word "hypnosis" and defined placebo as a form of cooperation. Having found that both procedures in effect produced the same results on the Phenomenology of Consciousness Scale, Kirsch and Lynn and their co-workers concluded: "Our research does provide preliminary evidence for the general equivalence of these very different rationales and suggests that it is possible to define hypnosis as a placebo and not risk attenuation of hypnotic responsiveness" (Accardi et al., 2013**,** p. 113).

The apparent paradox in this reformulation of hypnosis is that placebo effects are usually conceived as demonstrating the supreme importance of mental states. Given this, it would then seem almost contradictory that Kirsch should focus on the importance of placebo effects while at the same time denying that hypnosis is a special state that can influence brain and bodily processes. The apparent inconsistency stretches further if we consider that the concept of "automaticity" is used by Kirsch as the explanation of placebo effects. Automaticity refers to expectancies and placebo suggestions, which, after having for the most part circumvented the individual's awareness, gain an implicit power to bring about remarkable changes in behavior (Kirsch & Lynn, 1999). Automaticity as a concept delivers certain winnings: It may be possible to use it to form an allegiance or even an amalgamation with Kihlstrom's dissociation theory, but its acceptance comes at major cost. It fails to give any recognition to the probable role of heightened states of consciousness and to openness in integrating new beliefs into psychoneuroimmunological processes that are generally conceived of as being central to the placebo effect (Ray, 2014). Moreover, Ogden, a major authority in this area, has drawn attention to the unresolved philosophical contractions implicit in any reductionist theory of placebo effects. She writes: "If an individual's psychological state can influence their health, then perhaps mind and body should not be seen as separate entities but as interacting. This, in part, contradicts dualistic models of the individual. However this interaction still assumes that mind and body are distinct: to interact with each other, they still need to be defined as being separate" (Ogden, 2007 p. 308). A review of the evidence from neuroimaging studies of emotional self-regulation, psychotherapy, and the placebo effect, led Beauregard (2007, p. 233) to an even more explicit conclusion that "by changing our mind we are changing our brain."

It seems clear that there are philosophical differences between major researchers, which appear to be ultimately based on different views of the mind-body relationship. It may then well be the case that research teams have different expectancies when carrying out their studies and these may influence how they deal with participants.

Conclusions

We need to take stock. Given that 60 years of research into settling the issue of whether hypnosis is an altered state has made only unsubstantial progress, there remain enormous and seemingly insurmountable difficulties to be surpassed in order to clear the way through the "Rhineland jungle" and find the searched-for psi-conducive hypnotic state.

The above overview of the denigration of the importance of altered states in contemporary psychology should discourage any belief or assumption that research has established hypnosis as being a specific altered state of consciousness or indeed as one with state specific effects. Moreover, as we noted above, experimenters with different expectancies concerning the nature of hypnosis when carrying out essentially the

same study often obtain quite different results. In short, hypnosis research is beset with the same problem concerning experimenter effects that has restricted progress in parapsychology (Parker & Millar, 2014).

In dealing with the seemingly intractable issue of what hypnosis is, it becomes impossible to avoid confronting the ultimate problem of the nature of consciousness. A current article co-authored by Lynn (one of Kirsch's most consistent collaborators) expresses a view that seems to tackle this in an arguably more progressive way (compared to previous standpoints). In considering the implications of the hypnosis research for a theory of consciousness, the authors write: "A fundamental assumption of the work on hypnosis is that consciousness is pliable, capable of being perturbed and channeled by suggestions, mental images, memories, and focused associations to achieve therapeutic objectives" (Green, Laurence, & Lynn, 2014, p. 204).

As far as hypnotic psi is concerned, in view of the demonstrable complexity of hypnosis, we may have to face that the search for a specific state or for a set of reliable personality variables, which will relate to psi performances in any predictable and replicable way, may turn out to be a folly initiated by fool's gold. On the other hand, it seems only fair to emphasize that even if a specificity cannot be identified, the more dramatic phenomena of hypnosis seem to illustrate the potency of altered belief systems to effect major changes in psychological processes.

However, it would be premature to reach any conclusion before we have fully explored the two remaining strategies concerning the attempts by research to make progress in the area of hypnotic psi. As was suggested earlier in the context of analogy with the Rhineland jungle, one strategy that is important is to follow and examine the flow formed by the "main stream" of research. The other is to re-track and take a renewed look at some of the historical work on hypnosis for clues that might revitalize research with hypnosis and psi. We continue with a review of these focal points in Part 2.

References

Accardi, M., Cleere, C., Lynn, J., & Kirsch, I. (2013). Placebo versus "standard" hypnosis rationale: Attitudes, expectancies, hypnotic responses, and expectancies. *American Journal of Clinical Hypnosis*, *56,* 103–114.

Barabasz, A. F., & Barabasz, M. (2007). Hypnosis and the brain. In M. R. Nash and A. J. Barnier (Eds.), *Oxford handbook of hypnosis: Theory, research and practice.* Oxford, England: Oxford University Press.

Barber, T. X. (1969). *Hypnosis: A scientific approach*. New York, NY: Van Nostrand Reinhold.

Barber, T. X. (1999). A comprehensive three-dimensional theory of hypnosis. In I. Kirsch, A. Capafons, E. Cardeña-Buelna, & S. Amigo (Eds.), *Clinical hypnosis and self-regulation: Cognitive-behavioral perspectives* (pp. 21–48). Washington, DC: American Psychological Association.

Barrett, D. (2010). Dissociaters, fantasizers, and their relation to hypnotizablity. In D. Barrett (Ed.), *Hypnosis and hypnotherapy: Vol 1. History, theory and research* (pp. 77–96). Oxford, England: Praeger.

Beauregard, M. (2007). Mind does really matter: Evidence from neuroimaging studies of emotional self-regulation, psychotherapy, and placebo effect. *Progress in Neurobiology, 81*, 218–236.

Bowers, K. S. (1992). Dissociated control and the limits of hypnotic responsiveness. *Consciousness and Cognition, 1*, 32–39.

Buranelli, V. (1975). *The wizard from Vienna—Franz Anton Mesmer.* New York, NY: Coward, McCann, & Geoghegan.

Cardeña, E. (2005). The phenomenology of deep hypnosis: Quiescent and physically active. *International Journal of Clinical and Experimental Hypnosis, 53,* 37–59.

Cardeña, E. (2011). On wolverines and epistemological totalitarianism. *Journal of Parapsychology, 75*, 3–14.

Council, J. R., Kirsch, I., & Hafner, L. P. (1986). Expectancy versus absorption in the prediction of hypnotic responding. *Journal of Personality and Social Psychology, 50,* 182–189.

Crawford, H. J., Clarke, S. W., & Kitner-Triolo, M. H. (1988). EEG activity pattern differences in low and high hypnotizables: Reflections of cognitive strategy differences? Paper presented at the Fourth International Psychophysiology Conference, Prague, Czechoslovakia.

Dingwall, E. (1967). *Abnormal hypnotic phenomena*: *Vol. 1. Hypnotism in France*. London, England: Churchill.

Enqvist, B. (1996). *Pre-surgical hypnosis and suggestions in anesthesia*. (Doctoral dissertation). University of Stockholm, Sweden.

Evans, F. J., & Orne, M. T. (1971). The disappearing hypnotist: The use of simulating subjects to evaluate how subjects perceive experimental procedures. *International Journal of Experimental Hypnosis, 19,* 277–296.

Farthing, G. W. (1992). *The psychology of consciousness*. Englewood Cliffs, NJ: Prentice Hall.

Finer, B. (1988). Hypnosis and anaesthesia. In G. D. Burrows & L. Dennerstein (Eds.), *Handbook of hypnosis and psychosomatic medicine* (pp. 293–305). Amsterdam, Netherlands: Elsevier.

Forrest, D. (1999). *A history of hypnosis.* London, England: Penguin Books.

Fromm, E., & Nash, M. N. (Eds.). (1992). *Contemporary hypnosis research.* London, England: Guildford Press.

Fromm, E., & Shor, R. E. (Eds.). (1972). *Hypnosis—Research developments and perspectives*. Chicago, IL: Paul Elak.

Gauld, A. (1992). *A history of hypnotism.* Cambridge, England: Cambridge University Press.

Ghandi, B., & Oakley, D. A. (2005). Does hypnosis by any other name smell as sweet? The efficacy of "hypnotic" inductions depends on the label "hypnosis." *Consciousness and Cognition, 14*, 304–315.

Glisky, M. L., Tataryn, D. J., Tobias, B. A., Kihlstrom, J. F., & McConkey, K. M. (1991). Absorption, openness to experience, and hypnotizability. *Journal of Personality and Social Psychology, 60*, 263–272.

Green J. P., Barabasz, A .F, Barrett, D., & Montgomery, G. H. (2005). Forging ahead: The 2003 APA Division 30 definition of hypnosis. *International Journal of Clinical and Experimental Hypnosis, 53,* 259–264.

Green, J. P., Laurence, J-R., & Lynn, S. L. (2014). Hypnosis and psychotherapy: From Mesmer to mindfulness. *Psychology of Consciousness: Theory, Research and Practice, 2,* 199–212.

Green, J. P., & Lynn, S. J. (2008). Fantasy proneness and hypnotizablity: Another look. *Contemporary Hypnosis*, *25,* 156–164.

Gruzelier, J. (2005). Altered states of consciousness and hypnosis in the twenty-first century. *Contemporary Hypnosis, 22,* 1–7.

Gruzelier, J. (2006). Frontal functions, connectivity and neural efficiency underpinning hypnosis and hypnotic susceptibility. *Contemporary Hypnosis, 23,* 15–32.

Halligan, P. W., & Oakley, D. A. (2013). Hypnosis and cognitive neuroscience: Bridging the gap. *Cortex, 49*, 359–364.

Halligan, P. W., & Oakley, D. A. (2014). Hypnosis and beyond: Exploring the broader domain of suggestion. *Psychology of Consciousness: Theory, Research, and Practice, 1,* 105–122.

Haynes, J. D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, D. (2007). Reading hidden intentions in the human brain. *Current Biology, 17,* 323–328.

Heap, M. (2008). Hypnosis in the courts. In M. R. Nash & A. J. Barnier (Eds.), *Oxford handbook of hypnosis: Theory, research and practice* (pp. 745–766). Oxford, England: Oxford University Press.

Hilgard, E. R. (1992). Dissociation and theories of hypnosis. In E. Fromm & M. N. Nash (Eds.), *Contemporary hypnosis research* (pp. 69–101**)** London, England: Guildford Press.

Johnson, M. (1998). John Björkhem (1910–1963) In N-O Jacobson (Ed.), *Svensk Parapsykologi* (pp. 71–73). Kristianstad, Sweden: John Björkhem's Minnesfond.

Kahneman, D. (2011). *Thinking fast and slow*. London, England: Penguin Books.

Kallio, S., & Revonsuo, A. (2003). Hypnotic phenomena and altered states of consciousness: A multilevel framework of description and explanation. *Contemporary Hypnosis, 20,* 111–164.

Kallio S., & Revonsuo, A. (2005). Altering the state of the altered state debate: Reply to commentaries. *Contemporary Hypnosis, 22,* 46–55.

Kihlstrom, J. (2003). The fox, the hedgehog, and hypnosis. *International Journal of Clinical and Experimental Hypnosis, 53*, 166–189.

Kihlstrom, J. F. (2005). Is hypnosis an altered state of consciousness or what? *Contemporary Hypnosis, 22,* 34–38.

Kihlstrom, J. F. (2008). The domain of hypnosis, revisited. In M. R. Nash & A. J. Barnier (Eds.), *Oxford handbook of hypnosis: Theory, research and practice* (pp. 21–52). Oxford, England: Oxford University Press.

Kihlstrom, J. (2013). Neuro-hypnotism: Prospects for hypnosis and neuroscience. *Cortex, 49*, 365–374.

Kirsch, I. (1994). Hypnosis as a non-deceptive placebo: Empirically derived techniques. *American Journal of Clinical Hypnosis, 37,* 95–106.

Kirsch, I. (1997). Suggestibility or hypnosis: What do our scales really measure? *International Journal of Clinical and Experimental Hypnosis, 45*, 212–225.

Kirsch, I. (2005). Empirical resolution of the altered state debate. *Contemporary Hypnosis, 22,* 18–23.

Kirsch, I. (2009). *The emperor's new drugs: Exploding the antidepressant myth*. New York, NY: Basic Books.

Kirsch, I. (2011). The altered state issue: Dead or alive? *International Journal of Clinical and Experimental Hypnosis, 59,* 350–362.

Kirsch, I., Cardeña, E., Derbvshire, S., Dienes, Z., Heap, M., Kallio, S.,. . . Whalley, M. (2011). Definitions of hypnosis and hypnotizablity and their relation to suggestion and suggestibility: A consensus statement. *Contemporary Hypnosis and Integrative Therapy, 28*, 107–115.

Kirsch, I., & Council, J. R. (1992). Situational and personality correlates of hypnotic phenomena. In E. Fromm, & M. N. Nash (Eds.), *Contemporary hypnosis research* (pp. 267–291). London, England: Guildford Press.

Kirsch, I., & Lynn, S. J. (1995). The altered state of hypnosis—Changes in the theoretical landscape. *American Psychologist, 50,* 846–858.

Kirsch, I., & Lynn, S. J. (1999). Automaticity in clinical psychology. *American Psychologist, 54*, 504–515.

Kirsch, I., Mazzoni, G., & Montgomery, G. H. (2007). Remembrance of hypnosis past. *American Journal of Clinical Hypnosis, 49,* 171–179.

Kirsch, I., Mazzoni, G., Roberts, K., Dienes, Z., Hallquist, M. N., Williams, J., & Lynn, S. J. (2008). Slipping into trance. *Contemporary Hypnosis, 25*, 202–209.

Kirsch, I., Mobayed, C., Council, J. R., & Kenny, D. A. (1992). Expert judgments from subjective state reports. *Journal of Abnormal Psychology, 101*, 657–662.

Lifshitz, M., Bonn, N. A., Fischer, A., Kashem, I. F., & Raz, A. (2013). Using suggestion to modulate automatic processes: From Stroop to McGurk and beyond. *Cortex, 49,* 463–473.

Lynn, S. J., Fassler, O., & Knox, J. (2005). Hypnosis and the altered state debate: Something more or nothing more? *Contemporary Hypnosis, 22*, 39–45.

Lynn, S. J., Kirsch, I., & Hallquist, M. N. (2008). Social cognitive theories of hypnosis. In M. R. Nash & A. J. Barnier (Eds.), *Oxford handbook of hypnosis: Theory, research and practice* (pp.111–139). Oxford, England: Oxford University Press.

Lynn, S. J., & Rhue, J. W. (1986). The fantasy-prone person: Hypnosis, imagination, and creativity. *Journal of Personality and Social Psychology, 51*, 404–408.

Lynn, S. J., Vanderhoff, H., Shindler, K., & Stafford, J. (2002). The effects of an induction and defining hypnosis as a "trance" vs. cooperation: Hypnotic suggestibility and performance standards. *American Journal of Clinical Hypnosis, 44,* 231–240.

Mason, M. F., Norton, M. I., Van Horn, J. D, Wegner, D. M., Grafton, S. T., & Macrae, N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science, 315,* 393–395.

Mazzoni, G., Venni, A., McGeown, W., & Kirsch, I. (2012). Neuroimaging resolution of the altered state hypothesis. *Cortex ,49*, 400–410.

McConkey, K. M. (2008). Generations and landscapes of hypnosis: Questions we've asked, questions we should ask. In M. R. Nash & A. J. Barnier (Eds.), *Oxford handbook of hypnosis: Theory, research and practice* (pp. 53–80). Oxford, England: Oxford University Press.

McGeown, W. J., Mazonni, G., Vennen, A., & Kirsch, I. (2009). Hypnotic induction decreases default mode activity. *Consciousness and Cognition, 18,* 848–855.

Milgram, S. (1974). *Obedience to authority: An experimental view*. London, England: Harper & Row.

Munthe, A. (1929). *The story of San Michele.* London, England: John Murray.

Nadon, R., Hoyt, I., Register, P., & Kihlstrom, J. (1991). Absorption and hypnotizability: Context effects revisited. *Journal of Personality and Social Psychology, 60*, 144–153.

Nadon, R., & Kihlstrom, J. F. (1987). Hypnosis, psi, and the psychology of anomalous experience. *Behavioral and Brain Sciences, 10,* 597–599.

Nash, M. R., & Barnier, A. J. (Eds.). (2008). *Oxford handbook of hypnosis: Theory, research and practice.* Oxford, England: Oxford University Press.

Oakley, D. A. (2008). Hypnosis, trance and suggestion: Evidence from neuroimaging. In M. R. Nash & A. J. Barnier (Eds.), *Oxford handbook of hypnosis: Theory, research and practice* (pp. 365–392). Oxford, England: Oxford University Press.

Ogden, J. (2007). *Health psychology*. New York, NY: McGraw Hill.

Parker, A. (2000). A review of the ganzfeld work at Gothenburg University. *Journal of the Society for Psychical Research, 64,* 1–15.

Parker, A. (2003a). Hypnos och paranormala upplevelser: Finns det mer än ett historiskt samband? [Hypnosis and paranormal experiences: Is there more than a historical connection?] *Abstracts of the Nordiske Hypnosekongress / Nordic Hypnosis Congress*, Oslo, Norway.

Parker, A. (2003b). Psi and altered states. In L. Storm & M. Thalbourne (Eds.), *Parapsychology in the 21st century* (pp. 65–89). Jefferson, NC: McFarland.

Parker, A. (2014). Consciousness in mainstream books. Paper in preparation.

Parker, A., & Millar, B. (2014). Revealing psi secrets: Successful experimenters appear to succeed by using their own psi. *Journal of Parapsychology, 78*, 39–85.

Playfair, G. (1987/2011). *Medicine, mind & magic*. Wellinborough, England: Aquarian Press. Republished as *If this be magic.* Guildford, England: White Crow Press.

Polczyk, R., & Pasek, T. (2006). Types of suggestibility: Relationship among compliance, indirect, and direct

suggestibility. *International Journal of Clinical and Experimental Hypnosis, 54,* 392–415.
Pope, D. (1952). A Swedish explorer in parapsychology—Dr. John Björkhem. *Parapsychological Bulletin, 27*, 1–4.
Rainville, P., Hofbauer, R. K., Bushnell, M. C., Duncan, G. H., & Price, D. D. (2002). Hypnosis modulates activity in brain structures involved in the regulation of consciousness. *Journal of Cognitive Neuroscience, 14*, 887–901.
Rainville, P., & Price, D. D. (2003). Hypnosis phenomenology and the neurobiology of consciousness. *International Journal of Clinical and Experimental Hypnosis, 51,* 105–129.
Ray, O. (2014). How the mind hurts and heals the body. *American Psychologist, 59*, 29–40.
Ray, W. J., & Oathes, D. (2003). Brain imaging techniques. *International Journal of Clinical and Experimental Hypnosis, 5,* 97–104.
Raz, A. (2007). Suggestibility and hypnotizability: Mind the gap. *American Journal of Clinical Hypnosis, 49*, 205–210.
Raz. A. (2011). Does neuroimaging of suggestion elucidate hypnotic trance? *International Journal of Clinical and Experimental Hypnosis, 59,* 363–377.
Raz, A., Kirsch, I., & Nitkin-Kaner, Y. (2006). Suggestion reduces the Stroop effect. *Contemporary Directions in Psychological Science, 17*, 91–95.
Rhine, J. B. (1952). Extrasensory perception and hypnosis. In L. M. LeCron (Ed.), *Experimental hypnosis* (pp. 359–368). New York, NY: Macmillan.
Roche, S. M., & McConkey, K. M. (1990). Absorption: Nature, assessment, and correlates. *Journal of Personality and Individual Differences, 59,* 91–101.
Schor, R. E. (1972). The fundamental problem in hypnosis research as viewed from historic perspectives. In E. Fromm & R. E. Shor (Eds.), *Hypnosis—Research developments and perspectives* (pp. 15–42). Chicago, IL: Paul Elak.
Silvia, C., & Kirsch, I. (1992). Interpretative sets, expectancy, fantasy proneness, and dissociation as predictors of hypnotic response. *Journal of Personality and Social Psychology, 63,* 847–856.
Spanos, N. P., & Coe, W. C. (1992). A social-psychological approach to hypnosis. In E. Fromm & M. N. Nash (Eds.), *Contemporary hypnosis research* (pp. 102–130). London, England: Guildford Press.
Stanford, R., & Stein, A. (1994). A meta-analysis of ESP studies contrasting hypnosis and a comparison condition. *Journal of Parapsychology, 58*, 235–270.
Stolt, C-M., & Björkhem-Bergen, L. (2004). Hypnosis in Sweden during the twentieth century—The life and work of John Björkhem. *History of Psychiatry, 15,* 193–200.
Tasso, A. F., & Péerez, N. A. (2008). Parsing everyday suggestibility: What does it tell us about hypnosis? In M. R. Nash & A. J. Barnier (Eds.), *Oxford handbook of hypnosis: Theory, research and practice* (pp. 283–310). Oxford, England: Oxford University Press.
Tellegen, A., & Atkinson, G. (1974). Openness to absorbing and self-altering experiences (‘‘absorption’’), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology, 83*, 268–277.
Uneståhl, L-E., & Blundzen, P. (1996). Integrated mental training—neuro-biochemical mechanisms and psychophysical consequences. *Hypnosis, 23,* 148–158.
Wagstaff, G., & Cole, J. (2005). Levels of explanation and the concept of a hypnotic state. *Contemporary Hypnosis, 22,* 14–17.
Whalen, J. E., & Nash, M. (1996). Hypnosis and dissociation: Theoretical, empirical, and clinical perspectives. In L. K. Michelson & W. J. Ray (Eds.), *Handbook of dissociation* (pp. 191–206). New York, NY: Plenum Press.
Wickramasekera, I. (1989). Is hypnotic ability a risk factor for subjective (verbal report) psi, somatization, and health care costs? In L. Coly & J. D. S. McMahon (Eds.), *Psi and clinical practice*: *Proceedings of an international conference* (pp.184–197). New York, NY: Parapsychology Foundation.
Wild, T. C., Kuiken, D., & Schopflocher, D. (1995). The role of absorption in experiential involvement. *Journal of Personality and Social Psychology, 69,* 569–579.
Wilson, S. C., & Barber, T. X. (1983). The fantasy-prone personality: Implications for understanding imagery, hypnosis, and parapsychological phenomena. In A. A. Sheikh (Ed.), *Imagery: Current theory, research, and application* (pp. 340–390). New York, NY: Wiley.
Wobst, A. (2007). Hypnosis and surgery: Past, present, and future. *Anesthesia and Analgesia, 104,* 1199–1208.

Acknowledgements

The author wishes to thank in alphabetical order: Brian Millar, Nemo Mörck, Donald West, and two anonymous reviewers for their comments.

Department of Psychology
University of Gothenburg
Box 500
40530 Gothenburg, Sweden
Adrian.Parker@psy.gu.se

Abstracts in Other Languages

French

LA JUNGLE DU PSI HYPNOTIQUE: PREMIERE PARTIE.
LA RECHERCHE SUR L'HYPNOSE PERTINENTE POUR LE PSI

RÉSUMÉ : Les recherches contemporaines se sont efforcées d'établir un état hypnotique favorisant le psi à partir des notions d'hypnotisabilité et de dissociation. Cela vient en grande partie de l'héritage historique association les phénomènes hypnotiques et psi, une méta-analyse mal comprise et un certain nombre de variables communes. En réalité, l'hypnose conçue comme un état fait toujours l'objet de nombreux débats et un passage en revue de 60 ans de recherches intensives sur l'hypnose indique que l'hypnose serait plutôt à voir comme une combinaison complexe de divers facteurs : une « jungle de variables » plutôt qu'un état unitaire. Cela implique que la recherche d'un état spécifique favorisant le psi pourrait n'être qu'une entreprise vouée à l'échec. Toutefois, certains des efforts pour réduire l'hypnose à des attentes suscitées inconsciemment ou à de « l'automaticité » semblent refléter la tendance actuelle à dénigrer le rôle des processus conscients. Paradoxalement, les phénomènes les plus spectaculaires de l'hypnose pourraient plutôt illustrer l'aptitude des systèmes de croyance modifiée à modifier de façon majeure les processus psychologiques.

German

DER DSCHUNGEL DES HYPNOTISCHEN PSI: TEIL I.
PSI-RELEVANTE FORSCHUNG ZUR HYPNOSE

ZUSAMMENFASSUNG: Gegenwärtige Forschungsanstrengungen wurden in der Absicht unternommen, einen psi-förderlichen hypnotischen Zustand in erster Linie mit Hilfe von Hypnotisierbarkeit und Dissoziation nachzuweisen. Das meiste davon scheint auf eine frühe historische Verbindung zwischen hypnotischen und Psi-Phänomenen zurückzugehen, einer falsch interpretierten Meta-Analyse und einer Anzahl gemeinsamer Variablen. In Wirklichkeit ist die Auffassung von Hypnose als einem Zustand immer noch äußerst umstritten, und eine Übersicht über 60 Jahre intensiver Hypnoseforschung deutet darauf hin, dass Hypnose bestenfalls als komplexes Zusammenspiel verschiedener Faktoren gesehen werden kann – eher als ein „Dschungel aus Variablen" als ein einheitlicher Zustand. Die Suche nach einem spezifisch psi-förderlichen Zustand könnte sich daher als illusorisches Unternehmen herausstellen. Einige der Versuche, Hypnose auf unbewusst induzierte Erwartungen oder auf „Automatizität" zu reduzieren, scheint den gegenwärtigen Trend widerzuspiegeln, die Rolle bewusster Prozesse abzuwerten. Paradoxerweise können andererseits die dramatischeren Hypnosephänomene dadurch, dass sie grössere Änderungen bei psychologischen Prozessen bewirken, die Wirkmächtigkeit veränderter Glaubenssysteme illustrieren.

Spanish

LA SELVA HIPNÓTICA DE PSI: PARTE 1.
INVESTIGACIÓN EN HIPNOSIS PERTINENTE A PSI

RESUMEN: Los esfuerzos de investigación contemporáneos han intentado establecer un estado hipnótico propicio a psi principalmente en términos de hipnotizabilidad y disociación. Mucho de esto parece estar

fundado en la temprana asociación histórica entre los fenómenos hipnóticos y psi, un meta-análisis incomprendido, y una serie de variables compartidas. En realidad, el concepto de estado de hipnosis continúa siendo disputado y una revisión de los 60 años de investigación intensiva sobre la hipnosis indica que la hipnosis es una combinación compleja de varios factores: una "jungla de variables" más que un estado unitario. Esto implica que la búsqueda de un estado específico propicio a psi puede ser una empresa ilusoria. Sin embargo, algunos de los esfuerzos para reducir la hipnosis a expectativas elicitadas inconscientemente o "automaticidad" parecen reflejar la tendencia actual para denigrar el papel de los procesos conscientes. Paradójicamente, los fenómenos más dramáticos de la hipnosis pueden ilustrar el poder de sistemas de creencias alterados para efectuar cambios importantes en los procesos psicológicos.

Journal of Parapsychology, 79(1), 37–52

THE JUNGLE OF HYPNOTIC PSI: PART 2. RESEARCH ON RELATIONSHIPS BETWEEN PSI AND HYPNOSIS

By Adrian Parker

ABSTRACT: Research on hypnosis and on psi has led to disputable correlational findings. This review indicates all the contemporary studies have major methodological weaknesses and failed to produce replicable findings and to predict future psi scores. A critique is made of the correlational approach as largely relating guessing behavior to weak psychological variables of doubtful validity. The net outcome is that research into hypnotic psi, rather than fulfilling the promise of a functional relationship based on solid ground, has become bogged down with correlational variables, methodological flaws, and what appear to be nonreplicable findings. An argument is made for a refocusing of research efforts on the most promising variables, namely "positive dissociation" and "positive schizotypy." In the absence of full-scale sophisticated projects, the suggestion is that progress can be best made by the study of selected star participants or even so-called virtuoso performers. In this context the health versus pathology issue is a basic one that remains to be resolved. The issues raised are considered to be fundamental ones for a theory of consciousness.

Keywords: hypnosis, psi, schizotypal personality, trance, consciousness

In the first paper, it became clear that Rhine's assertion that hypnosis is a jungle of variables appears to be well-founded. The review of the 60-year dispute concerning whether or not the hypnotic state exists gave little prospect of penetrating the issues with these variables in the quest to find a psi-conducive state. It is proposed here that we now critically examine the actual attempts to make progress. This involves looking at the mainstream approaches within parapsychology and hypnosis research (or in jungle symbolism, trying to find a main stream through the jungle) and then reviewing the contemporary attempt to re-track approaches using clues left from the history of hypnosis in the belief that some important leads have been missed.

The Stagnated "Main Stream"

A major mainstream approach in the scientific development of parapsychology involves the focus of process research on finding significant correlations of psychological tests with psi scores. In contrast with the correlational approach, the original attraction of hypnotic induction and of the emergent ganzfeld methodology was the promise of discovering functional relationships. The ideal is that you do something to the participant and the result will be seen clearly in the scores of the psi test. With the ganzfeld, this "something" was hypothesized at various times to enable a state of "noise reduction" or a radical shift in the state of consciousness (Parker, 1994, 2000). Another apparent functional relationship was the presence of target–related imagery that appeared to be synchronously or even causally linked to the occurrence of specific imagery in the film clip. Even unexpectedly changing the sender seemed to evoke an appropriate response in the receiver ín the form of an apparent surprise: "Where have you been?" (Parker, Persson, & Haller, 2000).

The supposed psi-conducive state could be measured simply by requiring participants to respond to taped instructions with numerical ratings during every 10 min of ganzfeld stimulation (Parker, Millar,

& Beloff, 1976) but this proved disturbing and seemed to result in significant psi-missing scores (Parker, 1975). While postsession questionnaires are now often employed, these obviously lack precision in identifying the state or change in state associated with the correct imagery. Future research might employ more sophisticated nonverbal methods such as button-pressing, which could be used to directly assess the state at the actual time of the potential psi experience.

Historically, the ganzfeld technique grew out of the need to specify and lift out the most functional aspects of hypnosis and dream research as regards psi-conduciveness. The search for a *specific* hypnotic psi-conducive state has another history, and it is one which has led hypnosis research into uniting with the correlational approach. As a result, hypnotic psi research has encountered some of the hazards of the correlational approach that now need to be described.

In Part 1, research was reviewed showing how it is misleading to avoid the debate over whether the hypnotic state exists or not, because it has led to the rather naive assumption that this state empowers causal effects above and beyond what could be achieved by manipulations of individual variables in the waking state. Even assuming hypnosis is a special state, it appears rather clear from the reviewed research that this state is not a unitary one but is heavily influenced by a multitude of variables that intercorrelate.

There is a further serious and potentially fatal weakness in pursuing psi in the correlational way. This is the risk of falling into the quagmire formed by correlating amorphous sets of error variance. Tests of psi seldom give an effect beyond 10% of the error variance, and many psychological tests are anything but pure measures of what they claim to measure. Moreover, the correlations of psi scores with external variables such as hypnotizability rarely, if ever, account for more than 20% of the shared variance. Seen from this perspective, the research data become a quagmire of error variance with little in the way of solid ground to get a grip on the psi, and this may be in part because of its so-called elusiveness. The situation is made even worse when we consider, as I noted earlier, that in hypnosis research the responses to many of these psychological questionnaires, such as those of absorption and dissociation, have been shown to be influenced by the context in which they are given. As we noted in Part 1, when the questionnaires are given separately from the treatment procedure, the supposed significant correlation virtually vanishes.

This is not to say that nothing has been learned or can be learned from the correlational approach. One plausible assertion is that something vital to classical hypnotic phenomena may have been lost by basing contemporary hypnosis research on the study of "normal" students in the psychology laboratory rather than studying special participants, who gave rise to the more dramatic claims for hypnosis.

Back to the Starting Point

An outcome of the intensive debate during the early 2000s in the journal *Contemporary Hypnosis* was the recommendation that efforts by made to rediscover the type of hypnosis that characterized the somnambulists of Charcot's and Janet´s time (Kallio & Revonsuo, 2003, 2005). The somnambulists of today were now to be renamed "virtuosos." It would seem that a major goal of at least some state supporters is now to find these virtuosos. A profile that seems in many ways to define virtuosos has resulted from the efforts to select individuals with high scores on the Dissociated Experience Scale (DES), the Tellegen Absorption Scale (TAS), and even the Transliminality scale.

In using dissociation as a key concept and measure, a key aspect that has not been addressed is the question of whether or not virtuosos are psychologically disturbed individuals or unusually gifted healthy individuals. We recall how leading researchers such as Charcot and Bernheim had opposing views on this issue, and this historical division set the scene for the future of hypnosis research, with the issue now reappearing in the present context. If we are searching for healthy and gifted individuals experiencing high dissociative states then the DES is not the appropriate instrument. The DES was developed as a clinical instrument to aid in the diagnosis of dissociative identity disorder (Bernstein & Putnam, 1986). Individuals who score high on the DES usually have severe relationship problems, report having fearful parental attachments, and have a low sense of coherence in life (Ray, 1996). DES scores of 20 or more are usually regarded as the cutoff in the general population for psychopathological responding. Wright and Loftus (1999) have

been outspoken about the skewing of the DES when it is used on the normal population and accordingly developed an alternative form (the DES-C). However this was not used in some of the major contemporary work on hypnotic psi.

The Search for the Elusive Psi-Conducive Hypnotic State

In view of the enormous complexity of hypnosis, reviewed in Part 1, it is not surprising that any effort to elucidate the literature on psi and hypnosis as to their possible causal connections becomes, if not an already forlorn quest, at least an extremely daunting one.

Such efforts benefit considerably from the meticulous meta-analysis published by Stanford and Stein (1994). The authors found 25 studies that used hypnosis along with control comparison conditions such as relaxation. The main finding was a strong experimenter effect amongst the 12 investigators involved in these studies. As far as the role of the hypnotic state was concerned, 21 of the 25 studies used a within-subjects design rather than a between-subjects design, which made the outcomes liable to expectancy and order effects. Worse, many of the studies showed gross flaws in design and absence of precautions. In some cases these deficiencies were even associated with the differences in psi scores between the hypnosis and control groups that seriously weakened any definitive conclusions about the role of hypnosis. The order effects in one major study led Stanford and Stein to write: "The hypnosis-comparison contrast was significant only when the comparison condition preceded hypnosis. The significance was due, substantially, to psi-missing in the comparison condition" (Stanford & Stein, 1994, p. 235). It has been supposed that psi-missing can operate as an effect of the holding back of best performance when the control condition comes first (see Parker & Millar, 2014).

As many of the studies used only a few highly selected and highly hypnotizable participants, there are several alternatives to consider: Hypnotic psi may be part of the virtuoso performance; it may be that the hypnotic procedure with its positive expectancy is a means of enhancing an already existing potential; or as Stanford and Stein seem ready to endorse, hypnotic induction per se may have no effect beyond a person-by-situation interaction (Stanford & Stein, 1994, p. 261). This focus on the "person" factor in the equation brings in the relevance of the three factors that were highlighted above in the ongoing controversy over hypnosis: fantasy-proneness, absorption, and dissociation.

Taking the first of these, Barber's fantasy-proneness, questionnaires have been developed that quantify both the frequency and the type of anomalous experiences and altered states in fantasy-prone individuals (Pekala, 1991). Cluster analysis of the types of individuals sharing responses showed that about 10% of the "hypnotically sensitive" report paranormal experiences (Pekala, & Forbes, 1997; Pekala, Kumar, & Cummins, 1992). This and the finding that the correlations between hypnotizability and reports of psychic *experiences* account for about 9% of the variance, led the authors to favor the view that psychic experiences do actually occur more often in those who have a special hypnotic ability (Kumar & Pekala, 2001, p. 275).

However, the earlier problems we encountered in Part 1 while trying to make sense of such relationships reoccur when examining hypnotizability and psi experiences. Could absorption underlie the relationship between fantasy proneness and "hypnotic ability"? Are the correlations reliable indexes of the strengths of the relationships given that the context of administering tests on the same occasion gives stronger correlations than administering the tests on separate occasions? Worse, several items on the absorption scale can be interpreted as relating to paranormal experiences and thereby creating a degree of tautology or overlap between the scales.

Experimental studies should help settle the issue. However there is apparently only one such major study relating absorption to psi scores—in this case with the ganzfeld as the method of inducing psi (Dalton, Zingrone, & Alvarado, 1999). Participants were highly selected, with a background in art, music, drama, and even meditation, all of which might easily have contributed to the outcome. Three variables highly relevant to the debate on the hypnotic state were included: absorption, dissociation, and altered state. The results using overall first rank hits as the outcome measure were rather spectacular in that the number of hits was more than double the chance expectancy.

With such a clear effect, there can be some confidence that the correlations actually relate to psi rather than to extraneous factors. Decisively, the psi scores failed to relate to dissociation as measured by the DES, but they did show a moderately significant relationship to absorption. On the other hand, the attainment of an altered state through the use of the ganzfeld did not relate significantly to the psi scores, which was contrary to what had been predicted. Considering the scarcity of such strong psi effects, it is regrettable that the report (as appears to be the case with most of the work under the leadership of Dalton) is not published in a peer-reviewed journal.

It is also relevant to mention here that our Gothenburg real-time digital ganzfeld study was not successful in that the absorption scale did not distinguish psi scores, but it should be noted that most of the participants were already high scorers on absorption, thereby rendering the test insensitive as a predictor of scoring (Parker, Grams, & Pettersson, 1998).

Despite the above failure to find support for dissociation (DES) scores in predicting psi scores, this relationship may not be entirely invalid. An analysis of the DES scores for attendees at conferences on psychic experiences showed that they correlated moderately with the self-reported frequency of subjective psi experiences. However, the respondents here claimed that the psi experiences did not occur in truly dissociated states but "rather as an intrusion into an otherwise normal, conscious state" (Richards, 1991, p. 87).

It may be that there is an interaction between scores on hypnotizability, dissociation, paranormal belief, and psi experiences. The statistical analyses of scores on the Anomalous Experiences Scale indicate that these factors interact in a way that makes some individuals more sensitive than others to psychic experiences and abilities (Kumar & Pekala, 2001). A skeptical explanation would of course be that hypnotizability together with dissociation enhances fantasy proneness, and paranormal experiences are merely a feature of this fantasy proneness leading to the delusionary belief of being psychic.

Experimental studies are clearly much needed if we are to come closer to any definitive answers. Four such studies have been reported in the literature since Stanford and Stein carried out their meta-analysis. A problem with all these studies is that when they followed the design described earlier of comparing low and high hypnotizable groups, these and further subdivisions left a very small number of participants in the final comparison groups. Applying statistical analysis to this kind of data often results in "power failure," which typically arises through the use of small samples, low power, flexible designs, and flexible statistics. This increases the likelihood of false positive results, and the subsequent failure of such findings to replicate has been called the "winner's curse" (Button et al., 2013).

One of the first of these contemporary hypnotic psi studies, by Del Prete and Tressoldi (2005), used "hypnotic sessions" to produce a "hypnagogic state" in order to facilitate scoring on a forced-choice ESP task using static targets. The study involved 12 participants previously selected for their high scores on the absorption and transliminality scales. Two conditions were compared for their effect on psi scores. One of these conditions was what they called the "hypnagogic state" whereas the other condition was "self-induced relaxation." The hypnagogic condition gave a significant hit rate of over 37% whereas the control procedure (self-induced relaxation) gave results close to the chance expectancy of 25%. Although most contemporary authorities on hypnosis would clearly avoid confusing the hypnagogic state and hypnosis, the hits of the hypnagogic group correlated positively and significantly with the scores on the absorption scale (r = .76) and on the transliminality scale (r = .71). This might be considered surprising given that the participants had already been selected for high scores. The study is clearly remarkable given the significant findings with so few participants.

Tressoldi and Del Prete (2007) carried out a second study which can be regarded as an attempt at replicating and extending the above design. The difference was that the hypnotic instructions were supplemented with suggestions that were focused either on the use of ESP or the induction of an OBE to view the target. Although the authors concluded that their findings replicated the earlier ones, it is evident from the results that formally this was not so, because the scores in both conditions were close to chance expectancy, except for what would seem to be a post hoc session effect and marginally significant (by two-tailed tests) correlations of hits in the ESP condition with absorption and transliminality scores.

It is unfortunate that in both the above studies rather scanty details of the target selection and security aspects of the procedure are given. Moreover, it would seem following further inquiry that the same

target selection may have been used for all participants (Tressoldi, personal communication, March 10, 2014). If so, this would be a serious flaw as it would allow for order effects and possibly even stacking effects to occur in the data.

A further study was reported by Parra and Argibay (2013) using a design which compared the psi performance of groups selected on hypnotic suggestibility and subjective psychic experiences. The high suggestibility group (n = 20) scored at 40% where MCE was 25%, which was significantly higher than the low suggestibility group. It is unfortunate given the success of the experiment that the possibility of handling cues in the judging procedure was not eliminated (Parra & Argibay, 2013).

The method of splitting participants into contrasting groups on the basis of scores also formed the basis for a study of dissociation and hypnotizability by Cardeña and co-workers in Lund. Their first study compared the psi scores of participants who had been selected for their high or low scores on hypnotizability. The aim of this study was to make a further division amongst the participants into those with high versus low DES scores and evaluate the psi scores of this successive division. Although some minor post hoc significant results were found, the crucial and major hypothesis, that high hypnotizables and high dissociaters would excel over the low hypnotizables and low dissociaters, failed to gain any support from the data (Cardeña, Marcusson-Clavertz, & Wasmuth, 2009).

A further study was carried out by Marcusson-Clavertz and Cardeña (2011) along similar lines, but this time the ganzfeld state was used as an intended means of producing and evaluating psi performance. The study merits some detailed scrutiny given that the claims that were made for it ignited an academic and a public controversy (Halle et al., 2012). There appear to have been logistical problems from the beginning: a handful of participants, a study overloaded with variables and hypotheses, and a design that was part of a conventional study concerning "mind wandering." The overall results, based on the use of both direct hits and sum of ranks in the ganzfeld, lay close to chance expectancy. For most critics of the ganzfeld, this would be the beginning and the end of the significance of the study. This was, however, not a standard ganzfeld experiment with relatively normal individuals but used hypnotizability and dissociation in order to arrive at contrasting groups of individuals.

The above problem concerning "power failure" due to the low number of individuals in the various groups becomes then extremely acute. Moreover, given that individuals were chosen from the upper and lower ends of already skewed distributions, it is likely that some of the tests chosen by the authors violated the rules for the use of parametric statistics. When numbers are reduced to between 5 and 9 in the various groups, the investigation moves dangerously close to a case study. In this instance, the report is devoid of the qualitative information about individuals that such a study can provide, so it unfortunately falls between the two worlds.

The main hypothesis was that the combination of high hypnotizables and dissociaters would give superior psi scores, but this was not confirmed. Neither did, as predicted, the dissociation scores relate significantly to the psi scores. However, after controlling for the sheep-goat effect, a significant *negative* relationship was found between psi scores and hypnotizability, a finding that was counter to the hypothesis being tested. After looking at the distribution of scores on the Phenomenology of Consciousness Inventory (PCI), the authors then proceeded to speculate that some of the high dissociaters may have failed to engage their psi because they were unable to use the ganzfeld to enter an altered state. Although dissociation clearly did not relate to psi scores, the authors reported (apparently based on five participants) that the high dissociative-high hypnotizables seemed to be correct when they relied on their "hunches" rather than their imagery".

In terms of the prior hypotheses, the only indisputable positive finding is a form of sheep-goat effect in which belief in the individual's own success correlated significantly with the psi z scores (although the specific measure used to test this hypothesis may not have been prespecified).

The study appears to present many post hoc findings although they were not identified as being post hoc. Perhaps the most promising one of these concerns the previously described PCI, which in this experiment was used to access the altered state experienced during the ganzfeld sessions. Scores on the PCI were found to correlate significantly with the psi z scores, but only amongst the high hypnotizables. However,

as the authors freely admit, there are serious problems with this finding because the PCI was administered after the first (nonpsi) ganzfeld session as part of a "mind wandering experiment" and not after the actual psi-ganzfeld session. This makes it rather unclear as to what lies behind the significance. Was it the actual state during the psi experiment, the PCI scores, the expectancy created by the nonpsi ganzfeld session, or the "winner's curse" referred to above?

It is difficult to go further with such data but we can try to answer the fundamental questions raised earlier concerning the nature of dissociation: How do we measure it and what do we mean by it?

Is Dissociation a Healthy or Pathological State and Does Dissociation Relate to Psi?

If we look at the psychometric market, we find there are several different questionnaire measures of dissociation (Carleton, Abrams, & Asmundson, 2010; Harrison & Watson, 1992; Körlin, Edman, & Nybäck, 2007; Riley, 1988) as well as some potentially direct measures (Palmer, 1994**,** 2011, 2013).

We can begin with the previously described experiment by Marcusson-Clavertz and Cardeña (2011) which used the more pathologically oriented Dissociated Experience Scale (DES) for selecting high dissociaters. Although their report regrettably does not supply any information about their mental health, it should be added that far from all the high scorers met the criterion for a DSM diagnosis (Ray, 1996). On the other hand, the series of psychological studies of hypnosis at Lund found that the high-dissociation-high-suggestible group had "elevated pathological dissociation and fantasy proneness and reported a greater history of exposure to stressful events (Terhune, 2010, p. 31). It is worth remembering the study by Barrett (reviewed in Part 1) in which he interviewed highly hypnotizable persons and described about half of them as "fantasizers" and the other half as "dissociaters" (Barrett, 2010). The majority of the dissociaters reported nightmares and memories of childhood trauma and abuse (although no attempt was apparently made to check on the validity of these experiences). As the dissociaters were found to be more likely than the fantasizers to believe in the reality of their hypnotic hallucinations, this study provided some support for the idea mentioned earlier that virtuosos were hallucinating rather than merely imagining the suggestions.

The explanation for finding any potentially "normal" or "healthy" individuals amongst high DES scorers, such as may or may not have been the case in the Marcusson-Clavertz and Cardeña (2011) study, can lie in the apparent bimodality of the DES. More disturbed patients are found to fall on the part of the curve based on responses affirming the presence of amnesia, depersonalization, and derealization, and even post-traumatic stress disorder. By contrast, apparently normal and healthy individuals with high DES scores mainly affirm the items concerning absorption and imaginative experiences (Körlin et al., 2007; Ray, 1996; Waller, Putnam, & Carlson, 1996).

This suggests that using dissociation in this latter restricted sense of absorption and imaginative ability is a more promising route forward, and this is especially so when we bear in mind the fairly consistent results which relate absorption to subjective psi experiences.

The obvious conclusion that emerges is that if we wish to pursue further the quest for the psi-conducive state, there may be good reason to focus on the concept of dissociation, but then it is necessary to distinguish the positive and healthy aspects versus the pathological aspects (Wright & Loftus, 1999). In searching for an instrument to measure dissociation in the normal population, there exist, as well as the earlier mentioned DES-C, several other variations that seem to share the common view of relative healthy dissociation as encompassing absorption and attentional distraction (Carleton et al., 2010; Harrison & Watson; 1992; Riley, 1988). The most established of these test instruments is the Dissociative Processes Scale (DPS), which is derived from the DES. It shows good reliability by having very high alpha coefficients, and it is adapted to the normal population. In addition to the normality aspect, what speaks for the future use of the DPS is that it has as its main factors: absorption, obliviousness, and detachment (Watson, 2001).

Palmer used and evaluated a range of dissociation measures as part of a relentless effort to find a reliable means of detecting psi at an unconscious level. His rationale is that psi functions best when tested subliminally and in a dissociated state (Palmer, 1994**,** 2011, 2013). In the initial series of explorative studies

Palmer carried out, he designed a means of testing psi, the Perceptual ESP Test (PET), that might be sensitive to spontaneous dissociations or mind wandering. The test used shaped carets (directionally placed V-shaped symbols) to reduce the response bias that occurs with traditional ESP card symbols. To measure dissociation, Palmer used one of the above scales designed partly for the normal population, the Questionnaire of Experiences of Dissociation, and supplemented this with ongoing dissociative state reports during the responses to the PET. This essentially assessed the degree to which the receivers felt suddenly compelled to look at the specific area of the screen at the point in time when the sender was viewing the target carets. The experimental designs also incorporated various subliminal effects on ESP scores based on Palmer's earlier work with Martin Johnson using the Defense Mechanism Test. These were found to be effective only in participants who reported relatively high scores on the dissociation measures (Palmer, 1994).

The notion of utilizing dissociation as an outside force for influencing choices in ESP experiments then became a distinctive feature of Palmer's later work (Palmer, 2011). To make fuller use of this, he used a form of the Ouija board in which receivers were to try to identify the letters that composed one of the five possible target words being viewed by the sender. Palmer used this time as the "trait measure" of dissociation the Complex Partial Epileptic Signs scale with the TAS partialled out. For a "state measure" of dissociation Palmer simply asked about the degree of outside force participants felt that appeared to be directing the pointer across the Ouija board. The main incontrovertibly significant finding was a curvilinear relationship between the ESP scores and the influence of an outside force. This meant that those who felt the influence up to 40% of the time scored significantly on the psi test. Although this was a post-hoc finding, it is one that made sense, as those with very high scores could be conceived of as being unrealistically overconvinced of their ability.

In Palmer's current effort in this area dissociation was used again to predict psi scores with another derivation of the Ouija board (Palmer, 2013). In this case the Ouija board took the form of a modern computer tablet divided into squares where participants were required to locate clairvoyantly the square designated as the target (reminiscent of the childhood game of cruisers and battleships). For the trait measure, Palmer now turned to the above recommended Dissociative Processes Scale but retained the same state measure concerning the degree of influence of an outside force on the hand**.** When reconsidering the degree to which this force was felt, Palmer later thought that because the new participants seemed to differ in certain key aspects from the earlier ones, those reporting an influence over 40% of the time should now be included. With this criterion in place, the DPS proved to be the most significant predictor of ESP scores. The unique and positive feature of Palmer's work, although it suffers from multiple analyses, post hoc findings, and ad-hoc hypotheses, is that it combines a correlational approach with a novel functional state approach and focuses on normalcy.

These studies with this revised "normalization" of dissociation seem then to offer some land winnings but the revision may also lead to some conceptual reformulations. To grasp the wider applications of the proposed revision of the dissociation concept, we need to look at a third area of dispute concerning altered states: namely, the findings on sleep and dream processes. A similar, often heated, debate to those over the hypnotic state and the existence of psi is found here. In this case the dispute is over whether the dream states have a psychological meaning or if from a skeptical reductionist standpoint they are mere epiphenomena. This is a debate that reflects many of the same parallel issues that have occurred in the psi and hypnosis debates. Readers who have an insatiable interest in controversies are referred to the *Brain and Behavioral Sciences* issue on the topic (Hobson, Pace-Schott, & Stickgold, 2000) and the three-volume *The New Science of Dreaming* (Barrett & McNamara, 2007).

Entering this, the third potential jungle of research findings, is beyond the scope of this review so we choose to keep within the secure path mentioned above, that of the positive measure of dissociation. What speaks for this choice of path is that a positive concept of dissociation has given rise to a theory in dream research that promises to bring together the many disparate and incongruous findings, not only in dream research, but also in hypnosis and psi research. This approach is termed the "continuity of conscious states theory," because it highlights the degree of continuity in the human capacity for experiencing various altered states of consciousness (Watson, 2001, 2003).

The Continuity of Consciousness Theory

The theory proposes that individuals who have intense, novel, and unusual experiences are more likely to recall their dreams and integrate them into their waking life. This capacity enables a link between sleep and waking states and even becomes expressed in semidissociated states such as absorption in fantasy and daydreaming. The term "sleep-related experiences" is used to include a variety of altered states that may occur nocturnally. The continuity of consciousness is said to occur potentially in the recall of nightmares, vivid dreams, recurring dreams, lucid dreams, and problem-solving dreams. The theory has support in research findings that show dream recall to be related to fantasy proneness (Blagrove & Hartnell, 2000), transliminality (Soffer-Dudek & Shahar**,** 2009), and absorption (Schredl, Jochum, & Souguenet, 1997; Watson, 2003). With a singular exception (Knox & Lynn, 2014) the theoretical implications of this work do not, however, appear to have reached either hypnosis researchers or, to my knowledge, psi researchers.

What gives this continuity of consciousness theory some plausibility is the evidence that dreams have a direct influence on waking experiences as part of a reciprocal flow of consciousness in which the waking experiences influence dream content and these in their turn come to influence waking life (Schredl & Hofmann, 2003). There is here a tangible clue for finding the way forward, because sleep-related experiences appear to have a positive relationship to measures of "transliminality": the ease with which previously unconscious psychological material enters consciousness (Soffer-Dudek & Shahar, 2009).

Clearly, a large degree of individual variation is expected to occur in the propensity for experiencing the continuity of various altered states. For instance, the recall of sleep-related experiences significantly positively relates to dissociation (Giesbrecht & Merckelbach, 2006; Giesbrecht, Smeets, Leppink, Jelicic, & Merckelbach, 2007).

However some skepticism concerning the continuity and reciprocity of relationships here seems appropriate: Do sleep related experiences occur due to transliminality or do sleep disturbances create the dissociative experiences (Lynn, Lillienfeld, Merckelbach, Giesbrecht, & van der Kloet, 2012; Soffer-Dudek & Shahar, 2012)? Statistically, in terms of multiple regression equations, this should be a relatively easy question for future research to answer, but it does lead us further into the clinical area.

In the clinical area, there is a contradiction as to what might be expected from the pigeonholing of medical diagnosis. This concerns the evidence that symptoms of dissociation correlate significantly with some of the signs of schizophrenia (Spitzer, Haug, & Freyberger, 1997). The relationship between dissociation and signs of schizophrenia in the normal population appears relatively strong, but the strength of that relationship depends on the inclusion or exclusion of the common factor of depersonalization (Watson, 2003). This finding can be readily understood if we take a dimensional view of psychosis and schizophrenia rather than the traditional medical one (Bentall, 2009). The dimensional concept allows for the existence of so-called healthy, positive individuals whose symptoms often are limited to anomalous experiences (Claridge, 1997; Nelson & Rawlings, 2010). Whilst the occult beliefs of such individuals can merit at least in psychiatric circles a diagnosis of schizotypal personality, many of these individuals appear to be well-organized and structured, so as not to be of much clinical concern (Goulding, 2004). As such they often are called "happy schizotypes," although the expression "schizotypy" may be a misnomer considering the apparent healthy status of such individuals (Claridge, 1997).

Claridge, who carried out the pioneer work on "positive" or "happy" schizotypes, joint-authored a paper reporting that the schizotypy measure correlated with distressful nightmares (Claridge, Clark, & Davis, 1997). Yet, surprisingly, schizotypy also correlated significantly with dream enjoyment, leading the authors to construe "schizotypy as an essentially neutral personality characteristic, which includes among its cognitive features the tendency towards a wide range of unusual perceptual experiences and thought styles (Claridge et al., 1997, pp. 384–385). This line of reasoning may explain why schizotypy, assessed on the basis of the Unusual Experiences Scale, has shown moderate relationships with both creativity and absorption (Nelson & Rawlings, 2010).

Similar conclusions concerning the relationship of anomalous and psi experiences to schizotypal personality were reached by Simmonds-Moore (2009) but by a different route. Instead of relating these

experiences to a common ground in absorption, Simmonds-Moore explored the relationships between psi experiences and just about all the variables that have been implicated at one time or another as having a potential relationship to psi (dissociation, transliminality, boundaries, temporal lobe lability, and creativity). She concluded that thinner boundaries would naturally allow less filtering of information to the brain and form the basis of what she preferred to call the "anomaly prone personality." This filtering of extrasensory information can, however, go in a positive or negative direction depending on the degree of control over the anomalous experiences. Control is then the hallmark of healthy functioning (I would add "integration") that distinguishes this personality from the disturbed schizotypal personality. The problem is that while some of these variables may show weak correlations with anomalous experiences (Irwin, Schofield, & Baker, 2014), there is little reason to believe they can function as predictors of actual psi performance in the laboratory. Even Simmonds-Moore's own results on this cast serious doubt on the utility of these variables (Simmonds-Moore & Holt, 2007).

There can be many reasons for this failure, some of which were mentioned earlier. Even if the evidence is lacking, it might still be that the successful participants in psi experiments, especially those in hypnosis and ganzfeld experiments, belong to the "positive schizotypy group," rather than showing the more extreme symptoms of dissociation and psychosis proneness. It was after all from groups of "positive schizotypes" with their own psi experiences that we recruited most of our successful participants in our psi-ganzfeld experiments. They were as a group characterized, as might be expected, by high scores on magical ideation and absorption (Parker, 2000; Parker et al., 1998). By contrast, Simmonds-Moore and Holt (2007), who failed to replicate the relationship to schizotypy, recruited their relatively few volunteers from university staff and students.

Whatever the reason for this failure, it serves as a reminder to take heed of the demands for purely predictive findings—"psi on demand"—made by the critic Alcock (2003). It seems evident that correlational findings fall far short of this, if they are not in fact steps in the wrong directions.

What Can Be Learned From the Failure to Predict?

It is possible that an intensive study of virtuosos and star participants might turn the situation around if we use another approach focused on functional relationships. As we related in Part 1, for instance, the intervention of secretively and suddenly changing the sender in the ganzfeld procedure produced a seemingly appropriate response ("Where have you been?") in the receiver (Parker et al., 2000). This indicated, at least for these researchers, that psi is more than a statistical anomaly, in that it follows mental causality. An earlier study with two star performers gave some further apparent insights into the nature of psi beyond what could be achieved by correlational studies. A companion paper to this (Parker & Millar, 2014) related how two successful psi performers continued to score high with experimenters who previously had a track record of being so-called "psi-inhibitory." Moreover, for the few trials that I had time to carry out, I reversed roles with one of the star participants, and in the role of receiver I then obtained high scores. This all suggests that with the right atmosphere and right expectancy, there does appear to be a degree of robustness in the psi scores of some star performers that can potentially fulfill the demands of critics such as Alcock.

In the third of the Gothenburg ganzfeld-psi series of experiments, seven pairs of participants who had made direct hits were invited back, and four of these repeated their success in making again direct hits (Parker et al., 1998). With hindsight, we should have made an intensive study of these individuals, but the lack of funding led to the diversion of resources to other areas.

This was a missed opportunity but what we may want to learn from the above is that both correlational studies using tests of transliminality, dissociation, and schizotypy and studies applying functional methods to the ganzfeld and hypnosis should focus on following up the selected high psi performers. It is also clear that we need to be more innovative. It is often thought of as taboo to vary the conditions during the running of experiments, but as the anecdotal examples given earlier suggest, it is by making such unexpected interventions that we can actively explore causal effects on psi scores. Perhaps some of these suggestions will help free research from a morass of error variance.

The fundamental issue, which was also raised in Part 1, is are we prepared to consider that hypnosis and placebo effects can potentially engage awareness beyond its normal constraints and thereby produce what appear to be remarkable phenomena? If so then concepts such as transliminality, meditation, and mindfulness may offer alternative explanatory systems for the more dramatic effects. This is a line of thinking that a recent spate of articles seems to support (e.g., Alladin, 2014; Halligan & Oakley, 2014; Krippner, 2014). Nevertheless in fairness and completeness it should be mentioned there are of course even other views—for instance that hypnosis is a species of trance phenomena with diminished rather than heightened awareness (Crabtree, 2012).

Just how hypnosis relates to other so-called mediumistic trance states is obviously a matter that should be resolved by empirical research. Some preliminary modern work was, recently reported that has relevance here (Roxburgh, 2007; Roxburgh & Roe, 2011). Significantly, mediums were found to be higher than nonmediums on psychological wellbeing and to experience lower distress. No significant differences were found as regards boundary thinness or dissociation, but it should be noted the researchers used the pathological measure of dissociation, the DES, rather than the normality oriented DPS. Regrettably, beyond these studies, there is mainly historical material to go on (Garrett, 1938; Sidgwick, 1915).

In the debate over how certain forms of dissociation and schizotypy may relate to hypnosis, the health versus pathology issue achieves considerable theoretical importance. In focusing on the healthy and positive aspect of altered states, an argument can be made for the viewpoint, consistent with the above literature review, that these states enable the expression of latent potential (Tart, 1977), especially latent creativity (Lynn & Sivec, 1992).

Conclusions

Recently I discovered how the Swedish physician and foremost hypnosis researcher, John Björkhem, came more than 60 years ago to similar conclusions to the above and those of the earlier papers in this series. Björkhem's conclusions were apparently based on 30,000 experiments with 3,000 individual subjects carried out between 1930 and 1950 at Uppsala and Lund universities (Björkhem, 1953a, 1953b).

Some of the studies have been part of one of his doctorates and his subsequent work (Björkhem, 1943; Björkhem, 1951) but no systematic evaluation appears to have been published in peer-reviewed journals. Even so, in terms of the numbers of experiments and participants, all contemporary psi studies fade into insignificance by comparison. Björkhem's conclusions, given the enormity of his experience and of his data, are worth considering. Moreover, his conclusions correspond closely with those being arrived at here. They can be summarized as follows:

1. The most psi favorable state is not deep hypnosis but slight changes in consciousness.
2. Those subjects engaged in somnambulistic states have to learn to distinguish the psi state from that of fantasy and hallucinations.
3. Other persons present may exert an influence on the outcome.
4. Observations of other successful experiments promotes the appearance of the phenomena.
5. To produce the right state of mind for psi in participants, the experimenter must make deep psychological contact with the participants.
6. Extremely subtle testing conditions can also influence the nature of the phenomena.
7. The psi factor is a gift and appears to have a genetic endowment.
8. The way forward is to use gifted mediums.

Some of the above conclusions have been previously discussed but most merit further comment. If the first claim concerning slight changes in consciousness being a psi conducive state has validity, then any exclusive reliance on hypnosis may indeed be a false lead or an unnecessary complication. This point emphasizes not the state per se but slight changes in it. What has been termed the "change in state hypothesis" has up to now only been briefly studied with the ganzfeld with inconclusive results (Parker, 1994).

The hypothesis deserves more decisive investigation not only with the ganzfeld and hypnosis but in the field as a whole. In view of the findings relating psi data to absorption, it may be the case that participants show spontaneously changes in state that are psi conducive. In earlier unpublished work I carried out with Brian Millar, we individually tested a high scoring participant whom I had just discovered. The participant was asked to give state reports (as described by Tart, 1972) for each run of 25 trials with ESP cards. These state reports failed to predict scoring, but in making our analyses we were in fact entering the mire of meaningless correlations described earlier in this paper. Today, further examination of all the scores from the 20 runs carried out reveals there were only three scores of (9, 10, and 13 where MCE = 5) that we could be reasonably certain gave hits that were not due to chance. For the two highest of these scores (and none of all the other scores) the participant remarked that he noticed a sudden change from relaxation to tension. What is then needed is a study of the cases of other high scoring participants, with the specific objective of lifting out only the occasions when they are producing high scores or remarkable correspondences in the content of their mentation that cannot easily be attributed to chance.

When Björkhem (1953b) additionally adds: "In certain cases an influence from the experimenter seems to condition the emergence of the psi factor" (p. 86), he appears to have pre-empted the current focus on the experimenter effect and possibly experimenter psi. As far as we can judge his own role as an experimenter, it has to be said that by any measure Björkhem was an extraordinarily gifted and charismatic individual. He gained doctorates in theology and psychology, published five books, and became a licensed physician. He was said to be the most educated person in Sweden of the time and it would appear that it was only his interest in researching psi which excluded him from an academic position (Stolt & Björkhem-Bergen, 2004). Björkhem's remarkable success teaches us that we need to select and study the gifted experimenter as well as the gifted participant, but it teaches us more than this. As it was noted in Part 1, Björkhem's success even with hypnosis itself failed to work in a change of culture and in the ambience of Rhine's laboratory at Duke University. Clearly, we need also to examine the context and ambience in which psi-conducive experimenters best function.

The third point that Björkhem raised can be expressed in contemporary terms as that of "observational learning." This aspect has been nearly totally neglected in parapsychology but is a major principle in the teaching skills and self-efficacy, especially in the field of sports psychology. It was because of this that in initiating the Gothenburg ganzfeld work we arranged for a visit by the successful ganzfeld researcher Kathy Dalton accompanied by her videotaped recording of her best ganzfeld hits. It seems also relevant to mention how the PK success of Felicia Parise was inspired by observation of Nina Kulagina's apparent PK performance (Honorton, 1993).

The existence of a genetic influence on psi, which Björkhem testified to, has also received little investigation unless we include twin studies (Brusewitz, Cherkas, & Parker, 2014; Jensen & Parker, 2012; Playfair, 2013). In Björkhem's own case the claim may actually have an anecdotal aspect. Besides his academic and clinical achievements, Björkhem fathered five children, all of whom became academically successful. One of these was Örjan Bjorkhem, who seemed to possess his father's ability to find and even produce psi phenomena (Johnson, 1998). Interestingly, shortly before his premature death Örjan Björkhem gave an account for the first time in English of his PK successes, along with a detailed theory that proved to be remarkably similar to that of Batcheldor, although he was unfamiliar with that work (Björkhem, 1994). What appeared to unite both the Björkhems in their view of hypnosis is that it was a culturally available and effective procedure, which allowed the participants to shift their view of reality to include psi.

So is the search for the hypnotic psi-conducive state, independent of culture and the historical era, a futile search such as being lured by fools' gold? Perhaps so, but some valuable findings and more sophisticated ways of thinking have emerged. In studying apparently psi-gifted individuals, the 60 years of "hypnotic state wars" have given us greater precision in measurement as well as a deeper and broader perspective on these issues and, perhaps, some consensus. In terms of consensus, it is hard not to agree with the two major protagonists in the hypnosis debate, Kirsch and Kihlstrom, on the importance of rigorous methodologies and on the relevance of hypnosis for the wider understanding of how human cognition works (Kihlstrom, 2014; Kirsch, Mazzoni, & Montgomery, 2007; Lynn, Woody, Montgomery, & Gaudiano; 2014). Likewise,

I find myself agreeing with Cardeña (2014) as to the relevance of both hypnosis and psi for understanding the nature of consciousness and perhaps even for solving some of the major problems of psychology.

As outlined earlier, besides the methodological weaknesses, the disagreement concerns the complexity of the variables involved in psi and hypnosis, which, in a time of limited funding, creates a series of intractable problems. Expressed in the Rhineian symbolism chosen here, it bogs down research in the correlational quagmire. Accordingly, should further expeditions be made into this area, much better equipped and sophisticated designs with large groups of individuals will be required.

A more viable alternative might be the repeated testing of selected performers who are gifted at altering their state of consciousness. This would enable a focus on the personal history and characteristics of individuals who produce high-quality psi. In view of what was presented in a previous paper concerning the success of psi-conducive experimenters as their own successful subjects (Parker & Millar, 2014), then self-experimentation is a legitimate option that psi research needs to embrace. Luke (2011) succinctly expresses this in his praiseworthy presidential address to the Parapsychological Association: "experiential reclamation" is now part of the agenda. This can potentially not only generate new and valuable findings but can in the long term fulfill the currently vocal demands for replication.

Finally, it seems appropriate to reveal my own experience and bias. Although I have worked for many years with hypnosis in clinical and research contexts, I confess to having no final and precise opinion as to what hypnosis is. I do however have a strong opinion that an understanding of hypnotic and psi experiences goes hand-in-hand with learning about what consciousness is. It is often said (and attributed to Richard Feynman) that those who believe they understand quantum mechanics, do not. I assert that the same thing can be said about consciousness and its altered states.

Perhaps we can go one step further. If as is often claimed in contemporary physics and astronomy, our human consciousness potentially possesses the ability to fully understand the universe and is capable of a "theory of everything," then this naturally presupposes that the complexity of the universe has evolved along with our consciousness. Our consciousness, and the understanding of it, would thereby appear to be a vital and an integral part of the sought-after theory of everything.

References

Alcock, J. (2003). Give the null hypothesis a chance. *Journal of Consciousness Studies, 10*, 29–50.

Alladin, A. (2014). Mindfulness-based hypnosis: Blending science, beliefs, and wisdoms to catalyze healing. *American Journal of Clinical Hypnosis, 56*, 285–302.

Barrett, D. (2010). Dissociaters, fantasizers, and their relation to hypnotizablity. In D. Barrett (Ed.), *Hypnosis and hypnotherapy* (Vol. 1, pp. 15–35). Oxford, England: Praeger.

Barrett, D., & McNamara, P. (Eds.) (2007). *The new science of dreaming* (Vols. 1–3). Westport, CT: Praeger.

Bentall, R. (2010). *Doctoring the mind*. London, England: Penguin Books.

Bernstein, E. M., & Putnam, F. W. (1986). Development, reliability and validity of a dissociation scale. *Journal of Nervous and Mental Disease, 74,* 727–735.

Björkhem, J. (1943) *De hypnotiska hallucinationerna* [The hypnotic hallucinations]. Stockholm, Sweden: Litteraturförlaget.

Björkhem, J. (1951). *Det ockulta problemet* [The occult problem]. Uppsala, Sweden: Lindblads Förlag.

Björkhem, J. (1953a). The future of parapsychological research [abstract]. *Proceedings of the First International Conference of Parapsychological Studies* (pp. 84–85). New York, NY: Parapsychology Foundation.

Björkhem, J. (1953b). The twenty-year experiments at Uppsala and Lund [Abstract]. *Proceedings of the First International Conference of Parapsychological Studies* (pp. 86–87). New York, NY: Parapsychology Foundation.

Björkhem, Ö. (1994). A theory of the event horizon: A new way of producing macro PK phenomena *Proceedings of Presented Papers: The Parapsychological Association 37th Annual Convention*, 57–64.

Blagrove, M., & Hartnell, S. J. (2000). Lucid dreaming: Associations with internal locus of control, need for cognition and creativity. *Personality and Individual Differences, 28,* 41–47.

Brusewitz, G., Cherkas, L., Harris, J., & Parker, A. (2014). Exceptional experiences amongst twins. *Journal of the Society for Psychical Research, 77*, 220–234.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J ., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews, 14*, 365–376.

Cardeña, E. (2014). Hypnos and psyche: How hypnosis has contributed to the study of consciousness. *Psychology of Consciousness: Theory, Research, and Practice, 1*, 123–138.

Cardeña, E., Marcusson-Clavertz, D., & Wasmuth, J. (2009). Hypnotizability and dissociation as predictors of performance in a precognition task. *Journal of Parapsychology, 73*, 137–158.

Carleton, R. N., Abrams, M. P., & Asmundson, G. J. G. (2010). The Attentional Resources Allocation Scale (ARAS). Psychometric properties of a composite measure for dissociation and absorption. *Depression and Anxiety, 27*, 775–786.

Claridge, G. (Ed.). (1997). *Schizotypy: Implications for illness and health*. Oxford, England: Oxford University Press.

Claridge, G., Clark, K., & Davis, C. (1997). Nightmares, dreams, and schizotypy. *British Journal of Clinical Psychology, 36,* 377–386.

Crabtree, A. (2012). Hypnosis reconsidered, resituated, and redefined. *Journal of Scientific Exploration, 26*, 297–327.

Dalton, K., Zingrone, N. L., & Alvarado, C. S. (1999). Exploring out-of-body experiences, dissociation, absorption, and alterations in consciousness with a creative population in the ganzfeld. *Proceedings of Presented Papers: The Parapsychological Association 42nd Annual Convention,* 48–67.

Del Prete, G., & Tressoldi, P. (2005). Anomalous cognition in hypnagogic state with OBE induction: An experimental study. *Journal of Parapsychology, 69*, 329–339.

Garrett, E. (1938). *My life as a search for the meaning of mediumship*. New York: Oquaga Press.

Giesbrecht, T., & Merckelbach, H. (2006). Dreaming to reduce fantasy? Fantasy proneness, dissociation, and subjective sleep experiences. *Personality and Individual Differences, 41*, 697–706.

Giesbrecht, T., Smeets, T., Leppink, J., Jelicic, M., & Merckelbach, H. (2007). Acute dissociation after 1 night of sleep loss. *Journal of Abnormal Psychology, 116*, 599–606.

Goulding, A. (2004). Schizotypy models in relation to subjective health and paranormal belief and experiences. *Personality and Individual Differences, 37*, 157–167.

Halle, B., Hesslow, B., Karlström, G., Lidin, S. D., Lindgren, G., Löfstedt, C., ... Svensson, B. E. V. (2012, October, 31). Pseudovetenskap sprids okritiskt [Pseudoscience is being spread uncritically]. *Svenska Dagbladet*. Retrieved from http://www.svd.se/opinion/brannpunkt/pseudovetenskap

Halligan. P. W., & Oakley, D. A. (2014). Hypnosis and beyond: Exploring the broader domain of suggestion. *Psychology of Consciousness: Theory, Research, and Practice, 1,* 105–122.

Harrison, J. A., & Watson, D. (1992). *The Dissociative Processes Scale*. (Unpublished manuscript). Department of Psychology, University of Iowa, Iowa City, IA.

Hobson, J. A., Pace-Schott, E., & Stickgold, R. (2000). Dreaming and the brain: Toward a cognitive neuroscience of conscious states. *Behavioral and Brian Sciences, 23*, 793–1121.

Honorton, C. (1993). A moving experience. *Journal of the American Society for Psychical Research*, 87, 329–340.

Irwin, H. J., Schofield, M. N., & Baker, I. S. (2014). Dissociative tendencies, sensory-processing sensitivity and aberrant salience as predictors of anomalous experiences and paranormal attributions. *Journal of the Society for Psychical Research, 78,* 193–206.

Jensen, C. G., & Parker, A. (2012). Entangled in the womb: A pilot study on the possible physiological connections between identical twins with different embryonic backgrounds. *Explore: The Journal of Science and Healing, 8,* 26–31.

Johnson, M. (1998). John Björkhem (1910–1963). In N.-O. Jacobson (Ed.), *Svensk Parapsykologi* (p. 92). Stockholm, Sweden: Swedish Society for Psychical Research.

Kallio, S., & Revonsuo, A. (2003). Hypnotic phenomena and altered states of consciousness: A multilevel framework of description and explanation. *Contemporary Hypnosis, 20,* 111–164.

Kallio S., & Revonsuo, A. (2005). Altering the state of the altered state debate: Reply to commentaries. *Contemporary Hypnosis, 22,* 46–55.

Kihlstrom, J. (2014). Hypnosis and cognition. *Psychology of Consciousness: Theory, Research, and Practice, 1,* 139–152.

Kirsch, I., Mazzoni, G., & Montgomery, G. H. (2007). Remembrance of hypnosis past. *American Journal of Clinical Hypnosis*, *49*, 171–179.

Knox, J., & Lynn. S. J. (2014). Sleep experiences, dissociation, imaginal experiences and schizotypy: The role of context. *Consciousness and Cognition, 23,* 22–31.

Körlin D., Edman G., & Nybäck, H. (2007). Reliability and validity of a Swedish version of the Dissociative Experiences Scale (DES-II). *Nordic Journal of Psychiatry, 61*, 126–142.

Krippner, S. (2014). The mind-body-spirit paradigm: Crisis or opportunity? *American Journal of Clinical Hypnosis, 56,* 210–215.

Kumar, V. K., & Pekala, R. J. (2001). Relation of hypnosis-specific attitudes and behaviors to paranormal beliefs and experiences: A technical review. In J. Houran & R. Lange (Eds.), *Hauntings and poltergeists—multidisciplinary perspectives* (pp. 260–279). Jefferson, NC: McFarland.

Luke, D. (2011). Presidential address: Experiential reclamation and first person psychology. *Journal of Parapsychology, 75,* 185–200.

Lynn, S., Lillienfield, S. O., Merckelbach, H., Giesbrecht, T., & van der Kloet, D. (2012). Dissociation and dissociative disorders: Challenging the conventional wisdom. *Current Directions in Psychological Science, 21,* 48–53.

Lynn, S. J., & Sivec, H. (1992). The hypnotizable subject as creative problem-solving agent. In E. Fromm & M. N. Nash (Eds.), *Contemporary hypnosis research* (pp. 292–333*)*. London, England: Guildford Press.

Lynn, S. J., Woody, E., Montgomery, G., & Gaudiano, B. (2014). Hypnosis: Contributions to psychological science and clinical practice. *Psychology of Consciousness: Theory, Research, and Practice, 1,* 103–104.

Marcusson-Clavertz, D., & Cardeña, E. (2011). Hypnotizability, alterations in consciousness and other variables as predictors in a ganzfeld psi task. *Journal of Parapsychology, 75,* 235–260.

Nelson, B., & Rawlings, D. F. (2010). Relating schizotypy and personality to the phenomenology of creativity. *Schizophrenia Bulletin, 36,* 388–399.

Palmer, J. (1994). Explorations with the perceptual ESP test. *Journal of Parapsychology, 58,* 115–147.

Palmer, J. (2011). Motor automatisms as a vehicle of ESP expression. *Journal of Parapsychology, 75*, 45–60.

Palmer, J. (2013). Extrasensory perception, dissociation, and motor automatisms. *Abstracts of Presented Papers: The Parapsychological Association 56th Annual Convention*, 25.

Parker, A. (1975). Some findings relevant to the change in state hypothesis [Abstract]. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in parapsychology 1974* (pp. 40–42). Metuchen, NJ: Scarecrow Press.

Parker, A. (1994). Eyeballing the ganzfeld. *Proceedings of Presented Papers: The Parapsychological Association 37th Annual Convention*, 314.

Parker, A. (2000). A review of the ganzfeld work at Gothenburg University. *Journal of the Society for Psychical Research, 64,* 1–15.

Parker, A., Grams, D., & Pettersson, C. (1998). Further variables relating to psi in the ganzfeld. *Journal of Parapsychology*, *62*, 319–337.

Parker, A., & Millar, B. (2014). Revealing psi secrets: Successful experimenters appear to succeed by using their own psi. *Journal of Parapsychology, 78*, 39–85.

Parker, A., Millar, B., & Beloff, J. (1976). A three experimenter ganzfeld [Abstract]. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in parapsychology 1976* (pp. 52–54). Metuchen, NJ: Scarecrow Press.

Parker, A., Persson, A., & Haller, A. (2000). Using qualitative ganzfeld research for theory development: Top down processes in psi-mediation. *Journal of the Society for Psychical Research, 64,* 65–81.

Parra, A., & Argibay, J. C. (2013). A free-response ESP test in two hypnotic susceptibility groups: A pilot study. *Australian Journal of Parapsychology, 13*, 27–35.

Pekala, R. J. (1991). *The Phenomenology of Consciousness Inventory (PCI).* West Chester, PA: Mid-Atlantic Educational Institute.

Pekala, R. J., & Forbes, E. (1997). Types of hypnotically (un)susceptible individuals as a function of phenomenological experience: Towards a typology of hypnotic types. *American Journal of Clinical Hypnosis, 39,* 212–224.

Pekala, R. J., Kumar, V. K., & Cummins, J. (1992). Types of hypnotically susceptible individuals and reported attitudes and experiences of the paranormal and the anomalous. *Journal of the American Society for Psychical Research, 86*, 135–150.

Playfair, G. (2013) *Twin telepathy*. Guilford, UK: White Crow Books.

Ray, W. J. (1996). Dissociation in normal populations. In L. K. Michelson & W. J. Ray (Eds.), *Handbook of dissociation* (pp. 51–66). New York: Plenum Press.

Richards, D. G. (1991). A study of the correlations between subjective psychic experiences and dissociative experience. *Dissociation, 4*, 83–91.

Riley, K. C. (1988). Measurement of dissociation. *Journal of Nervous and Mental Disease, 176,* 449–450.

Roxburgh, E. (2007). The psychology and phenomenology of mediumship: An exploratory study. *Proceedings of Presented Papers: The Parapsychological Association 50th Annual Convention*, 203–207.

Roxburgh, E., & Roe, C. A. (2011). A survey of dissociation, boundary thinness, and psychological wellbeing in Spiritualistic mental mediumship. *Journal of Parapsychology, 75*, 279–300.

Schredl, M., & Hofmann, F. (2003). Continuity between waking activities and dream activities. *Consciousness and Cognition, 12,* 298–308.

Schredl, M., Jochum, S., & Souguenet, S. (1997). Dream recall, visual memory, and absorption in imaginings. *Journal*

of Personality and Individual Differences, 22, 291–292.
Sidgwick, E. M. (1915). A contribution to the psychology of Mrs Piper's trance phenomena. *Proceedings of the Society for Psychical Research, 28*, 1–652.
Simmonds-Moore, C. (2009). Anomalous experiences and boundary thinness in the mind and brain. In M. D. Smith (Ed.), *Anomalous experiences* (pp. 177–206). Jefferson, NC: McFarland.
Simmonds-Moore, C., & Holt, N. (2007). Trait, state and psi: An exploration of the interaction between individual differences, state preference and psi performance in the ganzfeld and a waking ESP control. *Journal of the Society for Psychical Research, 71,* 197–215.
Soffer-Dudek, N., & Shahar, G. (2009). What are sleep-related experiences? Associations with psychological distress and life stress. *Consciousness and Cognition, 18*, 891–904.
Spitzer, C., Haug, H. J., & Freyberger, H. J. (1997). Dissociative symptoms in schizophrenic patients with positive and negative symptoms. *Psychopathology, 30,* 67–75.
Stanford, R., & Stein, A. (1994), A meta-analysis of ESP studies contrasting hypnosis and a comparison condition. *Journal of Parapsychology, 58*, 235–270.
Stolt, C-M., & Björkhem-Bergen, L. (2004). Hypnosis in Sweden during the twentieth century—The life and work of John Björkhem. *History of Psychiatry, 15,* 193–200.
Tart, C. (1972). Measuring the depth of an altered state of consciousness, with particular reference to self-report scales of hypnotic depth. In E. Fromm, & R. E. Shor, (Eds.). *Hypnosis—Research developments and perspectives* (pp. 445–480). Chicago, IL: Paul Elak.
Tart, C. (1977). Discrete states of consciousness. In P. Lee, R. E. Ornstein, D. Galin, A. Deikman, & C. T. Tart (Eds.), *Symposium on Consciousness* (pp. 89–176) Harmondsworth, England: Penguin Books.
Terhune, D. B. (2010). *Dissociation and heterogenity in high hypnotic suggestibility* (Doctoral dissertation). Lund University, Sweden.
Tressoldi, P., & Del Prete, G. (2007). ESP under hypnosis: The role of induction instructions and personality characteristics. *Journal of Parapsychology, 71*, 113–137.
Waller, N. G., Putnam, F. W., & Carlson, E. B. (1996). Types of dissociation and dissociative types: A taxometric analysis of dissociative experiences. *Psychological Methods, 1,* 300–321.
Watson, D. (2001). Dissociations of the night: Individual differences in sleep-related experiences and their relation to dissociation and schizotypy. *Journal of Abnormal Psychology, 110,* 526–535.
Watson, D. (2003). To dream, perchance to remember: Individual differences in dream recall. *Personality and Individual Differences, 34,* 1271–1286.
Wright, D. B., & Loftus, E. F. (1999). Measuring dissociation: Comparison of alternative forms of the Dissociative Experiences Scale. *American Journal of Psychology, 112,* 497–519.

Acknowledgements

The author wishes to thank in alphabetical order: Brian Millar, Nemo Mörck, Donald West, and the two anonymous reviewers for their comments.

Department of Psychology
University of Gothenburg
Box 500
40530 Gothenburg, Sweden
Adrian.Parker@psy.gu.se

Abstracts in Other Languages

French

LA JUNGLE DU PSI HYPNOTIQUE : DEUXIÈME PARTIE.
LA RECHERCHE SUR LES RELATIONS ENTRE LE PSI ET L'HYPNOSE

RÉSUMÉ : La recherche sur l'hypnose et le psi a produit des résultats corrélationnels discutables. Ce passage en revue indique que toutes les études contemporaines ont des faiblesses méthodologiques majeures

et échouent à produire des résultats reproductibles et à prédire de futures performances psi. Nous faisons une critique de l'approche corrélationnelle comme pouvant être ramenée à la mise en lien de tentatives pour deviner des cibles et des variables psychologiques faibles dont la validité est douteuse. Le résultat net est que la recherche sur le psi hypnotique, plutôt que de satisfaire les promesses d'une relation fonctionnelle basée sur un terrain solide, s'est embourbée avec des variables corrélationnelles, des biais méthodologiques et des résultats apparemment non reproductibles. Nous affirmons donc que les recherches doivent être réorientées sur les variables les plus prometteuses, à savoir la « dissociation positive » et la « schizotypie positive ». En l'absence de projets sophistiqués à grande échelle, nous suggérons que des progrès pourraient être fait par les études sélectionnant des participants doués ou ceux qu'on appelle des virtuoses. Dans ce contexte, l'opposition problématique entre santé et pathologie reste à résoudre. Les problèmes soulevés sont considérés comme fondamentaux pour le développement d'une théorie de la conscience.

German

DER DSCHUNGEL DES HYPNOTISCHEN PSI: TEIL 2.
FORSCHUNGEN ZU BEZIEHUNGEN ZWISCHEN PSI UND HYPNOSE

ZUSAMMENFASSUNG: Forschungen zu Hypnose und Psi haben zu umstrittenen korrelativen Ergebnissen geführt. Diese Forschungsübersicht kommt zu dem Ergebnis, dass alle gegenwärtigen Studien erhebliche methodologische Schwachstellen aufweisen, zu keinen replizierbaren Ergebnissen geführt haben und nicht in der Lage waren, zukünftige Psi-Treffer vorherzusagen. Der korrelative Zugang wird kritisiert, da er im wesentlichen aus Rateverhalten in Bezug auf schwache psychologische Variablen von zweifelhafter Validität besteht. Unterm Strich bleibt übrig, dass die Forschung zu Psi unter Hypnose – anstatt die Hoffnung nach einem funktionellen Zusammenhang zu erfüllen, der auf eine sichere Grundlage hat – sich festgefahren hat mit Korrelationsvariablen, methodologischen Mängeln und offenbar nicht-replizierbaren Ergebnissen. Es wird für eine Neuausrichtung der Forschungsbemühungen auf die erfolgversprechendsten Variablen geworben, nämlich auf "positive Dissoziation" und "positive Schizotypie". Angesichts des Fehlens ausgearbeiteter und differenzierter Forschungsprojekte wird vorgeschlagen, dass Fortschritte am ehesten zu erzielen sind, wenn ausgewählte Starteilnehmer oder sogar sogenannte virtuose Spitzenkönner untersucht werden. In diesem Zusammenhang stellt sich die Frage nach Gesundheit-versus-Pathologie als grundlegend heraus, die gelöst werden muss. Die hier aufgeworfenen Fragen sind für eine Bewusstseinstheorie von fundamentaler Bedeutung.

Spanish

LA SELVA HIPNÓTICA DE PSI: PARTE 2.
INVESTIGACIÓN SOBRE LAS RELACIONES ENTRE PSI E HIPNOSIS

RESUMEN: La investigación sobre la hipnosis y psi ha producido resultados correlacionales discutibles. Esta crítica indica que todos los estudios contemporáneos tienen importantes deficiencias metodológicas y no pudieron producir resultados replicables y predecir resultados psi futuros. Hago una crítica del enfoque correlacional como relacionado en gran medida a adivinanzas con variables psicológicas débiles de dudosa validez. El resultado neto es que la investigación de psi hipnótica, en lugar de cumplir con la promesa de una relación funcional sólida, se ha empantanado en variables de correlación, fallas metodológicas, y lo que parecen ser resultados no replicables. Arguyo que se deben reorientar los esfuerzos de investigación en las variables más prometedoras, a saber la disociación "positiva" "y la "esquizotipia positiva. "En la ausencia de proyectos sofisticados a gran escala, sugiero que se puede progresar más estudiando participantes seleccionados estrellas o incluso a los llamados virtuosos. En este contexto, el problema de salud frente a la la patología es básico y queda por resolver. Las cuestiones planteadas son fundamentales para una teoría de la consciencia.

Journal of Parapsychology, 79(1), 53–64

BEWARE OF INFERENTIAL ERRORS AND LOW POWER WITH BAYESIAN ANALYSES: POWER ANALYSIS IS NEEDED FOR CONFIRMATORY RESEARCH

By J. E. Kennedy

ABSTRACT. Errors in inference can occur with any hypothesis testing method, including Bayesian analysis. The evaluation of expected rates of inferential errors is important when planning confirmatory research, but inferential errors have rarely been addressed in writings on Bayesian hypothesis testing. The present investigation applied classical and Bayesian hypothesis testing methods to binomial data with certain effects and to data simulating the null hypothesis. The Bayesian analyses generally had substantially lower power (probability of correctly detecting an effect), particularly for small effect sizes. For data with a small effect size and power of .80 for a classical analysis, the probability that the Bayes factor with a uniform prior correctly reached 3 or higher supporting the alternative model was only .173. The probability that the Bayes factor was 3 or higher incorrectly supporting the null model was .619. These findings verify that quantitative evaluation of expected inferential error rates is essential when designing confirmatory studies that use Bayesian analyses. The argument that biases in favor of the null model are appropriate for small effect sizes because of potential methodological problems is based on exploratory research and is not appropriate for well-designed confirmatory research that focuses on a pre-established effect size.

Keywords: Bayesian analysis, hypothesis test, inferential errors, statistical power, confirmatory research

Confirmatory research is the foundation for valid scientific findings. Exploratory research is usually the creative step that is the starting point for a line of research. However, exploratory research is prone to various questionable methodological practices (Ioannidis, 2012; Kennedy, 2014b; Wagenmakers, Wetzels, Borsboom, van der Maas, & Kevit, 2012). Confirmatory research provides the convincing evidence that makes science valid and self-correcting. Exploration and confirmation are both essential for science. In the past few decades, research in the social sciences has developed an unhealthy emphasis on exploratory research without adequate consideration of confirmation (Ioannidis, 2012; Wagenmakers et al., 2012). Fortunately, a more balanced perspective has been rapidly emerging (e.g., Open Science Collaboration, 2012), although some psychological researchers currently continue to advocate statistical methods that blur the distinction between exploration and confirmation (see Kennedy, 2015). Meta-analysis of exploratory studies does not eliminate the need for well-designed confirmatory research (Kennedy, 2013a).

Confirmatory methodology for experiments includes certain key practices: prespecification of the statistical methods and the criteria for acceptable evidence, sample sizes based on power analysis, public prospective registration of experiments, experimental procedures that make intentional or unintentional data alterations by one person difficult, documented formal validation of software, and sharing data for analyses by others (Kennedy, 2013a, 2013b, 2014b; KPU Registry, 2014). These practices are based on established confirmatory methodology in regulated medical research and on widely recognized principles for good research methodology. The Koestler Parapsychology Unit study registry now provides public prospective registration for parapsychological experiments (KPU Registry, 2012; Watt & Kennedy, 2015).

The quantitative evaluation of potential inferential errors is a fundamental factor for planning confirmatory research. Errors in inference and power analysis have long been standard topics in statistical textbooks for psychologists (e.g., Hays, 1963; Keppel, 1973). Exploratory research usually focuses on either *p*

values or effect sizes, with little or no consideration of sample size, power, and inferential errors. Properly designed confirmatory research incorporates an explicit, balanced recognition of the interactions among effect size, sample size, statistical significance, and inferential errors.

Evaluation of inferential errors and power quantifies the statistical validity of a planned hypothesis test. These evaluations determine the rates of correct and incorrect inferences if the true effect size is a certain value, and the corresponding rates if the null hypothesis is true. The evaluations can be done for different effect sizes to form a curve that covers the range of effects of interest. This curve represents the operating characteristics for the hypothesis test. For confirmatory research, a certain minimum effect size is often of interest and is the focus of the evaluation.

In well-designed confirmatory research, all analysis decisions that could affect the experimental results are made prior to data collection. These decisions include the specific statistical methods, criteria for acceptable evidence, specification of any data adjustments, and criteria for excluding any data from the analyses. If this information cannot be prespecified, the study is more exploratory than confirmatory. These methodological decisions should be publicly registered before data collection begins.

The use of Bayesian analysis is rapidly increasing in scientific research. The basic philosophy, assumptions, and models for Bayesian analyses were described conceptually in a previous paper (Kennedy, 2014a). That paper also pointed out the need for direct comparisons of Bayesian and classical methods for confirmatory research.

Kruschke (2011, p. 321) defines statistical power in a Bayesian context as "[t]he probability of achieving the goal [of the study], given the (suspected) true state of the world." He describes the value of using simulated data to evaluate power when planning research, but he does not address all types of inferential errors.

Most writings on Bayesian methods have focused on exploratory research without considering inferential errors or prespecifying the criteria for acceptable evidence. Few writings address confirmatory practices. One notable exception is the document by the U.S. Food and Drug Administration (2010) that has recommendations for the use of Bayesian methods when seeking approval of medical devices. The document recommends that the experimental design and protocol include the prior probabilities and statistical models that will be used, the criteria for acceptable evidence, and the operating characteristics for type I errors and power.

The present paper quantitatively compares the power and rates of inferential errors for classical and Bayesian analyses for examples of confirmatory experiments that use binomial analysis. The primary purpose is to investigate and to verify the need to evaluate statistical power and inferential error rates for Bayesian hypothesis tests.

Classical Power Analysis

The decision process for a classical hypothesis test for confirmatory research is to accept or to reject the null hypothesis. If the null hypothesis that the results are due to chance is rejected, the experiment is interpreted as providing evidence for the alternative or experimental hypothesis. Thus, the experiment has two possible outcomes—the null hypothesis is either accepted or rejected. Note that this decision process applies to well-designed confirmatory research but not to the more common exploratory analyses that are typically underpowered. In underpowered studies, one cannot accept or support the null hypothesis because nonsignificant results could be due to low power rather than to the null hypothesis being true.

Classical power analysis determines the sample size needed for an experiment to have a high probability of reaching the correct conclusion. The power analysis must consider the probability of rejecting the null hypothesis when the alternative hypothesis is true and the probability of accepting the null hypothesis when the null is true. Standard power calculations determine the sample size from four factors: (a) the desired power of the experiment, which is the probability that the null hypothesis will be rejected if the alternative or experimental hypothesis is true, (b) the effect size for the alternative hypothesis, (c) the alpha or significance level for the analysis, which is the probability of making a type I error that incorrectly rejects the null hypothesis when it is true, and (d) whether the statistical test is one- or two-sided.

For confirmatory research, the effect size for a power calculation is typically based on previous research and/or on an effect that would be meaningfully convincing or useful. The usual recommendations are that the power be at least .8 for confirmatory research and preferably higher, such as .95. The alpha or significance level is typically set at .05. However, the convention for alpha = .05 has become overly rigid and does not consider the nature of the phenomenon being investigated or the distinction between exploratory and confirmatory research. For confirmatory research on controversial effects, an alpha of .01 may be more appropriate. On the other hand, an alpha larger than .05 is often appropriate for small exploratory studies. The fixation on alpha = .05 has promoted ambiguity about whether research is exploratory or confirmatory.

When the effect size for the alternative hypothesis is estimated from previous data, the estimate may not be accurate. If the true effect size is greater than the estimate, the experiment will have greater power than the calculated power. If the true effect size is less than the estimate, the experiment will have lower power. Good practice is to consider the confidence interval for the effect size estimate and to use an effect size for the power calculation that is from the lower part of the confidence interval.

Power calculators are available online and as free programs—notably the G*Power program (Faul et al., 2007; Faul, Erdfelder, Lang, & Buchner, 2012). For binomial power calculations, some calculators use a normal approximation that is not accurate for small samples sizes. Exact calculations, such as performed by the G*Power program, are preferable for small samples sizes.

Background for Bayesian Analysis

Bayesian analysis is based on the philosophical position that probability reflects uncertainty in the human mind rather than uncertainty in the physical world (Kennedy, 2014a). Classical analysis is based on models of uncertainty in the physical world. Bayesian analysis requires models for the human mind in addition to models of processes in the physical world. The analysis starts with *prior probability distributions* that model the beliefs and uncertainties in a human mind prior to the experiment. These prior probabilities are updated based on the results of the experiment to produce *posterior probability distributions* that represent what a person should believe after the experiment. Obviously, a Bayesian analysis produces different results for different prior probabilities or beliefs. *Objective Bayesian methods* attempt to minimize the subjective aspects and potential biases of prior probability distributions.

The *Bayes factor* is a measure of the evidence from the current study and has a key role in Bayesian hypothesis testing. It is the ratio of the posterior probability for the experimental outcome if the outcome was produced by the alternative model divided by the posterior probability for the experimental outcome if the outcome was produced by the null model. The calculation of posterior probability for the alternative model requires a prior probability distribution for the effect size for the alternative model. The debates about Bayesian results often hinge on differing opinions about the prior probability distribution for an effect size. Objective Bayesian hypothesis tests typically focus on the Bayes factor.

Widely accepted conventions have not been established for the magnitude of the Bayes factor (or odds) that is considered acceptable evidence. Discussions of this topic usually refer to Jeffreys (1961), who said he used an odds of 3 the way a classical analyst uses $p = .05$ and an odds of 10 the way a classical analyst used $p = .01$. Most applications of Bayesian hypothesis tests have been for exploratory research and have not specified a criterion for acceptable evidence.

An important feature of Bayesian analysis is that the Bayes factor can be inverted to give the odds that the results were produced by chance, as assumed for the null model. The criterion for acceptable evidence can be applied to the null model as well as to the alternative model. This can provide evidence directly supporting the null hypothesis.

A Bayesian hypothesis test can have three possible outcomes. The Bayes factor can exceed the criterion supporting the alternative model, or it can exceed the criterion supporting the null model, or it can fall into the intermediate zone that does not convincingly support either model. An experiment with a small sample size will likely have the latter result.

Proponents of Bayesian analysis have sometimes argued that type I error and power are classical concepts that should not be considered with Bayesian analysis. That argument is basically a philosophical

stance that virtually eliminates any discussion of inferential errors or accountability for an analyst. Fortunately, the majority of analysts appear to recognize that inferential errors can occur with Bayesian analyses and should be addressed in scientific research, particularly confirmatory research. Kruschke (2011) discusses statistical power from a Bayesian perspective.

Evaluations of inferential error rates and power determine whether the methodological decisions about prior probabilities, specific statistical models, acceptance criteria, and sample size combine to form an effective process for making scientific inferences about a phenomenon. Researchers planning a confirmatory study must make decisions for each of these factors, and those choices affect the probability of making an incorrect inference. These decisions should be based on an understanding of how the different possible choices affect the operation of the hypothesis test. The effectiveness of a confirmatory hypothesis test needs to be quantitatively evaluated with all of the decisions for these factors operating together.

Investigation Plan

A basic strategy for evaluating statistical decision-making processes is to apply the processes to data with known properties, as in the common practice of using simulations to evaluate and compare statistical methods. This strategy can be used to evaluate any decision-making process, including Bayesian hypothesis tests. For purposes of evaluation, a hypothesis test can be conceptualized as a black box that has an input of data and outputs a decision. The black box can be any type of hypothesis test and is evaluated by observing the output when the input data have known properties. The same input data can be used to evaluate different hypothesis testing methods.

The first step is to develop a model for generating data that simulates an effect that the experiment is intended to detect, and a model that generates data for the null hypothesis. The planned statistical analysis is applied to data from each model, and the rates of correct and incorrect inferences determined. This pragmatic approach underlies traditional power analysis and indicates the statistical validity of the hypothesis test. An appropriate effect size or range of effect sizes for these evaluations is usually obvious for confirmatory research.

These evaluations estimate the probabilities of making correct and incorrect inferences for an experiment if the effect has a specific effect size. For binomial data, the effect is the value of the parameter P in a binomial model. Most binomial hypothesis tests assume that the effect being investigated has a fixed but unknown value of P. As discussed in the previous paper (Kennedy, 2014a), Bayesian prior probability distributions represent the uncertainties in the beliefs in a human mind, not variability or random effects in the phenomenon being investigated. The prior probabilities are part of the hypothesis test, not part of the models for generating data to evaluate the hypothesis test. The data for evaluating the hypothesis-testing methods simulate conditions in the external world. Thus, a fixed value of P is used to simulate a phenomenon for a typical binomial hypothesis test.

In the comparisons described below, the sample sizes for an experiment were determined using classical power analysis, and then the corresponding probabilities of correct and incorrect inferences for a Bayesian analysis were determined for that sample size and effect size. A Bayes factor of 3 was set as the criterion for significant or acceptable evidence for the alternative model or for the null model. The number of hits for obtaining a Bayes factor of 3 was found, and the probabilities of reaching that number were determined for both the null model and the modeled effect.

As noted above, a Bayesian hypothesis test has three possible outcomes. The evaluations here provide the probability of each possible outcome if the null hypothesis is true and if the modeled alternative hypothesis is true.

The cases described below are for studies with $P = .5$ for the null hypothesis, as is typical for parapsychological experiments with random event generators (REGs). Two different effect sizes were investigated as alternative models, $P = .53$ and $P = .503$. Both are within the range of effects reported as evidence in parapsychological experiments. For each effect size, the sample size was determined for three different sets of parameters for the classical power analysis: (a) alpha = .05 and power = .80, (b) alpha = .05 and power = .95, (c) alpha = .01 and power = .95. These options cover the range of power analysis that would

reasonably be used in designing a confirmatory experiment. The classical sample sizes were determined using the G*power program (Faul et al., 2012).

The comparisons were based on two-sided tests. This was done because two-sided tests are often recommended in general and are specifically appropriate for parapsychology given the established occurrence of psi missing. Psi missing has been prominent in studies with REGs. Also, two-sided analyses are the only option for the online binomial Bayes factor calculator that was used.

The Bayes factor analysis used the online binomial Bayes factor calculator provided by Rouder (2012). As is common for Bayesian binomial analysis, this calculator has a beta distribution for the prior probability and requires that the two beta parameters be specified. The parameters beta(1,1) were used for one analysis. These provide a uniform prior distribution that is frequently recommended as an objective Bayesian prior and is based on the assumption that any hit rate is equally possible between 0% and 100%. However, this "objective" prior distribution does not reasonably represent the beliefs of either proponents or skeptics of psi. A second analysis was done with parameters beta(22,22). These parameters represent a belief that effects between .4 and .6 are most likely. This distribution is symmetric and has 82% of the distribution between .4 and .6. The beta(22,22) prior is more representative of beliefs based on previous research than is the beta(1,1) prior. However, the selection of this prior distribution was somewhat arbitrary, and arguments can be made that a different prior distribution is preferable. Differing opinions about prior probability is the nature of Bayesian analysis. The purpose here is to investigate the value of power analysis and inferential error evaluations for Bayesian hypothesis tests. Identification and justification of an optimal prior distribution for a particular situation is a different task, and it may be affected by the findings from the present investigation. Online graphs and tables of the beta distribution (e.g., Casio Computer Company, 2014) can be used to display and explore the different options.

Once the number of hits for a Bayes factor of 3 was determined for the different sample sizes and beta parameters, the Stat Trek (2014) online binomial calculator was used to calculate the probabilities for the different possible outcomes under the null and alternative models. The binomial calculator gives cumulative probabilities that are equivalent to those obtained from running an extremely large number of simulations. The Appendix provides more detailed information about this process and gives the cutoff values that were used. Because the distributions investigated here are symmetric, the evaluations for a positive deviation give the same results as those for a negative deviation. For example, the evaluation of $P = .53$ gives the same inferential error rates and power as an evaluation of $P = .47$.

The Jeffreys prior that uses beta(.5,.5) is sometimes recommended as an objective prior and was also initially examined. This is a U-shaped distribution that is higher on the tails near 0 and 1 and thus is even less realistic than the uniform distribution. When it became apparent that this prior gives lower power than the uniform distribution, further consideration of it was rejected.

Results

The results are shown in Tables 1 and 2 for the two different effect sizes. The null model is designated as H_0 in the tables and the alternative model as H_1. The column for "Probability $BF(H_1) \geq 3$" under "If H_1 is True" gives the basic power for the Bayesian analysis, that is, the probability of obtaining a Bayes factor significantly (≥ 3) supporting the alternative model if the alternative model is true. The column for "Probability $BF(H_1) \geq 3$" under "If H_0 is True" gives the alpha level or probability of type I error for the Bayesian analysis, that is, the probability of obtaining a Bayes factor significantly in favor of the alternative model if the null model is true. The column for "Probability $BF(H_0) \geq 3$" under "If H_1 is True" is important because it gives the probability of obtaining a Bayes factor significantly in favor of the null model if the alternative model is actually true.

The Bayesian analyses generally have substantially lower power than the classical analyses, particularly for the uniform prior distribution. As shown in Table 1 for the hit rate of 53%, when the classical analysis has a power of .80, the power for the Bayes factor with a uniform prior is .393, and the probability of obtaining a Bayes factor that incorrectly supports the null model is .294. The probability of making a type I error with the uniform prior is .002 in all three cases and, thus, is much lower than the classical alpha level.

The power increases and error rates decrease for the beta(22,22) prior and as sample size increases. For the case with classical power of .95 and alpha of .01, the Bayesian power and probability of type I error for the beta(22,22) prior are .944 and .008 respectively, which are very close to the classical values.

The lower power and higher associated errors are dramatic for the smaller hit rate of 50.30%. As shown in Table 2, when the classical analysis has a power of .80, the power for the Bayes factor with a uniform prior is only .173, and the probability of obtaining a Bayes factor that incorrectly supports the null model is .619. With the beta(22,22) prior, the probability of the Bayes factor incorrectly supporting the null hypothesis is still greater than the probability of it correctly supporting the alternative model (.386 versus .318). For all the cases in Table 2, the Bayesian analysis has substantially less power, and in most of them it is substantially biased in favor of the null model. The probability of making a type I error is essentially zero.

Table 1

Inferential Errors and Power for Classical and Bayesian Binomial Analyses for H_1 Effect Size = 53.00% and H_0 (Chance) = 50.00%

Classical Power Analysis			Analysis for Bayes Factor (using *N* from the classical power analysis)						
Alpha	Power	*N*	Prior Beta Parameters	If H_1 is True (*P* = .53)			If H_0 is True (*P* = .50)		
				Probability $BF(H_1) \geq 3$	Probability $BF(H_0) \geq 3$	Probability $BF < 3$	Probability $BF(H_1) \geq 3$	Probability $BF(H_0) \geq 3$	Probability $BF < 3$
.05	.80	2,189	1,1	.393	.294	.313	.002	.976	.022
			22,22	.629	.069	.302	.014	.814	.172
.05	.95	3,613	1,1	.672	.106	.222	.002	.982	.016
			22,22	.844	.017	.139	.010	.866	.124
.01	.95	4,963	1,1	.846	.035	.119	.002	.984	.014
			22,22	.944	.004	.052	.008	.888	.104

$BF(H_1) \geq 3$ indicates that the Bayes factor is 3 or greater, supporting the alternative model.
$BF(H_0) \geq 3$ indicates that the Bayes factor is 3 or greater, supporting the null model.
$BF < 3$ indicates that the Bayes factor is less than 3 for both the alternative and null models.

Table 2

Inferential Errors and Power for Classical and Bayesian Binomial Analyses for H_1 Effect Size = 50.30% and H_0 (Chance) = 50.00%

Classical Power Analysis			Analysis for Bayes Factor (using *N* from the classical power analysis)						
Alpha	Power	*N*	Prior Beta Parameters	If H_1 is True (*P* = .503)			If H_0 is True (*P* = .500)		
				Probability $BF(H_1) \geq 3$	Probability $BF(H_0) \geq 3$	Probability $BF < 3$	Probability $BF(H_1) \geq 3$	Probability $BF(H_0) \geq 3$	Probability $BF < 3$
.05	.80	218,187	1,1	.173	.619	.208	.000	.998	.002
			22,22	.318	.386	.296	.002	.988	.010
.05	.95	361,059	1,1	.417	.337	.246	.000	.998	.002
			22,22	.601	.160	.239	.000	.992	.008
.01	.95	494,843	1,1	.643	.162	.195	.000	.998	.002
			22,22	.795	.061	.144	.000	.992	.008

$BF(H_1) \geq 3$ indicates that the Bayes factor is 3 or greater, supporting the alternative model.
$BF(H_0) \geq 3$ indicates that the Bayes factor is 3 or greater, supporting the null model.
$BF < 3$ indicates that the Bayes factor is less than 3 for both the alternative and null models.

Discussion and Conclusions

These findings verify that quantitative evaluation of inferential errors is needed for confirmatory research with Bayesian analysis, as well as for confirmatory research with classical analysis. The investigations found that the Bayesian analyses of the binomial data tended to have lower power and higher rates of associated inferential error than the classical hypothesis tests. For small effect sizes, the Bayesian analyses were strongly biased toward supporting the null model when the alternative model was true.

The analyses here should not be taken as evidence that classical hypothesis tests are generally superior to Bayesian methods, or as recommendations about prior probabilities or the magnitude that a Bayes factor must reach to provide acceptable evidence. Rather, my expectation is that useful prior probabilities and criteria for acceptable evidence can be found for Bayesian analyses if inferential errors and power are evaluated. The primary conclusion here is that these evaluations are essential for confirmatory research.

Small Effects

The differences in conclusions between Bayesian and classical analyses for small effects are well known among statisticians and are a manifestation of the *Jeffreys-Lindley paradox*. Diffuse prior probability distributions, such as a uniform distribution, enhance this paradox. Some proponents of Bayesian methods have argued that these differences between classical and Bayesian results indicate that classical methods are flawed (e.g., Jefferys, 1990; Wagenmakers & Grünwald, 2005). This argument appears to be based on the assumption that Bayesian methods are theoretically optimal and should be the standard for evaluating other statistical methods. However, a more common and less disputable strategy for evaluating statistical methods is to use data that simulate effects of interest. These more empirical evaluations are less susceptible to overly optimistic theoretical ideas. The evaluations here show that Bayesian analyses can be strongly biased in favor of the null model when the alternative model is actually true. Similar conclusions have been reached by others (e.g., Bem, Utts, & Johnson, 2011; Dobyns, 1992).

Some defenders of Bayesian methods attempt to put a positive spin on the bias against small effects by claiming that small effects are usually due to methodological artifacts or to slight meaningless effects for the phenomenon being investigated (e.g., Jefferys, 1990). That argument may have some plausibility for exploratory research, but it is not appropriate for well-designed confirmatory research that is attempting to verify previous findings. The argument is basically a speculation that is impervious to empirical data. Larger sample sizes cannot overcome these a priori assumptions that the effects are meaningless. The result is that small effects could be largely excluded from scientific findings with Bayesian analysis—a position that is inconsistent with the basic principles of empirical scientific research.

A more appropriate strategy for reaching valid scientific conclusions is to conduct well-designed confirmatory research rather than use biased statistical methods. Parapsychological research is just beginning to implement this type of confirmatory methodology.

Counterintuitive, Arbitrary Priors

The prior probabilities in Bayesian hypothesis testing can have counterintuitive consequences. For example, a uniform prior probability for effect size appears on the surface to represent a very open-minded prior belief that is typically described as "objective." However, a diffuse prior like this makes any specific effect size have a relatively low probability—including the specific effect size that is the experimental outcome. A low prior probability for the experimental outcome makes the Bayes factor favor the null hypothesis. In effect, the power of the hypothesis test is spread over an unrealistic "objective" range that results in the analysis favoring the null hypothesis. On the other hand, a narrow prior probability distribution can have a much higher or much lower probability for a particular effect size, depending on the details of the distribution and the specific effect.

The evaluation of inferential errors using simulated data for an effect of interest is very useful for revealing counterintuitive biases for a planned Bayesian hypothesis test. The Jeffreys-Lindley paradox is another manifestation of counterintuitive effects that can bias a hypothesis test.

In practice, the selection of a prior probability distribution is often substantially arbitrary. A wide range of prior probability distributions from diffuse to narrow usually can be justified as reasonably plausible. The beta(22,22) prior distribution was somewhat arbitrarily selected for the present investigation and will likely be considered inappropriate by some researchers. I cannot defend it as optimal or as specifically representing my personal beliefs. I can only say that it appears to me to be within the wide range of plausible priors.

Inferential errors and power can be useful factors in selecting prior probabilities for confirmatory research. Most writings on Bayesian analysis focus on the exploratory stage of research and recommend retrospective sensitivity analysis of prior probabilities. However, for confirmatory research, the sensitivity analyses need to be done at the planning stage and the selected priors included in the preregistered study information.

If an analysis has a high probability of inferential errors for an effect size that is of primary interest to the experimenter, the design and/or analysis need to be modified. For confirmatory research, this situation needs to be discovered at the planning stage rather than through retrospective sensitivity analysis. An efficient strategy may be to find a prior probability distribution that gives useful rates of inferential errors, and then consider whether the selected prior is within the range of plausible priors.

The potential for inferential errors also affects the interpretation of the final experimental results. For a classical analysis with low power, a nonsignificant result is ambiguous because the result could be due to low power or to the experimental hypothesis being false. A more fundamental point is that the absence of a careful power analysis typically indicates exploratory methodology that is prone to various questionable practices that can be difficult to detect from the final report. This is true whether or not the experimenter describes the research as exploratory. Similarly, for a Bayesian analysis that has a high probability of incorrectly supporting the null model when the alternative model is true, an experimental outcome supporting the null model is ambiguous, and the experiment is likely exploratory and prone to other questionable methodological practices that could be difficult to detect from the final report.

Some Bayesian analysts (e.g., Kruschke, 2011) consider the Bayes factor to be an undesirable hypothesis-testing method because of the high potential for bias. These analysts propose alternative Bayesian methods. If the proposed alternative methods are used for confirmatory research, prespecification of the criteria for acceptable evidence and evaluation of inferential errors and power are needed.

Limitations of Posterior Probabilities

An often underappreciated limitation of Bayesian analysis is that the mathematical functions representing the posterior probabilities do not represent the multifactorial contingencies of actual scientific beliefs. The validity and meaning of the Bayes factor (or *p* value) for an experiment depend on the methodology that was used. The statistical evidence from an experiment is contingent upon good methodology, but key methodological factors (e.g., preregistration of the planned analysis, software validation, and measures to prevent fraud) are not considered in the statistical models. This larger context that is not represented in the mathematical models can be decisive when evaluating the evidence from an experiment. Attempts to quantitatively adjust posterior probabilities and subsequent prior probabilities for these methodological factors are inevitably subjective and imprecise.

Parapsychological research with electronic REGs is a clear example of the importance of the larger context for research. Meta-analyses of experiments with REGs have consistently found that smaller studies have larger effects (Kennedy, 2013a). This pattern is a recognized symptom of methodological bias, but could also be a property of psi (Kennedy, 2013a). Either way, the data are not consistent with a straightforward statistical analysis. The evidence from parapsychological experiments with REGs depends more on a person's opinion about this property of the results than on the specific *p* values or Bayes factors that are produced in an analysis.

Reasonable expectations for convincing experimental evidence are: (a) confirmatory methodology as described in the introduction, and (b) reliable confirmatory results that have properties consistent with the assumptions for the statistical analyses. Parapsychological research and most psychological research have not yet met these standards (Kennedy, 2013a, 2014b; Wagenmakers et al., 2012).

Overall Summary

The key points in this paper can be summarized as:

1. Quantitative evaluation of expected inferential errors and power is essential for planning statistically-based confirmatory research, including research with Bayesian analyses. These evaluations can be done by determining the probability of correct and incorrect inferences for the planned analysis when applied to data with and without the effect that the study is designed to detect. These evaluations are relatively easy for binomial analyses.
2. Much work remains to be done to develop reasonably unbiased Bayesian methods. The results of Bayesian hypothesis tests currently represent explorations of the properties of poorly understood and somewhat arbitrarily selected mathematical functions more than beliefs that a person has or should have. Better understanding of errors in inference will contribute significantly to the development of Bayesian methods.

The first conclusion above is consistent with the U.S. Food and Drug Administration (2010) recommendations on the use of Bayesian methods when seeking approval of medical devices. These recommendations recognize the current uncertainties and limitations of Bayesian analyses and are appropriate for confirmatory research that is expected to receive critical scrutiny (Kennedy, 2014a).

Exploratory analyses present different challenges than those discussed here. One problem is that researchers often report exploratory analyses only if the outcomes are suggestive of an effect. This introduces a fundamental bias for false-positive errors and shows the need for confirmatory research. Attempts to provide convincing evidence from exploratory research by altering the type I error rates do not eliminate the bias from selective reporting and do not eliminate the need for confirmatory research. The debates about Bayesian versus classical statistics have often implicitly focused on attempts to develop convincing results from exploratory research. Psychologists and parapsychologists appear to be reaching the same conclusion as has previously been reached for regulated medical research: well-designed confirmatory research is required for convincing scientific evidence.

References

Bem, D. J., Utts, J., & Johnson, W. O. (2011). Must psychologists change the way they analyze their data? *Journal of Personality and Social Psychology*, *101*, 716–719. Retrieved from http://dl.dropboxusercontent.com/u/8290411/ResponsetoWagenmakers.pdf

Casio Computer Company. (2014). Beta distribution (chart) calculator. Retrieved from http://keisan.casio.com/exec/system/1180573226

Dobyns, Y. H. (1992). On the Bayesian analysis of REG data. *Journal of Scientific Exploration*, *6*, 23–45. Retrieved from http://www.scientificexploration.org/journal/jse_06_1_dobyns.pdf

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2012). G*Power 3 [Computer software]. Retrieved from http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3.

Hays, W. L. (1963). *Statistics*. New York, NY: Holt, Rinehart, and Winston.

Ioannidis, J. P. A. (2012).Why science is not necessarily self-correcting. *Perspectives on Psychological Science*, *7*, 645–654. Retrieved from http://pps.sagepub.com/content/7/6/645.full

Jefferys, W. H. (1990). Bayesian analysis of random event generator data. *Journal of Scientific Exploration*, *4*, 153–169. Retrieved from http://www.scientificexploration.org/journal/jse_04_2_jefferys.pdf

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, England: Clarendon Press.
Kennedy, J. E. (2013a). Can parapsychology move beyond the controversies of retrospective meta-analysis? *Journal of Parapsychology*, *77*, 21–35. Retrieved from http://jeksite.org/psi/jp13a.pdf
Kennedy, J. E. (2013b). [Letter to the editor]. *Journal of Parapsychology*, *77*, 304–306. Retrieved from http://jeksite.org/psi/jp13let.pdf
Kennedy, J. E. (2014a). Bayesian and classical hypothesis testing: Practical differences for a controversial area of research. *Journal of Parapsychology*, *78*, 170–182. Retrieved from http://jeksite.org/psi/jp14.pdf
Kennedy, J. E. (2014b). Experimenter misconduct in parapsychology: Analysis manipulation and fraud. Retrieved from http://jeksite.org/psi/misconduct.pdf
Kennedy, J.E. (2015). Critique of Cumming's "new statistics" for psychological research: A perspective from outside psychology. Retrieved from http://jeksite.org/psi/critique_new_stat.pdf
Keppel, G. (1973). *Design and analysis: A researcher's handbook*. Englewood Cliffs, NJ: Prentice-Hall.
KPU Registry. (2012). Registry for parapsychological experiments. Retrieved from https://koestlerunit.wordpress.com/study-registry/
KPU Registry. (2014). Exploratory and confirmatory analyses. Retrieved from http://www.koestler-parapsychology.psy.ed.ac.uk/Documents/explore_confirm.pdf
Kruschke, J. K. (2011). *Doing Bayesian data analysis: A tutorial with R and BUGS*. Burlington, MA: Academic Press.
Open Science Collaboration. (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science*, *7*, 657–660. Retrieved from http://pps.sagepub.com/content/7/6/657.full
Rouder, J. (2012). Bayes factor for a binomially distributed observation. Retrieved from http://pcl.missouri.edu/bf-binomial
Stat Trek. (2014). Binomial calculator: Online table. Retrieved from http://stattrek.com/online-calculator/binomial.aspx
U.S. Food and Drug Administration. (2010). *Guidance on the use of Bayesian statistics in medical device clinical trials.* Retrieved from http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071121.pdf
Wagenmakers , E., & Grünwald, P. (2005). A Bayesian perspective on hypothesis testing: A comment on Killeen (2005). *Psychological Science*, *17*, 641–642.
Wagenmakers, E., Wetzels, R., Borsboom, D., van der Maas, H. J., & Kevit, R. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science*, *7*, 632–638. Retrieved from http://pps.sagepub.com/content/7/6/632.full.pdf+html
Watt, C., & Kennedy, J. E. (2015). Lessons from the first two years of operating a study registry. *Frontiers in Psychology*, *6* (article 173), 1–4. DOI: 10.3389/fpsyg.2015.00173

Broomfield, CO, USA
jek@jeksite.org

Appendix

The online binomial Bayes factor calculator provided by Jeff Rouder (2012) can be used for two-sided binomial analyses. This calculator provides the Bayes factor that has the posterior probability of the experimental outcome with the null model divided by the posterior probability of the experimental outcome with the alternative model. For $P = .5$, this Bayes factor as a function of the number of hits is a symmetric bell-shaped curve. If a Bayes factor of 3 is considered acceptable evidence, outcomes in the middle of the curve with Bayes factors greater than 3 are evidence supporting the null hypothesis. The points on each side where the Bayes factor just reaches 3 are the cutoff values or critical values for supporting the null hypothesis. Bayes factors of .333 or less on the tails support the alternative hypothesis. If the Bayes factor is inverted to give the posterior probability with the alternative model divided by the posterior probability with the null model, these tail values are 3 or greater supporting the alternative hypothesis. The points on each side where the Bayes factor just reaches .333 or less are the cutoff values or criteria for supporting the alternative hypothesis. Outcomes with a Bayes factor that falls between 3 and .333 do not clearly support either the alternative or null hypothesis. For a given number of trials, prior beta parameters, and P for the null model, the cutoff values can easily be found by trying different numbers of hits until the Bayes factor

just reaches 3 or .333.

Once these cutoff points for the Bayes factor have been determined, the power and probability of errors in inference can be found using standard binomial calculators. The Stat Trek (2014) website provides a useful online binomial calculator that uses exact methods and normal approximations as appropriate.

Classical power can be calculated with a binomial calculator as the probability of reaching the cutoff for significance when *P* in the binomial model is the hit rate (effect size) assumed for the alternative hypothesis in the power evaluation. Likewise, the alpha level or probability of type I error can be calculated with a binomial calculator as the probability of reaching the cutoff for significance when *P* is the hit rate for the null hypothesis.

These same types of calculations can be done using the cutoff criteria for the Bayes factor. The use of the binomial calculator is straightforward once the Bayes factor cutoff values have been found. For the cutoff for the null hypothesis (Rouder's Bayes factor = 3), the relevant cumulative probability is less than or equal to the cutoff. For the cutoff for the alternative hypothesis (Rouder's Bayes factor = .333), the relevant cumulative probability is greater than or equal to the cutoff. The only complication is in handling both sides or tails. For the examples discussed here, the probability of outcomes on the negative side was negligible when the alternative model was applied. Both sides need to be considered when the null model is applied.

The Bayes factor cutoff values for the analyses in Tables 1 and 2 are listed below.

N	Prior Beta Parameters	*N* Hits for $BF(H_1) = 3$	*N* Hits for $BF(H_0) = 3$
2,189	1,1	1,167	1,147
	22,22	1,153	1,125
3,613	1,1	1,902	1,877
	22,22	1,885	1,851
4,963	1,1	2,595	2,566
	22,22	2,575	2,537
218,187	1,1	109,969	109,818
	22,22	109,859	109,680
361,059	1,1	181,676	181,486
	22,22	181,536	181,314
494,843	1,1	248,778	248,559
	22,22	248,617	248,361

Abstracts in Other Languages

French

ATTENTION AUX ERREURS INFÉRENTIELLES ET AUX FAIBLES PUISSANCES DANS LES ANALYSES BAYESIENNES : L'ANALYSE DE PUISSANCE EST NÉCESSAIRE POUR LA RECHERCHE CONFIRMATOIRE

RÉSUMÉ : Les erreurs dans les inférences peuvent survenir avec n'importe quelle méthode pour tester des hypothèses, dont l'analyse bayésienne. L'évaluation des taux attendus d'erreurs inférentielles est importante lorsque l'on prévoit des recherches confirmatoires, mais les erreurs inférentielles ont rarement été discutées dans la littérature sur l'analyse bayésienne. La présente étude applique les méthodes de test classique et bayésienne à des données binomiales avec certains effets et à des données simulant l'hypothèse nulle. Les analyses bayésiennes ont généralement une puissance substantiellement plus faible (probabilité de détecter

correctement un effet), particulièrement avec des petites tailles d'effet. Pour des données avec une petite taille d'effet et une puissance de 0,80 pour une analyse classique, la probabilité que le facteur de Bayes avec une probabilité a priori uniforme atteignant 3 ou plus soit correctement en faveur du modèle alternatif fut seulement de 0,173. La probabilité que le facteur de Bayes fut de 3 ou plus et soutienne incorrectement le modèle nul était de 0,619. Ces données vérifient que l'évaluation quantitative des taux d'erreur inférentielle attendus est essentielle lorsque l'on conçoit des études confirmatoires qui utilisent des analyses bayésiennes. L'argument selon lequel les biais en faveur du modèle nul sont appropriés pour des petites tailles d'effet du fait des potentiels problèmes méthodologiques est basé sur la recherche exploratoire et n'est pas appropriée pour des recherches confirmatoires bien conçues qui se concentrent sur une taille d'effet préétablie.

German

VORSICHT VOR INFERENZFEHLERN UND GERINGER TESTSTÄRKE BEI BAYESSCHEN ANALYSEN: EINE ANALYSE DER TESTSTÄRKE WIRD BEI BESTÄTIGUNGSFORSCHUNG BENÖTIGT

ZUSAMMENFASSUNG: Fehler bei Schlussfolgerungen können bei jeder hypothesenüberprüfenden Methode auftreten, auch bei der Bayesschen Analyse. Die Abschätzung der zu erwartenden Anzahl von Inferenzfehlern ist wichtig, wenn eine Bestätigungsforschung geplant wird, aber Inferenzfehler wurden nur selten in der Literatur über Bayessche Hypothesenüberprüfung diskutiert. In der vorliegenden Untersuchung wurden die klassische und die Bayessche Methode zur Hypothesenüberprüfung auf binomiale Daten bei bestimmten Effekten und auf Daten zur Simulation der Nullhypothese angewandt. Die Bayesschen Analysen weisen im allgemeinen eine deutlich geringere Teststärke auf (die Wahrscheinlichkeit, einen wahren Effekt zu erkennen), besonders bei kleineren Teststärken. Bei auf klassische Weise analysierten Daten mit einer geringen Effektstärke und einer Teststärke von .80 beträgt die Wahrscheinlichkeit, dass der Bayes-Faktor mit einem gleichförmigen Prior von 3 oder höher das alternative Modell zu Recht bestätigt, nur .173. Die Wahrscheinlichkeit, dass der Bayes-Faktor von 3 oder höher das Nullmodell fälschlicherweise bestätigt, beträgt .619. Diese Ergebnisse belegen, wie wichtig die quantitative Abschätzung erwarteter Inferenzfehler ist, wenn man Bestätigungsstudien unter Verwendung von Bayesschen Analysen plant. Das Argument, dass Abweichungen zugunsten des Nullmodells für geringe Effektgrössen angemessen sind wegen potentieller methodologischer Probleme, beruht auf exploratorischer Forschung und ist für eine wohlüberlegte Bestätigungsforschung unangemessen, die auf eine vorher festgelegte Effektstärke abzielt.

Spanish

CUIDADO CON LOS ERRORES INFERENCIALES Y BAJO PODER EN ANÁLISIS BAYESIANOS: SE NECESITA UN ANÁLISIS DE PODER PARA LA INVESTIGACIÓN CONFIRMATORIA

RESUMEN: Los errores en la inferencia pueden ocurrir con cualquier método de prueba de hipótesis, incluyendo al análisis Bayesiano. La evaluación de las tasas esperadas de errores inferenciales es importante en la planificación de la investigación confirmatoria, pero rara vez se han abordado los errores inferenciales en los escritos sobre la prueba de hipótesis Bayesiana. Apliqué métodos de prueba de hipótesis clásicos y Bayesianos a datos binomiales con efectos conocido y datos que simulaban la hipótesis nula. El análisis Bayesiano generalmente tuvo un poder sustancialmente menor (la probabilidad de detectar correctamente un efecto), en particular para los efectos pequeños. Para los datos con un efecto pequeño y poder de 0.80 en un análisis clásico, la probabilidad de que el factor de Bayes con un cálculo previo uniforme de 3 o más apoyando al modelo alternativo fue de sólo 0.173. La probabilidad de que el factor de Bayes fuera 3 o más apoyando incorrectamente al modelo nulo fue 0.619. Estos resultados verifican que la evaluación cuantitativa de los índices de error inferencial esperados es esencial en el diseño de estudios de confirmación que utilizan análisis Bayesianos. El argumento de que los sesgos a favor del modelo nulo son apropiados para los efectos pequeños por posibles problemas metodológicos se basa en la investigación exploratoria y no es apropiado para la investigación confirmatoria bien diseñada que se centra en un efecto de tamaño preestablecido.

Journal of Parapsychology, 79(1), 65–86

THE RELATIONSHIP BETWEEN LABILITY AND PERFORMANCE AT INTENTIONAL AND NONINTENTIONAL VERSIONS OF AN IMPLICIT PMIR-TYPE PSI TASK

By Glenn A. M. Hitchman, Chris A. Roe, and Simon J. Sherwood

ABSTRACT: A number of theories of psi such as Stanford's psi-mediated instrumental response (PMIR) model suggest psi can function without a person's awareness, and that their intent to exhibit psi may be counterproductive. However, few parapsychological studies have directly compared participants' performance at intentional and nonintentional versions of equivalent tasks. This study sought to address this issue whilst exploring the role of lability, suggested by Stanford to be predictive of a person's propensity to respond to extrasensory stimuli. 50 participants took part in both intentional and nonintentional versions of a 10-trial, binary, forced-choice precognition task. A contingent outcome task system involving positive pictures as reward for hit trials and negative pictures as punishment for miss trials was administered on a trial-by-trial basis. Participants scored marginally fewer hits than the mean chance expectation in both versions of the task, with no tangible difference in their performance between tasks. Furthermore, no relationship was found between the number of precognitive hits they achieved and their scores on a composite psychometric measure of lability, nor its constituent elements. However, participants' expectations that their luck could aid their performance, as well as their emotional reactivity, were significantly positively related to their tacit psi scores.

Keywords: extrasensory perception, nonintentional precognition, lability

Throughout the history of parapsychological research, experimenters have used both intentional and nonintentional methods in their attempts to capture psi phenomena. Whilst certain intentional means such as the ganzfeld approach (Bem & Honorton, 1994) have, on occasion, yielded successful outcomes, some theories of psi such as Stanford's (1974, 1977, 1982, 1990) psi-mediated instrumental response (PMIR) model suggest that the wilful intent to produce extrasensory effects may actually hinder the underlying process. A series of recent studies by Luke and colleagues utilised a nonintentional precognition protocol that was consistent with Stanford's conceptualisation of psi as primarily an unconscious process that functions in the service of an organism's needs by activating pre-existing behaviours in response to threats or opportunities in the environment (Luke, Delanoy & Sherwood, 2008; Luke & Morin, 2009; Luke, Roe & Davison, 2008). The method, which was employed relatively consistently throughout the four studies, involved a picture preference task in which participants were asked to select a preferred image from a set of four fractal patterns. Participants were unaware that this was actually a covert test of precognition, as immediately after they indicated their favourite picture the computer would randomly pick one of the four fractal images as a target. Stanford's notion of psi as a goal-oriented process was reflected by a contingent outcome design. At the end of a 10-trial session, participants who had scored more hits than the mean chance expectation (MCE) of 2.50 were rewarded by being able to rate positive images (either erotic pictures aligned to their sexual preference or humorous cartoons), whereas those who had achieved fewer hits than MCE were punished by having to take part in a boring number vigilance task. In each of these studies, participants' mean hit rates exceeded MCE, and three of the four studies yielded independently significant evidence of a nonintentional precognition effect. The combined mean hit rate across all four studies was 2.92 (SD = 1.46), significantly greater than the MCE of 2.50, $t(197) = 4.04$, $p = .00008$, two-tailed), with

an effect size (ES) of $r = .28$. Please note that throughout this paper, effect sizes for t tests are calculated according to the following formula:

$$ESr = \frac{\sqrt{t^2}}{t^2 + df}$$

In addition to describing the functionality of psi, Stanford's PMIR model also makes claims regarding various situational and individual difference factors which may either facilitate or constrain psi from occurring. The studies by Luke and colleagues utilised associated questionnaire methods as rudimentary attempts to assess the impact of two such factors—latent inhibition and lability. Latent inhibition can be understood as the general tendency of an organism to ignore or filter out information from further cognitive processing that it has learned is irrelevant to its ongoing situational concerns (Lubow, 1989). It was assessed indirectly via Goldberg's (1999) measure of openness to experience. Meanwhile, the lability construct (defined more precisely below) is often used as an antonym for, or on a continuum with, stability, and the linear and nonlinear subscales of the Creative Cognition Inventory (Holt, 2002) were used as a convenient proxy measure. Following the assumption that psi may function via extrasensory information transfer, it was predicted that those with higher levels of latent inhibition would be more prone to filtering out extrasensory data at an early stage and hence show diminished performance on the precognition task relative to those with lower levels of latent inhibition (see Holt, Simmonds-Moore, & Moore, 2007). Similarly, it was predicted that more labile individuals would have a greater propensity to respond to subtle extrasensory biases within their cognitive systems and would therefore perform better at the precognition task than more stable individuals.

In the Luke et al. (2008) study, a significant positive relationship was observed between openness to experience and precognition scores, $r = .46$, $p = .01$, two-tailed. However, this correlation was not replicated in the subsequent study by Luke and Morin (2009; $r = -.08$, $p = .64$, two-tailed). Meanwhile, when considering the relationship between tacit psi scores and the lability construct, Luke and Morin (2009) failed to find significant correlations between precognitive performance and either of the subscales of the Creative Cognition Inventory (linear subscale: $r = .25$, $p = .17$; nonlinear subscale: $r = .20$, $p = .27$). Despite the inconclusive results regarding the covariates of tacit precognition, Hitchman, Roe, and Sherwood (2012) were sufficiently encouraged by the overall success of the Luke and colleagues' protocol to attempt to replicate and extend the paradigm. Whilst the core facets of the method were preserved, the design of their study was refined in several ways, including the experimental software program being completely rewritten in an updated programming language to overcome fears that previous results may have been due to an artifact within the code, and the number of trials being increased from 10 to 15 to enhance the statistical power of the study. They did, however, retain all of the questionnaire measures that had been used throughout Luke and colleagues' four studies.

Participants in the Hitchman et al. (2012) replication scored more hits on the nonintentional precognition task than the MCE of 3.75 (mean hit rate = 4.02) but this difference was not significant, $t(49) = 1.14$, $p = .13$, one-tailed. Meanwhile, this study provided some indirect support for the assumed role of latent inhibition via a medium-sized positive correlation between the number of precognitive hits participants achieved and their scores on Goldberg's (1999) openness to experience scale, $r = .29$, $p = .02$, one-tailed. However, little evidence was found to indicate that lability could influence precognitive performance, as participants' nonintentional precognition hit rates were unrelated to their scores on the linear and nonlinear subscales of Holt's (2002) Creative Cognition Inventory: linear subscale; $r = .14$, $p = .16$; nonlinear subscale; $r = .03$, $p = .41$; both one-tailed. However, in both cases it was argued that the measures used to assess these constructs were too indirect to allow any firm conclusions in relation to their effects. In particular, creativity constitutes a very limited proportion of lability, which is a much broader concept reflecting a wider range of facets relating to the readiness for change in an organism. It may be that the particular elements of lability that have been assessed as covariates of psi performance to date are not the most influential, whereas other components of the construct that have not yet been considered may play a more pertinent role in the

psi process. It was therefore considered worthwhile in the present study to explore the relationship between performance at precognition tasks and a more comprehensive measure of lability, with particular attention to how it relates to the PMIR model.

Lability is a construct that was first popularised within parapsychology by Braud (e.g., Braud, 1980, 1981, 2002; Braud & Schlitz, 1983; Braud, Shafer & Mulgrew, 1983). According to Braud (1980, p. 1), lability represents "the ease with which a system can change from one state to another, the amount of 'free variability' in the system." Lability can thus be characterised as the a priori probability (ready capacity) that a system will change its state in a given situation. In this regard, Braud suggested that the brain-mind may, at times, be constrained by specific structural patterns. On these occasions, the brain-mind is said to be in an inert state in which it is resistant to change. Once such structures have been released, he proposes that the brain-mind may be at greater liberty to adapt or reorganise itself in relation to psi-relevant information or events (Braud, 2002). He therefore proposed that psi phenomena are more likely to manifest when individuals are characterised by a more labile (and hence less inert) state.

To illustrate this, Braud offered the concrete example of when a specific pattern of neuronal activity is required in the facilitation of a particular memory. If the neurons required for that pattern of neuronal activity are temporarily engaged in other structures or activities, the specific memory will be provisionally inhibited. Once those neuronal patterns become unstructured or deconstrained, the neurons required for the previously inhibited memory are made available, allowing for the memory to be triggered. The relevance of this particular example to the PMIR model is clear when we consider that one of the mechanisms through which Stanford (1990) claimed psi-mediated responses could be accomplished is via the triggering of pre-existing response mechanisms, which could include particular memory traces.

Despite the conceptual promise of lability, finding an appropriate measure of this construct is less straightforward. Braud et al. (1983) tended to focus on the notions of cognitive and perceptual lability. They measured cognitive lability via an assessment of the fluency of word associations, whereas perceptual lability was measured by assessing how frequently participants' perceptions of the Necker cube alternated between the two potential representations. More recently, Roe and Holt (2006; Holt & Roe, 2006) devised a broader measure of lability that combined various established psychometric measures believed to be indicative of an individual's lability. This composite measure was designed to include a variety of emotional, cognitive, physiological, neurological, and behavioural elements. Specifically, the NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992) was included, particularly as Openness to Experience, a constituent element, is thought to reflect an individual's willingness to engage with novel ideas and values, whereas Neuroticism is said to be indicative of emotional instability and a poor ability to control impulses, with both factors being particularly pertinent to the notion of lability.

The composite scale also included a measure of temporal lobe lability (Personal Philosophy Inventory; Persinger & Makarec, 1987). According to Persinger, those with greater levels of temporal lobe lability tend to be more impulsive, suggestible, emotionally sensitive, imaginative, and have a greater interest in philosophical ideas. He also suggests they may be more prone to psi experiences (Persinger, 1989). To account for lability of mood, two items on Mood Lability designed to screen for bipolar disorder (Akiskal et al., 1995) were included. This disorder is characterised by grand fluctuations of mood, indicating a highly labile state. Finally, following Braud's (1981) assertion that lability is related to novelty generation, two measures of creativity were included (Creative Cognition Inventory; Holt, 2007; Emotional Creativity Inventory; Averill, 1999).

These measures have subsequently been adjusted into a refined lability scale (Drennan, Roe, & Broughton, 2011) consisting of 71 items. This refined measure has an adequate level of internal consistency ($\alpha = .86$) with factor analysis revealing five main elements: Intuitive Cognition (26 items, $\alpha = .92$), Conceptual Cognition (18 items, $\alpha = .42$), Ego-Orientated Cognition (12 items, $\alpha = .78$), Emotional Interpretation (10 items, $\alpha = .71$) and Analytical Cognition (5 items, $\alpha = .79$). Whilst this refined and validated measure was not available at the time of the present study, given that its items were selected from the aforementioned composite elements, the composite scale may be considered a reasonable assessment measure, and it also enables further analysis of its constituent elements. Interestingly, these researchers found a medium-sized

negative correlation between the composite lability scale and spontaneous psychokinetic experiences. This reflects the trend that the majority of research in relation to lability to date has focused on its relationship with psychokinesis, with little research having been conducted in relation to extrasensory perception. This study consequently provided an ideal opportunity to evaluate the role of lability in a PMIR-type precognition task.

In addition to exploring the covariates of psi predicted by Stanford in his PMIR model, this experiment also presented an opportunity to test one of the key predictions of the model by comparing intentional and nonintentional versions of the psi task. According to the PMIR model, psi can function without the conscious intention or awareness of the individual and any consciously generated thinking or cognitive constraints in relation to the need for psi in a life event or experimental situation may significantly diminish the possibility for psi to be manifested. We could therefore expect both intentional and nonintentional psi to be possible in principle, although the potential for intentional psi may be restricted by the cognitive interference associated with participants' awareness of the need to fulfil a psi task.

Few studies in the parapsychological literature have directly compared performance across intentional and nonintentional versions of equivalent or similar tasks. Rao and Davis (1978) performed an experiment designed to assess experimenter effects across nonintentional and intentional psi tasks with a limited sample of 11 female participants. The intentional psi task consisted of a word-based ESP test in which participants were asked to explicitly guess a series of English and Telugu (an unfamiliar language to the participants) target words that were concealed from their conventional sensory faculties. The nonintentional psi task required participants to rank 40 items from a mood adjective check list on a 4-point scale which were later compared against a list of randomly generated target numbers ranging from 1–4 for each adjective. The results of the study indicated a differential language effect in the intentional psi task, with participants scoring significantly higher on English words than Telugu words, but only for one of the experimenters. For the nonintentional psi task, it was found that participants scored significantly higher when they gave different mood ranks in the second of two experimental sessions compared with those who gave the same ranks. Furthermore, the number of mood items checked differently across the two nonintentional psi task sessions was found to correlate significantly with the differential between scores across the two languages in the intentional psi task. These findings would seem to indicate a relationship between participants' performance at intentional and nonintentional psi tasks, with participants who showed a greater tendency towards the differential language effect in the intentional psi task also performing better at the nonintentional psi task. It is interesting to note that the results of the nonintentional psi task are also indicative of a lability effect, with participants who changed their ranks from one session to the next showing heightened performance relative to those who were more rigid in their responses. Overall, however, a direct comparison of the relative strength of intentional and nonintentional psi in this study is limited by the different nature of the two tasks. The present study therefore included both intentional and nonintentional psi tasks of the same type in order to test the predictions of the PMIR model in relation to consciously generated thinking and the cognitive constraints associated with intentionality in a much more direct way.

Regarding the methodological considerations of the present study, it is noteworthy that in the studies involving the Luke and colleagues paradigm described above, participants took part in the contingent positive or negative outcome task (a reward system similar to the feedback mechanisms employed in studies of intentional psi) only at the end of completing a run of 15 trials. However, meta-analyses of forced-choice precognition studies indicate that there is a significant positive relationship between the degree and immediacy of feedback participants receive and the effect size reported for the corresponding study (Honorton & Ferrari, 1989; Steinkamp, Milton, & Morris, 1998). Tart (1977, 2001; Tart, Palmer, & Redington, 1979) has also argued for the inclusion of immediate feedback in parapsychological experiments. His claim revolves around a learning paradigm in which some form of feedback on performance is almost always provided as reinforcement immediately after a behavioural response. The elimination or absence of feedback, in turn, is commonly used to extinguish a learned behavioural response to a given stimulus.

On the topic of feedback, the stimuli used in contingent tasks are also worthy of attention. In the Luke and colleagues' paradigm, participants in the negative reward condition had to take part in a boring

number-vigilance task. In their replication attempt, Hitchman et al. (2012) observed that, despite most participants indicating that this task was relatively unpleasant, some actually claimed to enjoy the task, and it was felt that the emotive intensity of this negative reward could be enhanced. To ease the transition to a trial-by-trial feedback mechanism that would necessitate a larger number of incidences of contingent outcomes, it was also important to reduce the length of time required for each instance. Consequently, unpleasant, scary, or gruesome images were selected from the IAPS set (Lang & Greenwald, 1993) for use in negative reward conditions. Moreover, Hitchman et al. (2012) speculated that the primary function of psi-mediated instrumental responses may be to avoid negative outcomes. It was therefore thought that the use of strongly negative, gruesome, or scary images provided on a trial-by-trial basis may lead to a stronger aversion to the negative reward conditions, and hence a greater tendency for participants to exhibit PMIR during each trial.

The overall aim of the present study was to refine what has so far been a promising experimental protocol whilst exploring the roles of intentionality and lability in precognitive performance. It was predicted that participants would score more hits than would be expected by chance on both versions of the psi task, although better performance was expected during the nonintentional version. Participants' lability scores were also expected to be related to their precognitive performance. Furthermore, consistent with the Hitchman et al. (2012) study, hypotheses regarding the relationship between participants' psi scores and their beliefs about psi and their own luck were tested. Finally, it was predicted that precognition scores would be positively related to participants' scores on Bem's (2003, 2011) emotional reactivity items.

Method

Design

A repeated-measures, quasiexperimental design was employed in which participants completed both a 10-trial nonintentional precognition task and a 10-trial intentional precognition task in sequential order. In both cases, the dependent variable was the number of direct hits they scored on the task, when the mean chance expectation (MCE) was five hits for each participant. A trial-by-trial reward manipulation was utilised such that each time participants scored a hit ($p = .5$), they were administered a positive reward of seeing a pleasant image whereas each time they scored a miss ($p = .5$) they were given a negative reward of seeing a negative image. For correlational purposes, questionnaire measures were used to collect individual difference data for the independent variables of paranormal belief (sheep-goat), a composite lability measure, and an emotional reactivity measure.

Participants

Twenty male and 30 female participants were recruited by opportunity sampling from friends, colleagues, associates, students at the University of Northampton, interested members of the public, and members of local hobby groups. Although two participants did not disclose their age, the mean age of the remaining participants was 27.21 years ($SD = 8.72$). Participants were invited to take part in "a psychological investigation of possible psychic ability and how it relates to an individual's personality and beliefs." No incentives were offered in exchange for participation and all participants were informed that they may see unpleasant images before giving their final consent to take part in the study.

Individual Difference Measures

Demographic questionnaire. This 2-item questionnaire asked about participants' age and gender.

Sheep-goat belief questionnaire. This 5-item questionnaire contained four questions corresponding to different aspects of the belief in psi variable as specified by Palmer (1972) in addition to a fifth item concerning whether or not participants believed their luck could influence the outcome of the psi task. Each item is scored on a true/false basis, yielding a total belief score that can range from 0 to 4.

Openness to experience (OE) scale. This 20-item questionnaire addresses an individual's openness to new experiences (Goldberg, 1999). Participants respond to items such as "Believe in the importance of art" and "Have a rich vocabulary" by indicating the extent to which each statement is an accurate description of themselves. Each item is rated on a 5-point Likert scale from "very inaccurate" to "very accurate," yielding a score which can range from 0 to 80. Coefficient alphas for subscales range from .77 to .86 (Goldberg, 1999), and these scores have been found to correlate with scores on the equivalent scale of the NEO personality inventory (r = .56; Gow, Whiteman, Pattie, & Deary, 2005).

Creative Cognition Inventory (CCI). This 29-item questionnaire addresses the use of different cognitive styles in the creative process, considering factors such as heightened internal awareness, intuition, and playfulness (Holt, 2002). Questions are categorised according to linear (4 items) and nonlinear (25 items) subscales. The linear scale relates to logical analysis, planning, and careful selection of ideas; it has acceptable internal consistency (α = .72). The nonlinear scale relates to paying attention to internal states, playful cognition, ideas arising in states along the dream-wake continuum, and a sense of ideas coming from "something other"; it also has acceptable internal consistency (α = .92). Both scales were found to have adequate construct, convergent, and discriminant validity (Holt, 2007). Respondents indicate the extent to which experiences such as "Trusting hunches or instincts" and "Paying attention to visual imagery" are important to their being creative. Items are scored on 5-point Likert scales from "not at all important" to "extremely important," yielding a total score that can range from 4 to 20 for the linear subscale and 25 to 125 for the nonlinear subscale.

Complex Partial Epileptic Signs (CPES) cluster of the Personal Philosophy Inventory (PPI). This 16-item questionnaire containing items which relate to experiences similar or analogous to those reported by patients with unusual activity in the temporal lobes, often achieved by means of direct electrical stimulation during surgery (Persinger & Makarec, 1987). Each item is scored on a yes/no basis, yielding a total score that can range from 0 to 16. Patients' responses to these items have been found to correlate significantly with measures of electroencephalographic activity localised to the temporal lobes (Makarec & Persinger, 1987). The CPES scale was found to have a satisfactory level of internal reliability and test-retest correlations of r = .85 to r = .95 after 10 days and r = .60 to r = .70 after 100 days (Persinger & Valliant, 1985).

Emotional Creativity Inventory (ECI). This 30-item scale contains items that address components of emotional creativity such as preparedness, novelty, effectiveness, and authenticity (Averill, 1999). Participants respond to statements such as "My emotional reactions are different and unique" on a scale from 1 ("strongly disagree") to 5 ("strongly agree"), yielding a total score that can range from 30 to 150. Only two of the items are reverse scored, but no evidence was found of response bias in test data (Averill, 1999). The inventory has been found to have a high level of internal reliability (α = .90) and test-retest reliability (r = .91 after 3 months).

Mood lability items. Three items derived from a self-report mood affect scale developed by Akiskal et al. (1995) based on the frequent up-and-down fluctuations in mood observed between sufferers' episodes of mood disorders such as bipolar II disorder (Inter-episode mood lability; Kraepelin, 1921). Participants indicate the extent to which each statement reflects patterns in their mood on a 3-point scale from "Not at all" to "Very much so," yielding a total score that can range from 3 to 9. These items have been found to be associated with clinical diagnoses of bipolar II disorder and family history of mood disorders (Benazzi, 2004; Benazzi & Akiskal, 2005).

Emotional reactivity items. These two items address individuals' awareness of their emotional reactivity to violent, scary, or gruesome content in photographs, movies, and videos (Bem, 2003, 2011). Participants respond on a scale from 1 ("not at all intensely aware") to 5 ("very intensely aware"). Bem advises the use of mean scores for correlational analysis, which can range from 1 to 5.

Materials for Test Session

PMIR visual basic program. The software program used in the Hitchman et al. (2012) study was modified for this experiment by the first author (available by request). The program's code incorporates

a number of procedural changes that enabled the experimental hypotheses to be addressed. As before, all questionnaires were integrated into the software, and the program was completely automated such that participants could give their responses and complete the experimental tasks without the presence, aid, or intervention of the experimenter.

International Affective Picture System (IAPS). The program presents a large set of emotive colour photographs from the IAPS set, the contents of which span numerous semantic categories including awe, excitement, contentment, amusement, fear, sadness, disgust, and anger (Lang & Greenwald, 1993). The images have been rated by independent judges for their perceived valence, arousal, and dominance. After all erotic images were filtered from the set on ethical grounds, the images were sorted into three groups: (a) positive images consisting of the most positively valenced pictures (mean pleasantness rating > 6), (b) negative images consisting of the most negatively valenced pictures (mean pleasantness rating < 4), and (c) neutral images consisting of those pictures with a mean pleasantness score around the midpoint of the rating scale (4.5 < mean pleasantness rating < 5.5). From each group, images were then paired together into sets such that each pair satisfied the following criteria: (a) both images depicted content which, in the authors' opinion, were drawn from similar semantic categories (e.g., a picture of rabbits and a picture of puppies both represent animal content), (b) the mean pleasantness and arousal ratings were very closely matched (for all pairs, pleasantness ratings were within .5 of a unit and arousal ratings were within 1.5 units), (c) the sum of the standard deviations of pleasantness and arousal ratings did not exceed 3.5 units (implying that the majority of individuals have similar emotional responses to the images). Using this method, 20 pairs of neutral images were selected to be used as target and decoy images for both nonintentional and intentional precognition trials. Similarly, 10 pairs of positive and 10 pairs of negative images were isolated for use as positive and negative reward images for the nonintentional precognition trials. Finally, 10 single positive and 10 single negative images were selected from the remaining unused images for use as positive and negative reward images in the intentional precognition trials.

Procedure

Participants were greeted and briefed in a quiet room, where it was explained to them that the experiment involved a test of psychic ability, but the implicit, precognitive nature of the nonintentional task was not disclosed. After being given the opportunity to pose any questions, participants were left alone in a quiet area to operate the computer program on a laptop computer. The principal investigator waited in a nearby room and was available to help if participants needed any further assistance.

The program displayed a written introduction before collecting participants' informed consent and presenting them with digitised versions of the questionnaire battery. Subsequent to participants answering all of the questions, the program instructed them to proceed to the experimental tasks. The program then gave instructions for what had been described to them in the briefing as a "preparatory" image preference indication task. Participants were told they would be shown some pictures in pairs and asked to choose their preferred image from each set. As a rationale for this task, participants were speciously told that their selections would help the program to choose appropriate targets for a later task.

Participants then took part in 10 nonintentional precognition trials. During each trial, they were shown one of the neutral pairs of images and selected which one of the two images they most preferred. An example neutral image pair is presented in Figure 1. Please note that for contractual reasons, indicative images are displayed rather than genuine IAPS pictures in all figures. At the beginning of each trial, the position of the cursor was reset to the centre of the screen to avoid biasing participants towards selecting either of the images. The trial in which each specific pair of images occurred was randomised for each participant, and the place each of the two images from each pair appeared on the screen (left or right) was also randomised for each trial.

Figure 1. Example stimuli for the precognition task.

Unbeknown to the participants, this image preference task constituted a forced-choice, nonintentional precognition task, as each time they indicated their preferred image from the pair the computer immediately selected one of the images at random as the target. Directly after each trial, participants were administered a positive or negative reward for their performance by means of a secondary image preference task. If the participants' selection matched the computer's random selection, they entered the positive reward condition, and the secondary image preference task consisted of participants indicating their preferred image from one of the pairs of positive images (Figure 2). However, if the participants' selection didn't match the computer's random selection, they entered the negative reward condition, and the secondary image preference task consisted of participants indicating their preferred image from one of the pairs of negative images (Figure 3). No data were collected regarding participants' selections during the contingent reward tasks.

Figure 2. Example stimuli for the positive reward condition.

Figure 3. Example stimuli for the negative reward condition.

As per the Hitchman et al. (2012) study, randomisation of the on-screen image array positions and computer target selections was achieved using the random number generation function within VB.NET, which is seeded by the CPU timer. Goodness of fit tests were conducted on the experimental data to evaluate whether or not adequate randomisation had been achieved. The tests revealed there was no bias in the number of times each target was selected by the computer, $\chi^2(1, N = 1000) = 0.68$, $p = .41$), nor a left/right bias in the positioning of stimulus images on screen, $\chi^2(1, N = 1000) = .06$, $p = .80$.

After participants had received their positive or negative reward for the 10th nonintentional trial, the computer displayed a dialogue box that informed them they had completed the task and asked them to call back the experimenter. The experimenter informed the participants of the implicit, precognitive nature of the task they had just completed and answered any questions to ensure that this had been thoroughly understood. Here, it was ensured that participants had been unaware of the true nature of the task they had just completed, as a means to validate that they had not intentionally attempted to exhibit precognition within trials. Participants were then told that the final part of the experiment involved a very similar task. However, now they were aware of how the program functioned, rather than simply selecting their favourite of the neutral images, participants were asked to use their will and intent to try to predict (by whichever means they wished) which image they thought the computer would select, thus achieving as many of the positive contingent rewards as possible. Once they had confirmed their understanding of the instructions, participants were provided with a password that was required for them to proceed to the final task on the computer.

After participants entered the password, the computer reiterated the instructions for the final task, prompting them to intentionally attempt to select the images which they believed would lead them to the positive outcomes. The program once again proceeded through 10 precognitive trials by displaying pairs of closely matched neutral images in random order. As it was no longer necessary to maintain any secrecy, in the reward conditions, rather than giving participants a choice between pairs of positive or negative images, a single positive or negative reward image was displayed on the screen for 3 s before automatically proceeding to the next trial.

Once participants had participated in and been positively or negatively rewarded for all 10 intentional precognition trials, the computer informed them that they had completed the experiment and prompted them to call back the experimenter. The experimenter then provided a full debrief. Typically, participants discussed their impressions of how well they thought they had performed in each version (nonintentional and intentional) of the task. Before leaving, all participants were asked not to discuss the nature of the experiment with other potential participants.

Ethics

The project was designed to adhere to the British Psychological Society's Code of Ethics and Conduct (BPS, 2009) and received ethical approval from the University of Northampton Research Ethics Committee. Participants were briefed prior to giving their informed consent as part of the program. All data were collected anonymously and participants were made aware of their right to withdraw from the experiment at any time without having to provide a reason. Importantly, participants were also forewarned in the briefing period that they may see negative, gruesome, and scary images during the experiment, and it was also mandatory that they ticked a box indicating that this had been explained to them and they were happy to continue to take part in the study to proceed past the digital consent form.

Results

The total number of hits participants scored on the nonintentional and intentional versions of the psi task was recorded, along with their scores on the individual difference measures. One participant did not provide answers to a number of items of the nonlinear subscale of the Creative Cognition Inventory and two participants did not provide sufficient data in relation to the Emotional Creativity Inventory. Due to the large number of omitted responses for these participants, substitution of scores was not considered feasible and consequently those participants' data have been excluded from analyses where appropriate.

The primary hypotheses predicted that participants would select more target images during the nonintentional and intentional versions of the precognition task than would be expected by chance alone. Fifty participants each completed 10 nonintentional and 10 intentional precognition trials with an associated probability of correctly selecting the target image of .50. Thus, with a total of 500 trials for each version of the task, the MCE was 250 hits. For the nonintentional precognition task, the actual number of hits was 247, with a mean hit rate per participant of 4.94 hits (SD = 1.17). As participants scored marginally fewer hits than the mean chance expectation (MCE = 5.00 hits), the experimental hypothesis was not supported. The representative effect size by t test was r = -.02 compared to r = .16 observed by Hitchman et al. (2012). In the intentional version of the task, participants accumulated a total of 245 hits, marginally lower than the MCE of 250. The mean hit rate per participant was 4.90 hits (SD = 1.42). As participants scored slightly fewer hits than the MCE, the experimental hypothesis was not supported. The representative effect size was r = -.03. On average, participants performed marginally better at the nonintentional version of the task (M = 4.94 hits, SD = 1.17 vs. M = 4.90 hits, SD = 1.42), but the difference in performance across the two conditions was not statistically significant, $t(49) = 0.16$, p = .44, one-tailed. Furthermore, there was only a very slight positive correlation between intentional and nonintentional precognition scores, $r(48) = .07$, p = .31, one-tailed.

The next set of hypotheses concerned relationships between psi task performance and individual difference measures related to lability. For transparency in interpreting the correlations between psi task scores and individual difference scores reported below, Table 1 presents a correlation matrix of these individual difference measures to highlight where there may be shared variance.

It was predicted that participants' performance on the (a) nonintentional and (b) intentional versions of the precognition task would be positively correlated with their scores on a composite measure of lability. Scores on the constituent elements of lability (Emotional Creativity, Mood Affect, Openness to Experience, Creative Cognition and Complex Partial Epileptic Signs) were amalgamated into a single score. Table 2 shows that composite lability scores were positively related to participants' performance on the nonintentional version of the precognition task, although the relationship was not statistically significant, $r(45) = .17$, p = .13, one-tailed. Conversely, contrary to the alternative hypothesis, composite lability scores were found to be negatively related to participants' performance on the intentional version of the precognition task, $r(45) = -.13$. A Steiger calculation (Clark-Carter, 2010) revealed that the difference between the two correlations is not significant, $t(44) = 1.39$, p = .09, one-tailed. Furthermore, none of the constituent elements of the composite lability scale were found to be significantly related to either intentional or nonintentional precognition scores.

Table 1
Pearson Correlations and Two-Tailed Significance Values Between Predictors of Psi Task Performance (N = 50)

	Sheep-Goat	ECI	Mood Affect	OE	CCI Linear	CCI Non-linear	PPI
ECI	.34*						
Mood Affect	.07	.04					
OE	.15	.49***	-.03				
CCI Linear	-.16	.05	-.02	-.10			
CCI Nonlinear	.37**	.63*****	.06	.53****	.01		
PPI	.51****	.38**	.05	.42**	-.24	.50***	
ER	.09	34*	.04	.26	-.09	.17	.04

Note. ECI: Emotional Creativity Inventory; OE: Openness to Experience; CCI: Creative Cognition Inventory; PPI: Personal Philosophy Inventory; ER: Emotional Reactivity
*$p < .05$. **$p < .01$. ***$p < .001$. ****$p < .0001$. *****$p < .00001$.

Table 2
Pearson Correlations Between Psi Task Scores and Predicted Covariates (Steiger One-Tailed Significance)

	N	Nonintentional	Intentional	Steiger's *t*
Sheep-goat	50	.07	-.06	0.68
		(.31)	(N/A)	(.25)
ECI	48	.06	-.13	0.91
		(.34)	(N/A)	(.18)
Mood Affect	50	.03	-.10	0.68
		(.42)	(N/A)	(.25)
OE	50	.10	-.12	1.10
		(.26)	(N/A)	(.14)
CCI Linear	50	.18	.22	0.24
		(.11)	(.06)	(.41)
CCI Nonlinear	49	.13	-.08	1.03
		(.18)	(N/A)	(.15)
PPI	50	.18	-.20	2.01*
		(.11)	(N/A)	(.03)
Composite Lability	47	.17	-.13	1.39
		(.13)	(N/A)	(.09)
Emotional Reactivity	50	.29*	-.04	1.68
		(.02)	(N/A)	(.05)

Note. ECI: Emotional Creativity Inventory; OE: Openness to Experience; CCI: Creative Cognition Inventory; PPI: Personal Philosophy Inventory
*$p < .05$.

It was also predicted that emotionally reactive participants would perform better at both versions of the precognition task. A mean score of the two emotional reactivity questions was calculated and correlated against participants' precognition scores. For nonintentional trials, there was a significant positive correlation between participants' precognitive performance and their mean emotional reactivity scores, $r(48) = .29$, $p = .02$, one-tailed. Contrary to the alternative hypothesis, for intentional trials there was a small negative correlation between the two variables, $r(48) = -.04$. A Steiger calculation revealed that the difference between the two correlations was on the threshold of statistical significance, $t(47) = 1.68$, $p = .05$, one-tailed.

Turning to the belief measures, consistent with the Hitchman et al. (2012) study, it was predicted that performance on the (a) nonintentional and (b) intentional versions of the precognition task would be positively correlated with participants' belief in psi (as measured by criteria I-IV of the Sheep-Goat questionnaire). Table 2 shows a very small positive correlation between belief in psi and performance on the nonintentional version of the precognition task and, unexpectedly, a very small negative correlation between belief in psi and performance on the intentional version of the precognition task, nonintentional: $r(48) = .07$, $p = .31$, one-tailed; intentional: $r(48) = -.06$.

Also based on the Hitchman et al. (2012) study, it was hypothesised that participants who believed they could use their psi to affect the outcome of the experiment would perform better at the (a) nonintentional and (b) intentional versions of the precognition task. Participants responded true or false to the statement "I believe that my psychic ability can affect the outcome of this experiment" and were grouped according to their response. The majority of participants ($n = 40$) indicated they were goats (i.e., their psychic ability couldn't influence the outcome of the experiment) compared with just 10 sheep. Sheep performed marginally better at the nonintentional version of the precognition task than goats (sheep mean precognition task score = 5.30, $SD = 0.82$; goat mean precognition task score = 4.85, $SD = 1.23$), but the difference was not significant, $t(48) = 1.09$, $p = .14$, one-tailed. Contrary to the alternative hypothesis, in the intentional version of the task, goats outperformed sheep (sheep mean precognition task score = 4.60, $SD = 1.35$; goat mean precognition task score = 4.98, $SD = 1.44$).

A similar hypothesis predicted that participants who believed that their luck could influence the outcome of the test would achieve a higher score on the (a) nonintentional and (b) intentional versions of the precognition task. For nonintentional trials, the 30 "luck sheep" (60%), who indicated that they believed their luck could affect the outcome of the experiment, scored nonsignificantly higher on the psi task than the 20 "luck goats" (40%), who indicated to the contrary (luck sheep mean psi score = 5.13, $SD = 1.31$; luck goat mean psi score = 4.65, $SD = .88$; $t(48) = 1.45$, $p = .08$, one-tailed. For intentional trials, luck sheep also outperformed luck goats (luck sheep mean psi score = 5.20, $SD = 1.32$; luck goat mean psi score = 4.45, $SD = 1.47$). The result of an independent samples t test indicated that luck sheep performed significantly better than the luck goats, $t(48) = 1.88$, $p = .03$, one-tailed.

As noted by Hitchman et al. (2012), the parametric Pearson correlation test may not be valid for scales with a score range of less than 20 points (Clark-Carter, 2010). This applies to the Sheep-Goat, Mood Affect, Personal Philosophy Inventory and Emotional Reactivity measures. Whilst Pearson correlations are reported above for consistency and comparison with other correlations, Spearman nonparametric correlations were also calculated. The results indicated the same pattern: Sheep-goat: nonintentional, $r_s(48) = .03$ $p = .42$, one-tailed; intentional, $r_s(48) = -.07$. For Mood Affect: nonintentional, $r_s(48) = .03$, $p = .42$, one-tailed; intentional, $r_s(48) = -.08$; Personal Philosophy Inventory: nonintentional, $r_s(48) = .16$, $p = .14$, one-tailed; intentional, $r_s(48) = -.22$; Emotional Reactivity: nonintentional, $r_s(48) = .36$; $p = .005$, one-tailed; intentional, $r_s(48) = -.10$.

Post Hoc Analysis of Participants' Image Preference Biases

The present study made use of closely matched pairs of authentic images as targets in the precognition task as opposed to the sets of four fractal images used in previous studies. It was therefore important to assess whether participants exhibited any systematic biases towards either image from each pair. Table 3 presents the number of times each image from each set was selected. The results of chi-square analyses

indicated that participants appeared to exhibit a systematic preference for image A in set 8, χ^2 (1, N = 50) = 3.92, p = .05 and image B in set 11 χ^2 (1, N = 50) = 5.12, p = .02. Hitchman et al. (2012) reported that participants showed similar biases within 2 out of the 15 sets of fractal target images in their study. Consequently, it would appear that the use of authentic images in place of fractal target images did not bring about a tangible increase in the extent of participants' preferences for specific images within each set.

Post Hoc Analysis of Practice/Decline Effects

Given that by necessity in this repeated measures design participants always took part in the nonintentional precognition task before completing the intentional precognition task, concerns were raised over the potential influence of practice and/or decline effects. In order to assess this, the total number of hits achieved by all participants for each trial was considered. Figure 4 indicates that there were no clear patterns in performance across either nonintentional trials (T1-T10) or intentional trials (T11-T20). In support of this, chi-square analysis indicated there were no significant differences between overall performance in each trial, χ^2 (19, N = 50) = 8.65, p = .98. Therefore, no evidence of either practice or decline effects was observed across the experiment.

Table 3
Chi Square Analyses of Participants' Preferences for a Specific Image From Each Target Pair

Image set	Image A selected	Image B selected	χ^2	p
1	31	19	2.88	.09
2	20	30	2.00	.16
3	29	21	1.28	.26
4	19	31	2.88	.09
5	22	28	0.72	.40
6	24	26	0.08	.78
7	26	24	0.08	.78
8	32	18	3.92	.05
9	30	20	2.00	.16
10	24	26	0.08	.78
11	17	33	5.12	.02
12	31	19	2.88	.09
13	25	25	0.00	1.00
14	29	21	1.28	.26
15	30	20	2.00	.16
16	19	31	2.88	.09
17	21	29	1.28	.26
18	30	20	2.00	.16
19	31	19	2.88	.09
20	26	24	0.08	.78

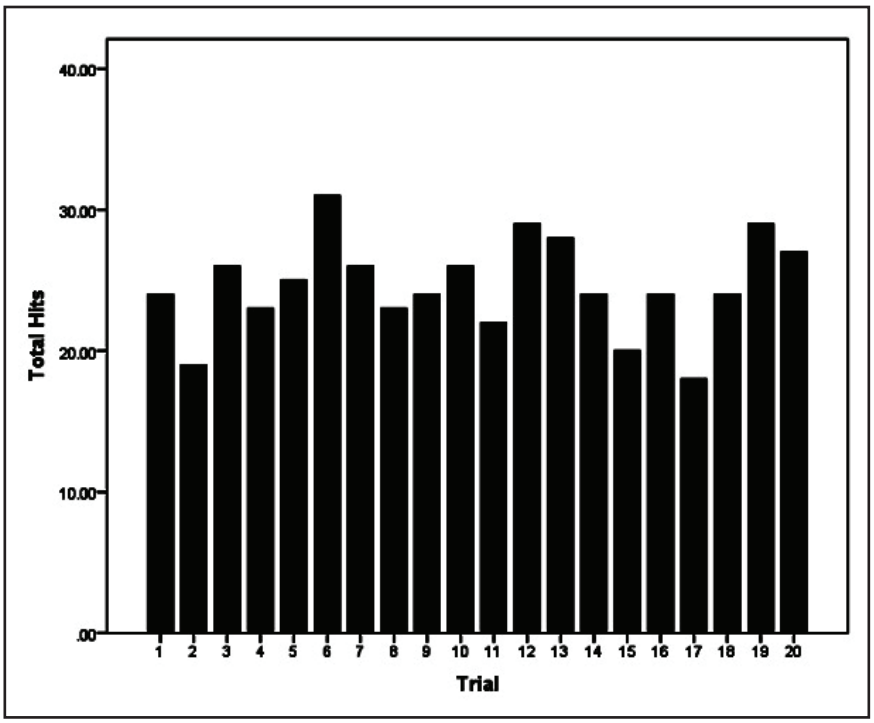


Figure 4: Total hits achieved by all participants for each trial.

Consistent with the Hitchman et al. (2012) study, multiple statistical tests have been conducted without a correction applied to the alpha levels for multiple analyses. Milton and Wiseman (1997) have noted that the standard Bonferroni adjustment should be considered conservative, whilst Abdi (2007) claims the Bonferroni correction is not appropriate when the inferential tests conducted are not entirely independent. Readers are advised that due to the exploratory nature of this study, all significant results reported within this manuscript are claimed as tentative pending replication, especially as the chance of a type 1 error is increased as a consequence of the multiple analyses carried out.

Discussion

The main experimental hypotheses of this study concerned a comparison of performance at equivalent nonintentional and intentional precognition tasks and the relationship between precognition and lability scores. Within Stanford's specification of the PMIR model, it is claimed that knowledge of need-relevant circumstances and an intention to fulfil such needs may play an inhibitory role in the psi-mediated instrumental response process. It was therefore expected that performance on the nonintentional precognition task would be higher than on the intentional precognition task. Overall, participants selected slightly fewer of the target images than the MCE in both nonintentional and intentional versions of the precognition task, providing no evidence of either intentional or nonintentional psi. Furthermore, although participants performed slightly better in the nonintentional version of the task as hypothesised, there was not a statistically significant difference between scores. Meanwhile, participants' lability scores were found to be unrelated to both their intentional and nonintentional precognitive performance.

Although the results of the Hitchman et al. (2012) study in relation to the main psi effect were in the predicted direction, they were nevertheless nonsignificant. The findings of the present study therefore represent a further decline in tacit precognition scores to below-chance levels and a second failure to replicate the significant effects demonstrated by Luke and associates using a largely similar method. Colborn (2004) has reviewed a multitude of factors that may account for general patterns of declining results across parapsychological paradigms. Amongst these, there are several potential explanations that may account for

these dissimilarities in results obtained by different researchers, including false positives (type I errors), decline effects, design modifications, and experimenter effects. In the case of the former, it may simply be a chance occurrence that Luke and associates were able to achieve psi indicative results that are nevertheless spurious and not reflective of a genuine and robust effect. However, this interpretation is questionable given that the results were replicated across a series of four studies, and are consistent with a much larger database of similar psi-indicative research in the forced-choice psi paradigm (e.g., Honorton & Ferrari, 1989).

With respect to design modifications, one key area in which this study differed from previous studies was in the implementation of trial-by-trial feedback. Data from several meta-analyses of forced-choice intentional psi studies indicate that the immediacy of feedback in relation to the psi task is a key variable in determining the size of effects (Honorton & Ferrari, 1989; Steinkamp, 2005; Steinkamp et al., 1998). The use of trial-by-trial feedback also helped to overcome a potential issue that the fundamental need being fulfilled in the psi task may be the avoidance of punishment, rather than the seeking of the highest reward conditions. With trial-by-trial feedback, it is necessary for participants to score a hit in each and every trial to avoid being negatively rewarded, whereas the feedback system in previous studies only required participants to score above chance across the entire run of trials to escape punishment. Despite the conceptual advantages of this feedback mechanism, participants, on average, failed to outperform MCE in the present study, whereas above-MCE results have been reported in each of the studies of this type employing end-of-run feedback. It may be, then, that in the context of this type of experiment, a series of smaller rewards carry less weight than a single reward of a longer duration.

This study also differed in terms of the task used in the negative reward condition. In the previous study, participants who underperformed the mean chance expectation over 15 trials took part in a boring number vigilance task. In the present study, however, participants were shown a negative image from the IAPS picture set which contained violent, gruesome, or scary content. This was intended to enhance the emotive impact of the negative outcome and hence increase participants' aversion to the negative reward condition. This rationale was supported by the finding that a significant, positive relationship was observed between participants' performance on nonintentional trials and their mean scores on Bem's (2003, 2011) emotional reactivity items.

However, the use of trial-by-trial feedback ensured that it was very unlikely ($p < .001$) that participants could entirely avoid negative rewards, whereas in previous end-of-run scenarios, there was a 50% chance of avoiding the negative contingent task. It is therefore possible that the presence of unpleasant stimuli throughout the experiment potentially induced a general state of anxiety or a psi equivalent of learned helplessness in some participants, which may have manifested itself within trials. As a result, rather than holding a consistent state of openness, being sensitive to extrasensory stimuli and responding to them accordingly, participants may instead have maintained feelings of apprehension and defensiveness. It is also a possibility that certain positive or negative reward images could have primed participants' selections of any subsequent neutral target images. Furthermore, a number of participants indicated a level of curiosity towards the negative images, suggesting that some may have been more attracted to the negative reward condition than the positive reward condition. In hindsight then, the use of emotive images as feedback may be better suited to an end-of-run feedback system to avoid building and reinforcing a state of anxiety rather than openness and potentially priming subsequent decision making.

With respect to the target images themselves, on the basis of participant and reviewer feedback in response to the Hitchman et al. (2012) study, sets of four fractal patterns were replaced by pairs of authentic images in the present study. This was intended to enhance the ecological validity of the precognition task. However, in a similar manner to the Hitchman et al. (2012) study, it was found that participants exhibited a systematic bias towards one of the images in two of the target sets. Although in the case of both studies target images were selected on the basis of being closely matched in terms of their content and valence and arousal ratings, in any set of images which differ tangibly in appearance, there remains an opportunity for a range of preferential biases (e.g. colour, gender) to influence participants' selections. Although it could come at the cost of psi discriminability, it is recommended that future studies attempt to further increase the normalisation of image sets, such as by using pairs of mirrored images, as in Bem (2011).

On the topic of ecological validity, it is worthwhile to consider a further design element at this stage. In the present study, participants took part in a total of 20 equivalent trials. Much of the early PMIR work (e.g. Stanford & Thompson, 1973) typically involved a single opportunity for participants to use psi to achieve a need-relevant goal, rather than a series of repetitious trials. Although many of the case reports from which Stanford developed the PMIR model appeared to involve only one opportunity for the exhibition of psi to achieve a favourable outcome (Stanford, 1974), repetition per se is not necessarily atypical. For example, lots of small delays in a bookshop could just as easily lead to a serendipitous meeting as a single delay of a longer duration. Nevertheless, it is recommended that future studies pay very careful attention to the ecological validity of psi tasks to ensure that they reflect the ways in which psi is assumed to occur in everyday life situations.

Turning to the role of intentionality in this paradigm, this study failed to yield evidence of the potentially inhibitory influence of need-relevant information or any conscious cognitions in relation to such information in psi-mediated scenarios. According to Stanford (1990), all that is necessary for an adaptive psi-mediated outcome to occur is a behaviour, and any consciously generated thinking, cognitive constraints or need-relevant information in relation to such a behaviour could inhibit the potential for it to be mediated by psi. In the present study, participants were foretold in their briefing that a psi task (the intentional psi task) would follow what was described as a "preparatory" image preference task (the nonintentional psi task). Participants performed similarly when offered a minimal amount of informational cognitive priming (as in the nonintentional task) and when given full disclosure regarding the nature of the task (as in the intentional task). Although no measure of cognitive activity was employed to objectively assess whether or not there were differences in the conscious cognitions of participants in the nonintentional and intentional conditions, it is unlikely that they would have felt the need to doubt or disobey the instructions they were provided. Consequently, it is reasonable to assume that in the nonintentional condition participants were simply indicating their image preferences, whereas in the intentional condition they were actively engaged in trying to foretell which image would be randomly selected by the computer. According to the authors' interpretation, the cognitive constraints which could inhibit psi were defined rather broadly in the specification of the PMIR model, so it is difficult to assess whether or not the different versions of the task would have resulted in participants engaging in different cognitive activities to an extent which Stanford believed may have a tangible impact on the psi process. Furthermore, in the absence of objective empirical data in relation to participants' cognitive processes during the tasks, it is not possible to discount the possibility that the limited information participants were given about the eventual intentional psi task may have led them to experience conscious cognitions that influenced either the speed of their decisions or the decisions themselves within the nonintentional psi task. Whilst greater efforts could be made to further reduce the cognitive priming given to participants in relation to the implicit psi task in future studies, we must still face the issue that if we accept the psi hypothesis as valid, it would not be possible to entirely avoid cognitive priming, as information in relation to tacit tasks could be available by extrasensory means.

It is also important to give mention to a design compromise in the present study. Given resource and time constraints, a repeated measures design was employed in order to meet sample requirements. Ideally, conditions in repeated measures designs should be counterbalanced in order to compensate for the potential for range effects such as practice, sensitisation and carryover effects from confounding results (Clark-Carter, 2010; Greenwald, 1976; Poulton, 1973). However, given that the nonintentional version of the precognition task relied on the naivety of the participants, it was entirely necessary for participants to take part in the nonintentional version of the task prior to receiving the briefing for the intentional version of the task. As a result, it is possible that some participants may have exhibited a greater level of performance at the intentional version of the task owing to them having practised an equivalent task in the previous condition. Conversely, some participants may have exhibited a diminished level of performance in the intentional trials owing to a decline effect, boredom, desensitisation to the reward stimuli or any other performance-based carry over effects associated with having performed a similar task in the previous condition. Due to this task order confound, any conclusions based on these results cannot be relied upon with confidence. Although chi square analysis found no evidence of consistent improvements or declines in performance across the experiment, authors typically recommended against the use of within-subjects designs when the juxtaposition

of conditions in not the main factor of interest (Greenwald, 1976; Poulton, 1973). The potential for these effects to manifest themselves within the data could only be eliminated by randomly allocating participants to take part exclusively in one of the two conditions in a between-subjects design.

The secondary hypotheses in this study concerned the performance of individuals at the nonintentional and intentional versions of the precognition task in relation to individual difference measures. Stanford (1990) had proposed rigidity in thought and behaviour as one of the principal inhibitory factors in the PMIR model. However, no support was found for the hypothesised effect of lability, with participants' scores on the composite measure not being found to correlate significantly with their performance at nonintentional or intentional precognition trials. Moreover, none of the constituent elements of the lability scale were found to covary to a significant extent with precognition scores. Focusing purely on effect sizes rather than statistical significance does not provide much more encouragement: all correlations were below $r = .2$ in absolute size. Similarly, no relationships were found between precognition task performance and participants' paranormal beliefs.

However, a significant, positive relationship was observed between participants' precognitive performance and their mean scores on Bem's (2011) Emotional Reactivity items, but only for nonintentional trials. Given the transition to using more emotionally potent images in the negative reward condition in this study, it was deemed particularly pertinent to have a gauge of whether individuals who were more reactive to negative emotive content would be more aversive to these images, and hence avoid the negative reward condition more frequently. As this relationship was only found to be significant for the nonintentional condition, the condition more similar to Bem's (2011) precognitive habituation task in which a similar effect has been observed, it was interesting to note that a Steiger calculation revealed the difference between the correlations for intentional and nonintentional trials approached a statistically significant level. This may suggest that participants' emotional reactivity does not interact as strongly with their tacit psi performance when they are consciously aware of the need to use some form of precognition.

As was noted by Hitchman et al. (2012), attempting to identify predictors of performance in a psi task in an experiment where no overall psi effect has been observed is not straightforward, as it is unclear whether psi phenomena were entirely absent from the experimental scenario, or simply if the majority of participants failed to demonstrate this ability. Indeed, Palmer (2009) has echoed these concerns by bemoaning the unreliability of psi performance and the effect of this on attempts to assess its covariates. Unless a relatively consistent psi effect can be identified, attempting to assess the roles of the awareness of the need to use psi, the intent to use psi and individual difference correlates of psi will remain problematic to achieve. What's more, a further concern with assessing trait-based measures as predictors of experimental variables is that that having a particular trait does not necessarily mean that the trait will be expressed under all circumstances. Consequently, future studies may wish to attempt to validate that any measure of lability used for a similar purpose is predictive of participants' propensity to exhibit a relevant labile state within the context of the experimental task, as well as ensure that the task is of such a nature that participants in that labile state are more likely to achieve a successful outcome.

To confound this issue, it is important to note that the reliability and validity of psychometric measures are typically assessed in isolation (i.e., when not administered amidst multiple other tests). However, Council's (1993) paper on context effects (CEs) highlights that correlations between psychological tests can vary considerably depending on whether they have been administered in the same testing session. The distortion of outcomes due to CEs has not been studied extensively, but it is important to keep in mind that the use of multiple measures within the same session can have unknown consequences and potentially threaten the construct validity of tests.

In addition, administering person-based measures before a test of psi can potentially result in a range of reactivity-related issues including, but not limited to, demand characteristics, priming, and rumination. Future researchers may seek to keep the measurement of predictor variables more distant from the experimental situation, perhaps in a separate controlled testing session. Furthermore, it is worthwhile to note that although participants received standardised instructions before beginning the experimental tasks, there were some minor differences in the information participants were provided with prior to attending their experimental sessions. Most importantly, for some participants who took part following the recom-

mendations of their associates, it was not possible to inform them before their arrival that the study involved the likelihood of seeing unpleasant images. Consequently, it is possible that their decision to consent to take part in the experiment may have been influenced by the fact that they had already made a considerable effort to attend the testing session prior to being fully informed. As well as being an ethical concern which should be avoided in future, it is also possible that this and other subtle differences in the information participants were provided prior to attending the testing session may have affected their orientation to the experimental tasks and their subsequent performance.

Nevertheless, for the second time, a significant difference was found between the psi task performance of luck sheep, those who believed they could use their luck to influence the outcome of the experiment, and luck goats, those who didn't believe they could do so. This finding, then, further strengthens the notion that a person's expectation in their ability to use luck in a particular situation may play a greater role in their success as opposed to the specific ways in which they conceive of luck (see Hitchman et al., 2012). It should be considered worthwhile to include the luck sheep-goat variable in subsequent studies to assess whether it is able to withstand the test of time and prove to be a robust and reliable effect.

Overall, this study has advanced the nonintentional precognition paradigm in several ways and, in turn, has raised a number of additional questions. Firstly, this experiment addressed the difference between nonintentional and intentional psi tasks, particularly in relation to the cognitive activities of the participants engaging in the tasks. No difference was found between performance at either version of the precognition task, raising doubt over the PMIR model's assertion that cognitive priming and focused intent can diminish the potential for psi-mediated instrumental responses to be executed. However, given the potential confound of task order and the fact that no overall evidence of psi was found within either version of the task, such doubts should only be cast cautiously. Concurrently, the first signs of a general pattern of declining results in this paradigm were observed. A number of potential explanations for such findings have been considered and should be continually monitored as the paradigm develops. Looking ahead, experimenter effects may be one of the most interesting avenues to explore in future studies in an effort to account for differential results across contrasting experimental teams. Nevertheless, a meta-analysis considering the six Luke and colleagues and Hitchman and colleagues studies conducted using this paradigm to date suggests that the overall paradigm still presents significant evidence of tacit psi, with a Stouffer Z of 3.75, $p = .00008$, mean effect size $r = .19$.

This study has also contributed to the consideration of the factors that may aid or hinder the instigation of psi-mediated instrumental responses, with a particular focus on the role of lability. Overall, little evidence was found that lability or any of its constituent elements had a bearing on participants' precognitive performance. Conclusions in relation to the effect of individual difference covariates of psi are clearly restricted in a study devoid of any evidence of psi per se. Nevertheless, effects of luck beliefs and emotional reactivity were observed, which should be considered as worthwhile variables to include in studies henceforth. Going forward, researchers may wish to turn their attention towards developing more reliable performance-based measures of the other individual difference covariates that are predicted to influence the PMIR process, particularly latent inhibition. In doing so, every effort should be made to tailor trait-based measures to the context in which they are expected to be expressed and to minimise the potential for measurements to impact the assessment of other experimental variables.

References

Abdi, H. (2007). Bonferroni and Sidak corrections for multiple comparisons. In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp.103–107). Thousand Oaks, CA: Sage.

Akiskal, H., Maser, J., Zeller, P., Endicott, J., Corvell, W., & Keller, M. (1995). Switching from "unipolar" to bipolar II: An 11-year prospective study of clinical and temperamental predictors in 559 patients. *Archives of General Psychiatry, 52*, 114–123.

Averill, J. R. (1999). Individual differences in emotional creativity: Structure and correlates. *Journal of Personality, 67*, 331–371.

Bem, D. J. (2003). Precognitive habituation: Replicable evidence for a process of anomalous cognition. *Proceedings of Presented Papers: The Parapsychological Association 46th Annual Convention*, 6–20.

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*, 407–425.

Bem, D. J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer, *Psychological Bulletin, 115*, 4–18.

Benazzi, F. (2004). Validating Angst's ups & downs personality trait as a new marker of bipolar II disorder. *European Archives of Psychiatry and Clinical Neuroscience, 254*, 48–54.

Benazzi, F., & Akiskal, H. S. (2005). A downscaled practical measure of mood lability as a screening tool for bipolar II. *Journal of Affective Disorders, 84*, 225–232.

BPS (2009). *Code of ethics and conduct: Guidance published by the Ethics Committee of the British Psychological Society*. Leicester, England: British Psychological Society.

Braud, W. G. (1980). Lability and inertia in conformance behavior. *Journal of the American Society for Psychical Research, 74*, 297–318.

Braud, W. G. (1981). Lability and inertia in psychic functioning. In B. Shapin and L. Coly (Eds.), *Concepts and theories of parapsychology* (pp. 1–36). New York, NY: Parapsychology Foundation.

Braud, W. (2002). Psi-favorable conditions. In V. G. Rammohan (Ed.), *New frontiers of human science: A Festschrift for K. Ramakrishna Rao* (pp. 95–118), Jefferson, NC: McFarland.

Braud, W. G., & Schlitz, M. J. (1983). Psychokinetic influence on electrodermal activity. *Journal of Parapsychology, 47*, 95–119.

Braud, W. G., Shafer, D., & Mulgrew, J. (1983). Psi functioning and assessed cognitive lability. *Journal of the American Society for Psychical Research, 77*, 193–208.

Clark-Carter, D. (2010). *Quantitative psychological research: The complete student's companion*. New York, NY: Psychology Press.

Colborn, M. (2004). The decline effect in spontaneous and experimental psychical research. *Journal of the Society for Psychical Research, 71*, 1–21.

Costa, P., & McCrae, R. (1992). *Professional manual: Revised NEO Personality Inventory (NEO PI--R) and NEO Five-factor Inventory (FFI)*. Odessa, FL: Psychological Assessment Resources.

Council, J. R. (1993). Context effects in personality research. *Current Directions in Psychological Science, 2*, 31–34.

Drennan, S., Roe, C., & Broughton, R. (2011). Lability, paranormal beliefs and psychokinetic experiences: Development of the Lability Scale using an online survey-based study. *Abstracts of Presented Papers: The Parapsychological Association 54th Annual Convention*, 25–26.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg, Netherlands: Tilburg University Press.

Gow, A. J., Whiteman, M. C., Pattie, A., & Deary I. J. (2005). Goldberg's IPIP Big-Five factor markers: Internal consistency and concurrent validation in Scotland. *Personality and Individual Differences, 39*, 317–329.

Greenwald, A. G. (1976). Within-subjects designs: To use or not to use? *Psychological Bulletin, 83*, 314–320.

Hitchman, G. A., Roe, C. A., & Sherwood, S. J. (2012). A re-examination of nonintentional precognition with openness to experience, creativity, psi beliefs and luck beliefs as predictors of success. *Journal of Parapsychology, 76*, 109–145.

Holt, N. J. (2002). *Creativity and boundary permeability*. Unpublished manuscript, University of Northampton, UK.

Holt, N. (2007). *Creativity, states of consciousness and anomalous cognition: The role of epistemological flexibility in the creative process*. Unpublished doctoral thesis, University of Northampton, England.

Holt, N. J., & Roe, C. A. (2006). The sender as a PK agent in ESP studies: The effects of agent and target system lability upon performance at a novel PK task. *Journal of Parapsychology, 70*, 69–90.

Holt, N., Simmonds-Moore, C., & Moore, S. (2007). Psi and cognitive disinhibition: Exploring the filters of consciousness hypothesis. *Proceedings of Presented Papers: The Parapsychological Association 50th Annual Convention*, 192–194.

Honorton, C., & Ferrari, D. C. (1989). "Future telling": A meta-analysis of forced-choice precognition experiments, 1935–1987. *Journal of Parapsychology, 53*, 281–308.

Kraepelin E. (1921). *Manic-depressive insanity and paranoia*. Edinburgh, Scotland: E & S Livingstone.

Lang, P. J., & Greenwald, M. K. (1993). *International affective picture system standardisation procedure and results for affective judgement: Technical reports 1A-1C*: University of Florida, Centre for Research in Psychophysiology, Gainesville, FL.

Lubow, R. (1989). *Latent Inhibition and conditioned attention theory*. Cambridge, England: Cambridge University Press.

Luke, D. P., Delanoy, D., & Sherwood. S. J. (2008). Psi may look like luck: Perceived luckiness and beliefs about luck in relation to precognition. *Journal of the Society for Psychical Research ,72*, 193–207.

Luke, D. P., & Morin, S. (2009, September). *Luck beliefs, PMIR, psi and the sheep-goat effect: A replication*. Paper presented at the meeting of the Society for Psychical Research, Nottingham, England.

Luke, D. P., Roe, C., & Davison, J. (2008). Testing for forced-choice precognition using a hidden task: Two replications. *Journal of Parapsychology, 72*, 133–154.

Makarec, K., & Persinger, M. A. (1987). Electroencephalographic correlates of temporal lobe signs and imaginings. *Perceptual and Motor Skills, 64*, 1124–1126.

Milton, J., & Wiseman, R. (1997). *Guidelines for extrasensory perception research*. Hatfield, England: University of Hertfordshire Press.

Palmer, J. (1972). Scoring in ESP tests as a function of ESP. Part 2: Beyond the sheep-goat effect. *Journal of the American Society for Psychical Research, 66*, 1–26.

Palmer, J. (2009). Winning over the scientific mainstream, *Journal of Parapsychology, 73*, 3–8.

Persinger, M. A. (1989). Geophysical variables and behavior: Predicting the details of visitor experiences and the personality of experients: The temporal lobe factor. *Perceptual and Motor Skills, 68*, 55–65.

Persinger, M. A., & Makarec, K. (1987). Temporal lobe epileptic signs and correlative behaviors displayed by normal populations. *Journal of General Psychology, 114*, 179–195.

Persinger, M. A., & Valliant, P. M. (1985). Temporal lobe signs and reports of subjective paranormal experiences in a normal population: A replication. *Perceptual and Motor Skills, 60*, 903–909.

Poulton, E. C. (1973). Unwanted range effects from using within-subject experimental designs. *Psychological Bulletin, 80*, 113–121.

Rao, K. R., & Davis, J. W. (1978). The differential effect and experimenter effects in intentional and nonintentional psi tests. *Journal of Parapsychology, 42*, 1–19.

Roe, C. A., & Holt, N. J. (2006). The effects of strategy ("willing" versus absorption) and feedback (immediate versus delayed) on PK performance. *Journal of Parapsychology, 70*, 49–67.

Stanford, R. G. (1974). An experimentally testable model for spontaneous psi events I: Extrasensory events. *Journal of the American Society for Psychical Research, 68*, 34–57.

Stanford, R. G. (1977). Conceptual frameworks of contemporary psi research. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp. 823–858). London, England: Van Nostrand Reinhold.

Stanford, R. G. (1982). An experimentally testable model for spontaneous extrasensory events. In I. Grattan-Guiness (Ed.), *Psychical research: A guide to its history, principles and practices* (pp. 195–205). Wellingborough, UK: Aquarian Press.

Stanford, R. G. (1990). An experimentally testable model for spontaneous psi events: A review of related evidence and concepts from parapsychology and other sciences. In S. Krippner (Ed.), *Advances in parapsychological research 6* (pp. 54–161). Jefferson, NC: McFarland.

Stanford. R. G., & Thompson, G. (1973). Unconscious psi-mediated instrumental response and its relation to conscious ESP performance. *Proceedings of Presented Papers: The Parapsychological Association 16th Annual Convention*, 99–103.

Steinkamp, F. (2005). Forced-choice ESP experiments: Their past and their future. In M. A. Thalbourne & L. Storm (Eds.), *Parapsychology in the 21st century: Essays on the future of psychical research* (pp. 124–163). Jefferson, NC: McFarland.

Steinkamp, E., Milton, J., & Morris, R. L. (1998). A meta-analysis of forced choice experiments comparing clairvoyance and precognition. *Journal of Parapsychology, 62*, 193–218.

Tart, C. (1977). Toward conscious control of psi through immediate feedback training: Some considerations of internal processes. *Journal of the American Society for Psychical Research, 71*, 375–407.

Tart, C. (2001). *Psi: Scientific studies of the psychic realm*. Bloomington, IN: iUniverse.com.

Tart, C. T., Palmer, J., & Redington, D. J. (1979). Effects of immediate feedback on ESP performance: A second study. *Journal of the American Society for Psychical Research, 73*, 151–165.

Acknowledgements

This research was funded by the Bial Foundation grant 105/08. We would like to gratefully acknowledge this support. We should also like to thank the Leslie Church and T. D. Lewis trustees for supporting this project.

Academic Department
UCFB Wembley, Wembley Stadium
Wembley, Middlesex, HA9 0WS, UK
G.Hitchman@ucfb.com

Abstracts in Other Languages

French

LA RELATION ENTRE LA LABILITE ET LA PERFORMANCE A DES VERSIONS INTENTIONNELLES ET NON-INTENTIONNELLES D'UNE TACHE PSI IMPLICITE DE TYPE PMIR

RESUME : Nombre de théories du psi, à l'instar du modèle de la réaction instrumentale médiatisée par le psi (PMIR) de Stanford, suggèrent que le psi peut fonctionner sans qu'une personne n'y prête attention, et que l'intention d'exhiber du psi peut être contre-productive. Toutefois, peu d'études parapsychologiques ont directement comparé les performances des participants à des versions intentionnelles et non-intentionnelles des tâches équivalentes. Cette étude cherche à résoudre cette question tout en explorant le rôle de la labilité, dont Stanford a suggéré qu'elle serait prédictive de la propension d'une personne à réagir à des stimuli extra-sensorielles. 50 participants prirent part à des versions intentionnelles et non-intentionnelles d'une tâche de précognition à choix forcée, binaire, à 10 essais. Un système de tâche contingente au résultat impliquant des images positives comme récompenses pour les essais réussis, et des images négatives comme punitions pour les essais manqués, fut administré pour chaque essai. Les participants eurent des scores légèrement plus faibles que ceux attendus par le hasard seul dans les deux versions de la tâche, avec aucune différence tangible dans leurs performances entre les tâches. De plus, aucune relation ne fut trouvée entre le nombre de succès précognitifs obtenus et leurs scores sur une mesure psychométrique composite de la labilité, ni sur ses éléments constituants. Cependant, les attentes des participants selon lesquelles leur chance pouvait contribuer à leur performance, ainsi que leur réactivité émotionnelle, furent positivement associées à leurs scores psi tacites.

German

DIE BEZIEHUNG ZWISCHEN LABILITÄT UND LEISTUNG BEI ABSICHTLICHEN UND NICHTABSICHTLICHEN VERSIONEN EINER IMPLIZITEN PSI-AUFGABE IM PVIR-MODELL

ZUSAMMENFASSUNG: Eine Anzahl von Psi-Theorien, wie Stanfords Modell der psi-vermittelten instrumentellen Reaktion (PVIR), legt nahe, dass Psi ohne die bewusste Aufmerksamkeit einer Person funktionieren kann und dass deren Absicht, Psi zu produzieren, dies eher verhindert. Einige parapsychologische Studien haben jedoch die Trefferleistung der Teilnehmer bei absichtlichen und nichtabsichtlichen Versionen entsprechender Aufgaben direkt verglichen. Diese Studie war diesem Thema gewidmet, wobei sie die Rolle der Labilität untersuchte, die Stanford zufolge die Tendenz einer Person vorhersagte, auf außersinnliche Reize zu reagieren. 50 Teilnehmer nahmen sowohl an absichtlichen wie unabsichtlichen Versionen einer Präkognitionsaufgabe teil, die aus 10 binären Durchgängen mit begrenzter Wahl bestand. Ein zufälliges ergebnisabhängiges Aufgabensystem, das positive Bilder als Belohnung bei Einzeldurchgängen mit Treffern und negative Bilder bei solchen mit Nieten vorsah, wurde auf der Basis der einzelnen Durchgänge verwendet. Die Teilnehmer blieben mit ihren erzielten Treffern bei beiden Aufgabenversionen knapp unter der mittleren Zufallserwartung, wobei sich kein Unterschied in ihrer Trefferleistung zwischen den Aufgaben nachweisen ließ. Zudem zeigte sich kein Zusammenhang zwischen der Anzahl der jeweils erzielten präkognitiven Treffer und ihrem Abschneiden bei einem zusammengesetzten psychometrischen Maß für Labilität, auch nicht bei den einzelnen Bestandteilen. Die Erwartungen der Teilnehmer, dass ihnen ihr Glück bei ihrer Leistung beistehen könnte, wie auch ihre emotionale Reaktivität, bezogen sich positiv auf ihre unbewussten Psi-Leistungen.

Spanish

LA RELACIÓN ENTRE LABILIDAD Y RENDIMIENTO EN VERSIONES INTENCIONAL Y NO INTENCIONAL DE UNA TAREA PSI IMPLICITA TIPO PMIR

RESUMEN: Una serie de teorías de psi tales como el modelo de Stanford de respuesta psi mediada instrumentalmente (PMIR) proponen que psi puede funcionar sin el conocimiento de la persona, y que la intención de exhibir psi puede ser contraproducente. Sin embargo, pocos estudios parapsicológicos han comparado directamente el desempeño de los participantes en versiones intencionales y no intencionales de tareas equivalentes. Este estudio trató de abordar esta cuestión y explorar el papel de la labilidad, que Stanford piensa que predice la propensión de una persona a responder a los estímulos extrasensoriales. Cincuenta personas participaron en ambas versiones intencionales y no intencionales de una tarea de elección forzada binaria precogntiva. Administramos un sistema de tareas con resultados contingentes de imágenes positivas como recompensa para los aciertos e imágenes negativas como castigo para los falllos, prueba por prueba. Los participantes tuvieron marginalmente menos aciertos que lo que se esperaría al azar en las dos versiones de la tarea, sin diferencia tangible en su desempeño entre las tareas. Por otra parte, no se encontró relación entre el número de aciertos de precognición y sus puntuaciones en una medida psicométrica compuesta de labilidad, ni de sus elementos constitutivos. Sin embargo, las expectativas de los participantes de que su suerte podía ayudar a su desempeño, así como su reactividad emocional, correlacionaron positivamente con sus puntuaciones psi tácitas.

Journal of Parapsychology, 79(1), 87–104

SUGGESTION, BELIEF IN THE PARANORMAL, PRONENESS TO REALITY TESTING DEFICITS, AND PERCEPTION OF AN ALLEGEDLY HAUNTED BUILDING

By Neil Dagnall, Kenneth Drinkwater, Andrew Denovan, and Andrew Parker

ABSTRACT: The present study investigated whether suggestion, level of belief in the paranormal, and proneness to reality testing deficits influenced participants' expectation of haunt-related phenomena. Participants watched a short slideshow outlining the history of a fictitious, abandoned hospital. Suggestion occurred in the final sentence of the presentation narration and stated that the hospital administrative building had either a history of ghostly activity or structural problems. Following the slideshow, to ensure participants attended to the suggestion, they read a transcript of the presentation narration. The experimenter then informed participants that they would see the internal features of the administrative building via a soundless, black and white video tour. On conclusion of the filmed sequence, participants completed measures assessing environmental perceptions and phenomena, haunt-related opinions and feelings, belief in the paranormal (Revised Paranormal Belief Scale), and proneness to reality testing deficits (Inventory of Personality Organization). Within the experimental phase, only level of paranormal belief and proneness to reality testing deficits affected haunt-related ratings; suggestion had no effect. Second phase inquiry, using path analysis, revealed that haunting history (the extent to which participants believed the administrative building had a history of being haunted) mediated the relationship between paranormal belief and expectation of haunt-related phenomena.

Keywords: suggestion, paranormal belief, reality testing, haunting

Definitions of ghosts vary over time and across cultures (Houran & Lange, 2001). The term "ghost" refers traditionally to the notion that spirits of the dead (human and animal) persist after corporeal death and exert an influence on the physical world. More precisely, as defined by Laythe and Owen (2012), haunting experiences denote internally perceived phenomena (e.g., sensations of a presence) or externally witnessed phenomena (e.g., objects moving), ascribed to spirit activity.

Belief in and experience of ghosts persists within modern society. Indeed, opinion polls report consistently that a substantial proportion of the general population believe in the existence of ghosts (Williams, Ventola, & Wilson, 2010). Illustratively, a 2005 Gallup survey, incorporating telephone interviews with 1,002 American adults, found that 32% of interviewees believed that ghosts (spirits of dead people) could return to certain places/situations; 37% considered houses could be haunted (Moore, 2005). These figures are commensurate with an earlier 2001 Gallup survey (Newport & Strausberg, 2001). MORI polls evidence similar levels of endorsement in Britain. The 2007 Survey on Beliefs, comprising telephone interviews with a representative quota sample of 1,005 adults, noted that 38% of interviewees believed in ghosts and 36% claimed to have seen a ghost (MORI, 2007). These figures concur with the 1998 MORI Paranormal Survey, which found that 40% of respondents believed in ghosts and 37% reported personal experience of ghosts (MORI, 1998). The prevalence of ghost-related beliefs and relatively frequent reporting of haunting experiences indicates the socially important nature of haunting phenomena and designates ghosts/hauntings as an important research area worthy of academic consideration.

Empirical attempts to explain ghost and haunt-related perceptions centre frequently on psychological factors. One significant variable is suggestibility. Generally, research has found associations between suggestibility and belief in the paranormal (Dafinoiu, 1995) and that experimental manipulation of

verbal suggestion can influence perception and recall of paranormal phenomena. For example, Wiseman, Greening, and Smith's (2003) study of séance phenomena, using self-selected delegates attending a Fortean Times convention, observed that participants who affirmatively answered the question "Do you believe that paranormal phenomena sometimes occur during séances?" were more susceptible to verbal suggestion about a séance-consistent phenomenon, movement of a hand-bell, than nonbelievers. Participants also reported experiencing unusual phenomena often associated with "genuine" séances; about a fifth believed the staged séance contained authentic paranormal phenomena, and a significantly greater percentage of believers considered this to be the case.

Similarly, Wiseman and Greening (2005) found that verbal suggestion affected perception of alleged paranormal key bending. Participants viewed footage of a performer (professional close-up magician) and an interviewer sitting at a table containing several objects (keys, pack of cards, cutlery, etc.). The camera showed a close-up of the performer's hands as he selected the key. They then used specious psychokinetic ability to produce a bend in the stem of the key (sleight of hand produced the distortion). The performer placed the key back on the table and the videotape footage concluded with a 60-s close-up of the distorted key. In the suggestion condition, a verbal comment on the soundtrack implied that the key continued to bend. The no-suggestion condition did not include this "bending" comment. Participants in the suggestion condition were more likely to report the key continuing to bend. In addition, participants reporting bending were highly confident that their testimony was reliable and were less likely to report the fake psychic's suggestion.

Wilson and French (2008) examined whether suggestion influenced recall of a psychic reading. Participants watched a scripted video of a reading followed by a "manipulated" interview in which the sitter commented upon the accuracy of the reading. In one version, the sitter declared correctly that the psychic mentioned the name Sheila and that Sheila was their mother's name. In the other version, the sitter asserted incorrectly that the psychic stated that their mother's name was Sheila. Wilson and French (2008) found believers showed a strong tendency to misremember the manipulated section of the reading, regardless of whether they received misinformation or not. Nonbelievers tended to remember the reading more accurately if no misinformation was present; however, in the misinformation condition their memories were as distorted as those of the believers.

Collectively, research shows that verbal suggestion can influence the perception and recall of supposed paranormal phenomena, especially when the suggestion is consistent with existing paranormal beliefs. Pertinently, research indicates that suggestion may play a critical role in the reporting of haunting phenomena (O'Keefe & Parsons, 2009). For example, Lange and Houran (1997) investigated whether the suggestion that a location was haunted would be sufficient to induce poltergeist-like perceptions (e.g., sensed presence). Participants attending an indoor performance theatre took part via random allocation to either the suggestion (paranormal activity) or control (renovation) condition. Each group went on a tour in which they visited five main theatre areas and completed an experiential questionnaire assessing their psychological and physiological perceptions. More intense experiences across measures demonstrated that mere suggestion could stimulate paranormal-type experiences.

Similarly, Terhune and Smith (2006), using a psychomanteum (mirror gazing task), established that suggestion could induce apparitional experiences. In the suggestion condition, instructions specified that mirror gazing could result in anomalous sensations (including seeing an apparition); the nonsuggestion condition advised only about the possibility of experiencing unusual bodily sensations/perceptual distortions. Suggestion influenced participants' perceptions; the suggestion condition produced more reports of visual and auditory apparitions.

Research has shown suggestibility (hypnotic and imaginative) is associated with the induction of anomalous/unusual experiences (Kumar & Pekala, 2001; O'Connor, Barnier, & Cox, 2008) and the stimulation of hallucinatory experiences (McConkey & Barnier, 2004). A notable example is the White Christmas paradigm (Barber & Calverey, 1964). The White Christmas test instructs participants to close their eyes and imagine hearing the famous Bing Crosby White Christmas song. After 30 s, participants rate the intensity of their imagery. Typically, a significant percentage of participants report hearing the song clearly. Mintz and

Alpert (1972) reported that the majority of schizophrenic patients (85%) and a nontrivial minority (40%) of controls reported a clear auditory image during the test (Merckelbach & van de Ven 2001). This finding illustrates that "normal" people will readily report suggested auditory events (Barber & Calverey, 1964).

Cumulatively, research demonstrates that the presentation of haunt-related suggestions can induce and heighten haunting-related perceptions (O'Keefe & Parsons, 2009). The relationship between suggestion, paranormal belief, and the reporting of unusual (ghost-related) experiences, however, is not a simple one and results across studies have been inconsistent. A classic illustration of this is the large-scale study conducted by Wiseman, Watt, Greening, Stevens, and O'Keefe (2002). They found that the frequency with which participants reported experiencing unusual phenomena in the past and attributed their experiences to ghosts varied as a function of level of belief. Believers reported experiencing seven of the eight unusual phenomena (feelings, sense of presence, sounds, changes in temperature, smells, sights, and tastes). The only phenomenon not reported more frequently by believers was an unusual sense of dizziness. Believers were also significantly more likely to attribute their experiences to ghosts. Similarly, when walking around Hampton Court Palace, an allegedly haunted location, believers noted more haunt-related experiences and demonstrated a greater tendency to attribute these experiences to a ghost.

Wiseman et al. (2002) also manipulated suggestion. In the positive suggestion condition, researchers told participants that an area was associated with increased reports of unusual phenomena. Contrastingly, in the negative suggestion condition, the researchers stated that the area was not associated with unusual phenomena. Suggestion had no effect on participants' expectations of experiencing unusual phenomena during the experiment or their tendency to attribute unusual phenomena to ghosts. In addition, the belief by suggestion interaction produced mixed results. When participants visited the allegedly haunted area, more believers reported unusual experiences in the positive suggestion condition. Generally, however, findings proved nonsignificant.

Overall, the results of haunt-related suggestion research are inconclusive and difficult to interpret. Generally, findings suggest that believers in the paranormal are particularly susceptible to suggestion and more easily deceived. However, this assumption is inconsistent with previous critical thinking research, which has failed to find consistent differences between believers and nonbelievers (Hergovich & Arendasy, 2005; Moore, Thalbourne, & Storm, 2010). In this context, philosophical bias may affect susceptibility to suggestion. Believers are prone to wrongly endorsing paranormal phenomena, whilst sceptics are inclined to deny the existence of paranormal effects. Hence, sceptics would be unlikely to detect genuine effects (if they exist).

The present study built on the work of Wiseman et al. (2002) in a number of important ways. First, to avoid problems linked to prior influence, the location used was fictitious. Real, historical locations (Hampton Court, Edinburgh Vaults, etc.) are often associated with hauntings and knowledge of this may influence participants' perceptions (Houran, Wiseman, & Thalbourne, 2002). Secondly, the manipulated suggestion appeared consistently throughout the presentation phase. It occurred on three occasions: the end of the narrative, within the narrative text, and as part of instructions prior to producing ratings. Contrastingly, Wiseman et al.'s (2002) use of suggestion was subtle (embedded within a talk about their experiment) and stated only once. Thirdly, only the haunting condition referred to the notion of haunting. This avoided any potential confound arising from stating that the location was not haunted. Mention of haunting, regardless of valence, may influence expectations, attention, and report rates. Previous work has indicated that attentional mechanisms play an important role in haunting experiences (Houran & Lange, 1996, 2001; Terhune, Ventola, & Houran, 2007). For this reason, the present study employed a haunted versus control (structure) manipulation.

Finally, to avoid the possible influence of external variables (magnetic fields, drafts, etc.) the current study took place in a controlled experimental setting. This ensured that participants' experiences were similar and prevented the introduction of physical conflates (temperature fluctuation, lighting variations, magnetic, electric fields, etc.). Using a controlled, nonhaunted location avoided issues arising from individual differences in perceptual acuity/sensitivity. There is evidence to suggest that believers in the paranormal may be more reactive to perceptual stimuli generally, and specifically more sensitive to possible paranormal

effects (Houran, Hughes, Thalbourne, & Delin, 2006; Thalbourne, 2010).

A further development was the inclusion of a reality testing measure, the Inventory of Personality Organization, IPO-RT (Lenzenweger, Clarkin, Kernberg, & Foelsch, 2001). Several studies report an association between proneness to reality testing deficits and unconventional beliefs, particularly belief in the paranormal and endorsement of urban legends and conspiracies (Drinkwater, Dagnall, & Parker, 2012; Irwin, 2004), and there have been no reported failures. Reality testing refers to the preference to test critically the coherent credibility of beliefs (Irwin, 2004). Reality-testing deficits bias individuals away from analytical-rational processing towards intuitive-experiential interpretations of anomalous events. Believers in this context are dependent upon, or favour, intuitive-experiential processing and consequently appraise perceptions and experiences less critically (Lindeman, 1998). Accordingly, they are more inclined to report unusual perceptual sensations (seeing things that do not exist, hearing things when there is no apparent reason, etc.). Similarly, emotion-based reasoning predicts level of paranormal belief (Irwin, Dagnall, & Drinkwater, 2012). Thus, believers tend to endorse paranormal occurrences because of their emotional rather than rational appeal (Sappington, 1990). In this context, the authors anticipated that participants scoring high on proneness to reality testing deficits would score higher on haunt-related expectancy measures. Similar to paranormal belief, the relationship between proneness to reality testing deficits and susceptibility to suggestion was less certain and, hence, worthy of investigation. This was tested in Phase 1 alongside the paranormal hypotheses.

The study comprised two distinct but related phases. Phase 1, experimental manipulation, examined the effects of verbal suggestion on paranormal believers (vs. nonbelievers) and perception of the fictitious hospital building. It was hypothesised, consistent with previous research, that believers (vs. nonbelievers) would expect the building to contain more haunt-related phenomena, and the differences would be greater with suggestion. Given the inconsistent nature of suggestion effects, the latter prediction was tentatively stated.

As outlined above, Phase 1 also addressed the relationship between suggestion and proneness to reality testing deficits. Previous research has found a positive association between belief in the paranormal and proneness to reality testing deficits (Drinkwater et al., 2012; Irwin, 2003, 2004). This relationship suggests, consistent with belief-in-the-paranormal research, that participants scoring high on proneness to reality testing deficits would be more susceptible to manipulation of suggestion. In the context of the present study, the suggestion that the building was haunted was likely to coincide with personal subjective, intuitive-experiential evaluations and existing paranormal-related schemas of haunted locations (antiquated, deserted, isolated, former hospital, etc.). Thus, we hypothesized that participants scoring high for proneness to reality testing deficits would correspondingly rate the expectation of haunt-related phenomena higher, and that this effect would be heightened when it was suggested that the location was haunted.

Phase 2 explored relationships between paranormal belief, proneness to reality testing deficits, the degree to which participants believed the building was haunted, and anticipation of haunt-related phenomena.

Method

Participants

One hundred eighty participants, recruited via opportunity sampling from undergraduate students and staff at Manchester Metropolitan University (faculties of Humanities, Languages & Social Sciences and the Department of Psychology), participated in the study. Mean age 20.89 years (SD = 4.99), range 18–48. The sample was 75% female (M = 20.96, SD = 5.42, 18–48 years) and 25% male (M = 20.71, SD = 3.48, 18–38 years). Participants responded to general e-mails, noticeboard advertisements, or personal requests to take part in a psychological study.

Hospital Photographs

A slideshow containing 24 photographs of hospitals was constructed. Pictures were obtained via an Internet search using the term "disused hospitals." Selected images were public domain, colour, non-distorted, taken during daytime, and representing a range of internal and external architectural features. Col-

lectively, images appeared to depict a single derelict hospital. Each image was mounted and fitted to a single PowerPoint 2010 slide (4:3 on screen; 25.40 cm x 19.05 cm). Slide presentation was set to 5 s and the slideshow ran for 2 min in total. A standard narration, read by a male member of staff, accompanied the slideshow and outlined the building's (fictitious) history. Recording of the narration was via a tablet and the initial recording was edited using Audacity software.

Suggestion Manipulation

The suggestion stated that the building had either a history of ghostly activity (haunted manipulation) or structural problems (structure manipulation): "In the years prior to its closure patients, staff and visitors reported ghostly sightings/structural problems and the building developed a reputation for being haunted/outdated. The centre of spectral activity/architectural concerns was the Administrative Building." To ensure equivalence between conditions, changes were restricted to three statements of the key manipulation phrase. These occurred at the end of the narrative, within the narrative text, and as part of instructions prior to producing ratings.

Video Tour of Administrative Building

After the slideshow, participants undertook a virtual (video) tour of the Administrative Building. The video comprised handheld camera footage of a building walkthrough, was in colour, and contained no sound. Filming was from the first person perspective and no people appeared in the shoot. The building was actually a disused location on one of the University campuses. Prior to the video shoot, the internal features of the building (i.e., doorways, staircase, furniture, room size, and décor) and the photographs of disused hospitals were carefully matched. This control ensured that participants were unlikely to perceive a mismatch between the slideshow and the virtual tour. Editing reduced the original recording to 10 min; coverage featured a range of internal features. Filming started on the ground floor, moved down to the cellar, and then progressed up through the second and third floors. Shooting concluded as the camera started to move back down the stairs. Recording used a Panasonic HC-V130EB-R Camcorder. Editing was via CyberLink PowerDirector.

Test Booklet

After watching the slideshow and finishing the virtual tour, participants completed a test booklet containing the following measures.

Environmental Perceptions and Phenomena Scale (EPPS). The EPPS is an 18-item measure adapted from previous research on hauntings and contains questions on a range of perceptions and sensations typically associated with hauntings (Laythe & Owen, 2012; Wiseman et al., 2002). Instructions asked participants to consider how they would feel when visiting and exploring the hospital Administrative Building late at night: "Within the Administrative Building would you expect to encounter the following phenomena?" Participants rated the extent to which they expected to encounter the following unusual, haunt-related phenomena (sensations and perceptions): unexplained voices; feeling of a sensed presence; fleeting visual shadows; heard footsteps; bouts of ringing in the ears; headache/migraine; perception of being touched; bouts of dizziness; periodic feelings of foreboding; heard bangs/raps; heard music from an unexplained source; unexplained odours; mist, lights, shadows, or other unexplained visual phenomena; notice of animals reacting to something that isn't there; sudden extreme cold or heat; heard or seen unexplained movement; and feelings of being watched. In this context, the EPPS provided a measure of the degree to which participants expected to encounter haunt-related phenomena within the Administrative Building. Participants rated each phenomenon on a Likert scale (0% = "certainly not" to 100% = "certainly".

History of Haunting (HH). A further question assessed whether participants believed the location was haunted: "To what extent do you believe that there is a history of paranormal activity (ghostly activity) at the location observed in the video (the Administrative Building)?" Participants responded on a 7-point

Likert scale (1 = "definitely not" to 7 = "definitely").

Haunting-Associated Opinions and Feelings (HAOF). A single item assessed general attitudes to the existence of ghosts: "To what extent do you believe that ghosts exist?" Responses were again assessed using a 7-point Likert scale (1 = "definitely not" to 7 = "definitely").

Two items measured the degree of anxiety ("How anxious do you feel when you think about ghosts?") and fear evoked by the notion of ghosts ("How fearful of ghosts are you?"). Responses were recorded on a 7-point Likert scale (1 = "not at all" to 7 = "extremely").

Revised Paranormal Belief Scale (R-PBS). The R-PBS is a self-report measure containing 26 questions measuring belief in seven facets of paranormal belief: Traditional Religious Belief, Psi Belief, Witchcraft, Spiritualism, Superstition, Extraordinary Life Forms, and Precognition (Tobacyk, 2004). The scale is a modified version of Tobacyk and Milford's (1983) Paranormal Belief Scale. Items are presented as statements (e.g., "I believe in God" and "Black magic really exists") measured on a 7-point Likert scale ranging from 1 ("strongly disagree") to 7 ("strongly agree"). Prior to analysis, scores were recoded 0–6 to facilitate Rasch scaling (Irwin, 2009). Final potential scores ranged from 0–156, with higher scores reflecting greater belief in the paranormal. Rasch scaling produced a 2-factor solution: New Age Philosophy, NAP, and Traditional Paranormal Belief, TPB (Lange, Irwin, & Houran, 2000). NAP (11 items) assesses belief in psi, reincarnation, altered states, and astrology, whilst TPB (5 items) measures traditional Western religious concepts such as the devil and witchcraft (Irwin, 2004). The Rasch scaling procedure (Andrich, 1988) produces scores ranging from 6.85 to 47.72 on NAP, and 11.16 to 43.24 on TPB. Previous research has established that the R-PBS is psychometrically and conceptually satisfactory (Tobacyk, 2004).

The Inventory of Personality Organization Reality Testing Subscale (IPO-RT). The IPO-RT (Lenzenweger et al., 2001) assesses proneness to reality testing deficits, and the capacities to differentiate self from non-self and intrapsychic from external stimuli, and to maintain empathy consistent with ordinary social criteria of reality (Kernberg, 1996). The IPO-RT also measures information-processing style (e.g., "I have heard or seen things when there is no apparent reason for it') without reference to psychotic symptomology. The scale contains 20 items assessed via a 5-point Likert scale (1 = "never true" to 5 = "always true"). Scores range from 20 to 100 (low scores indicate high reality-testing ability). The IPO-RT is temporally stable with nonclinical populations and is psychometrically established, demonstrating good retest reliability and construct validity (Lenzenweger et al., 2001).

Subjective Paranormal Experiences Scale (SPES). In addition to the previously outlined measures, participants completed the SPES, which measures participants' general paranormal history and assesses incidence of subjective paranormal experiences. The measure was not analysed in the present study because the number of respondents reporting hauntings was low.

Procedure

Ethical approval was granted as part of a wider research project examining the relationship between paranormal beliefs and anomalous experiences. Prior to testing, all participants provided informed consent and were aware of their right to withdraw from the study at any time.

Participant testing occurred in groups or individually and haunt condition allocation was random (suggestion manipulation: haunted vs. structure).

Prior to testing, instructions informed participants that the session would involve attending to a brief slideshow and accompanying audio narrative about an old disused hospital. Participants first watched the slideshow containing the hospital photographs and simultaneously listened to the supporting audio narrative. The first presentation of the suggestion manipulation occurred within the concluding sentence of the narrative. On conclusion of the presentation, participants read a text version of the narrated commentary in preparation for the virtual (video) tour of the old Administrative Building. This served to reinforce the previously narrated suggestion (second presentation of the suggestion manipulation). After allowing sufficient reading time, the recording (virtual tour) was played. Following the virtual tour, instructions guided participants to complete the test booklet containing the study measures. The instructions preceding the test

measures stated the suggestion manipulation a third and final time.

Participants completed the environmental measures (EPPS and HH) and the HOAF first to avoid priming paranormal ideation. To avoid potential order effects the sequence of the remaining measures (R-PBS and IPO-RT) was counterbalanced. Guidelines instructed participants to complete all questions and that there was no time limit. At the end of the testing session, testers debriefed the participants, thanked them for engaging with the research, and provided follow-up details.

Results

Primary Analysis: Reliability and Scale Descriptives

Prior to the main analysis, Cronbach's alpha (α) assessed the internal reliability of the measures. All scales proved psychometrically acceptable: The Environmental Perceptions Phenomena scale (EPPS; α = .93), proneness to reality testing deficits measure (IPO-RT; α = .90) and Revised Paranormal Belief Scale, (R-PBS; α = .89) demonstrated excellent internal reliability. The two R-PBS factors (New Age Philosophy, NAP; and Traditional Paranormal Belief, TPB) produced alpha coefficients in the good (NAP, α = .83) to acceptable TPB (α = .70) range (see also George & Mallery, 2003). Scale descriptive statistics appear in Table 1.

Table 1
Scale Descriptive Statistics (*N* = 180)

	α	*M*	*SD*	Range
EPPS	.93	43.73	18.09	2.35–82.94
R-PBS	.89	51.97	24.01	4.00–116.00
NAP	.83	21.60	3.88	6.85–31.41
TPB	.70	22.59	4.56	11.16–39.23
IPO-RT	.90	42.81	11.91	20.00–83.00
HH	NA	4.12	1.84	1.00–7.00
OG	NA	4.18	2.04	1.00–7.00
FA	NA	3.61	1.86	1.00–7.00
FF	NA	3.41	1.92	1.00–7.00

Note. EPPS = Environmental Perceptions Phenomena Scale, R-PBS = Revised Paranormal Belief Scale, NAP = New Age Philosophy, TPB = Traditional Paranormal Belief, IPO-RT = Proneness to Reality Testing Deficits, HH = History of Haunting, OG = Opinion Ghosts Exist, FA = Feeling Anxiety, FF = Feeling Fear.

EPPS correlated positively with level of paranormal belief and IPO-RT, indicating that expectancy of haunt-related phenomena increased as a function of level of paranormal belief and reality testing deficit. Paranormal belief (R-PBS, NAP, and TPB) correlated positively with IPO-RT. All observed interscale correlations were in the moderate range (see Table 2).

Experimental Manipulation: Paranormal Belief, Proneness to Reality Testing Deficits, Suggestion, and Haunting-Related Ratings

Consistent with previous research, median splits classified participants as either low or high (below vs. above median) on R-PBS and IPO-RT. This procedure is an established methodological convention within parapsychological and individual differences research (Wilson & French, 2014; Wiseman &

Greening, 2005) that has recently received disapproval. Ubiquitously cited criticisms of median split often reference MacCallum, Zhang, Preacher, and Rucker (2002), who outline potential negative analytical consequences arising from dichotomization of continuous variables (i.e., loss of information about individual variability, ensuing loss of power and effect size, and the undermining of measurement reliability). Accordingly, MacCallum et al. (2002) concluded that dichotomization is rarely defensible because it produces misleading results.

Table 2
Matrix of Pearson Product Moment Correlations Between Study Variables

	1	2	3	4	5	6	7	8	9
1. EPPS									
2. R-PBS	.57*								
3. NAP	.52	.85							
4. TPB	.44	.81	.65						
5. IPO-RT	.40	.35	.37	.21					
6. HH	.66	.59	.52	.49	.20				
7. OG	.52	.61	.54	.52	.21	.74			
8. FA	.41	.32	.23	.31	.22	.39	.43		
9. FF	.33	.29	.17	.29	.19	.33	.39	.84	

Note. EPPS = Environmental Perceptions Phenomena Scale, R-PBS = Revised Paranormal Belief Scale, NAP = New Age Philosophy, TPB = Traditional Paranormal Belief, IPO-RT = Proneness to Reality Testing Deficits, HH = History of Haunting, OG = Ghosts Exist, FA = Feeling Anxiety, FF = Feeling Fear.
*All $p < .01$.

A recent paper by Iacobucci, Posavac, Kardes, Schneider, and Popovich (2014) challenges and re-evaluates this view. Iacobucci et al. (2014) conducted an examination of median splits, which revealed that spurious effects were negligible and limited (usually) to instances where predictor variables correlated significantly among themselves, that is, cases of acute multicollinearity. Indeed, even when variables are nonorthogonal, the robust nature of a 2 by 2 factorial design typically negates extreme distortions (Iacobucci et al., 2014). Finally, it is worth noting that multicollinearity generally does not promote type I errors. Instead, it reduces effect sizes and increases the likelihood of type II error (Iacobucci et al., 2014).

We assessed the validity of our findings in two ways. Firstly, following Iacobucci et al.'s (2014) recommendation, we demonstrated orthogonality by computing correlation coefficients between the median split variables (belief in the paranormal and proneness to reality testing deficits) and the experimental factor (level of suggestion). The correlations were nonsignificant: R-PBS, $r(178) =. 02$, $p = .37$; IPO-RT, $r(178) = .00$, $p = .50$.

As a second validity check, consistent with Wilson and French (2014), we conducted a series of multiple regressions (not reported), which replicated the pattern of results observed from the ANOVAs reported below (see Tables 3 and 4).

ANOVA analysis was preferred to regression analysis because suggestion was an experimental manipulation and we were concerned primarily with exploring potential interactions. Separate ANOVAs for level of paranormal belief (R-PBS) and proneness to reality testing deficits (IPO-RT) were justified because the correlation between these two factors was in the low (see Hinkle, Wiersma, & Jurs, 2003) to moderate (see Cohen, 1992) range. Thus, effects of level and suggestion on environmental perceptions (EPPS) and history of haunting (HH) were analyzed via separate 2 (level of paranormal belief/level of reality testing: high vs. low) x 2 (suggestion: structure vs. haunted) independent ANOVAs.

Paranormal Belief

Means and standard deviations appear in Table 3.

Table 3
Level of Paranormal Belief (R-PBS) as a Function of Haunt-Related Attitudes (Environmental Perceptions Phenomena Scale, EPPS, and History of Haunting, HH)

	Level of Paranormal Belief					
	Below Median		Above Median		Overall	
	M	*SD*	*M*	*SD*	*M*	*SD*
EPPS						
Suggestion Type						
Structure	35.58	18.21	51.15	16.02	43.28	18.78
Haunted	36.62	16.67	52.63	14.33	44.19	17.48
Overall	36.12	17.35	51.88	15.14		
HH						
Suggestion Type						
Structure	3.53	2.02	5.07	1.32	4.29	1.87
Haunted	3.10	1.73	4.88	1.40	3.95	1.81
Overall	3.31	1.88	4.98	1.36		

Environmental Perceptions Phenomena Scale (EPPS). No significant main effect on EPPS was found for suggestion (structure, $M = 43.28$, $SD = 18.78$ vs. haunted, $M = 44.19$, $SD = 17.48$), $F(1, 176) = 0.27$, $p > .05$, $\eta p^2 = .002$. A significant main effect was found for level of paranormal belief, $F(1, 176) = 41.68$, $p < .001$, $\eta p^2 = .19$. Participants scoring above the median on paranormal belief scored higher on the EPPS ($M = 51.88$, $SD = 15.14$) than participants below the median ($M = 36.12$, $SD = 17.35$). Higher EPPS scores indicate a higher expectancy of haunt-related phenomena within the Administrative Building. No significant interaction was found between suggestion and level of paranormal belief on EPPS, $F(1, 176) = .01$, $p > .05$, $\eta p^2 = .00$.

History of Haunting (HH). No significant main effect on HH was found for suggestion (structure, $M = 4.29$, $SD = 1.87$ vs. haunted, $M = 3.95$, $SD = 1.81$), $F(1, 176) = 1.56$, $p > .05$, $\eta p^2 = .009$. A significant main effect was found for level of paranormal belief, $F(1, 176) = 45.53$, $p < .001$, $\eta p^2 = .21$. Participants scoring above the median on paranormal belief scored higher on HH ($M = 4.98$, $SD = 1.36$) than participants below the median ($M = 3.31$, $SD = 1.88$). Higher HH scores indicate a higher expectancy of haunt-related phenomena within the Administrative Building. No significant interaction was found between suggestion and level of paranormal belief on HH, $F(1, 176) = 0.25$, $p > .05$, $\eta p^2 = .001$.

Proneness to Reality Testing Deficits

Means and standard deviations appear in Table 4.

Environmental Perceptions Phenomena Scale (EPPS). No significant main effect on EPPS was found for suggestion (structure, $M = 43.28$, $SD = 18.78$ vs. haunted, $M = 44.19$, $SD = 17.48$), $F(1, 176) = 0.12$, $p > .05$, $\eta p^2 = .001$. A significant main effect was found for IPO-RT, $F(1, 176) = 17.92$, $p < .001$, $\eta p^2 = .092$. Participants scoring above the median on IPO-RT scored higher on the EPPS ($M = 49.28$, $SD = 16.98$) than participants below the median ($M = 38.32$, $SD = 17.57$). Higher EPPS scores indicate a higher

expectancy of haunt-related phenomena within the Administrative Building. No significant interaction was found between suggestion and IPO-RT on EPPS, $F(1,176) = .05$, $p > .05$, $\eta p^2 = .00$.

Table 4
Level of Proneness to Reality Testing Deficits (IPO-RT) as a Function of Haunt-Related Attitudes (Environmental Perceptions Phenomena Scale, EPPS, and History of Haunting, HH)

	Level of RT					
	Below Median		Above Median		Overall	
	M	*SD*	*M*	*SD*	*M*	*SD*
EPPS						
Suggestion Type						
Structural	37.57	18.54	49.12	17.34	43.28	18.78
Haunted	39.05	16.73	49.43	16.82	44.19	7.48
Overall	38.32	17.57	49.28	16.98		
HH						
Suggestion Type						
Structural	3.80	2.04	4.80	1.53	4.29	1.87
Haunted	3.74	1.88	4.16	1.73	3.95	1.81
Overall	3.77	1.95	4.47	1.66		

History of Haunting (HH). No significant main effect on HH was found for suggestion (structure, $M = 4.29$, $SD = 1.87$ vs. haunted, $M = 3.95$, $SD = 1.81$), $F(1, 176) = 1.69$, $p > .05$, $\eta p^2 = .010$. A significant main effect was found for IPO-RT, $F(1, 176) = 6.86$, $p = .01$, $\eta p^2 = .038$. Participants scoring above the median on IPO-RT scored higher on the HH ($M = 4.47$, $SD = 1.66$) than participants below the median ($M = 3.77$, $SD = 1.95$). Higher HH scores indicate a higher expectancy of haunt-related phenomena within the Administrative Building. No significant interaction was found between suggestion and IPO-RT on HH, $F(1, 176) = 1.15$, $p > .05$, $\eta p^2 = .007$.

Haunt Associated Opinions and Feelings (HAOF)

In light of the lack of a suggestion effect and for brevity's sake, subsequent analysis focused on only belief in the paranormal and proneness to reality testing deficits. A series of independent *t* tests assessed differences on the opinion and feeling measures (HAOF), as shown in Table 5.

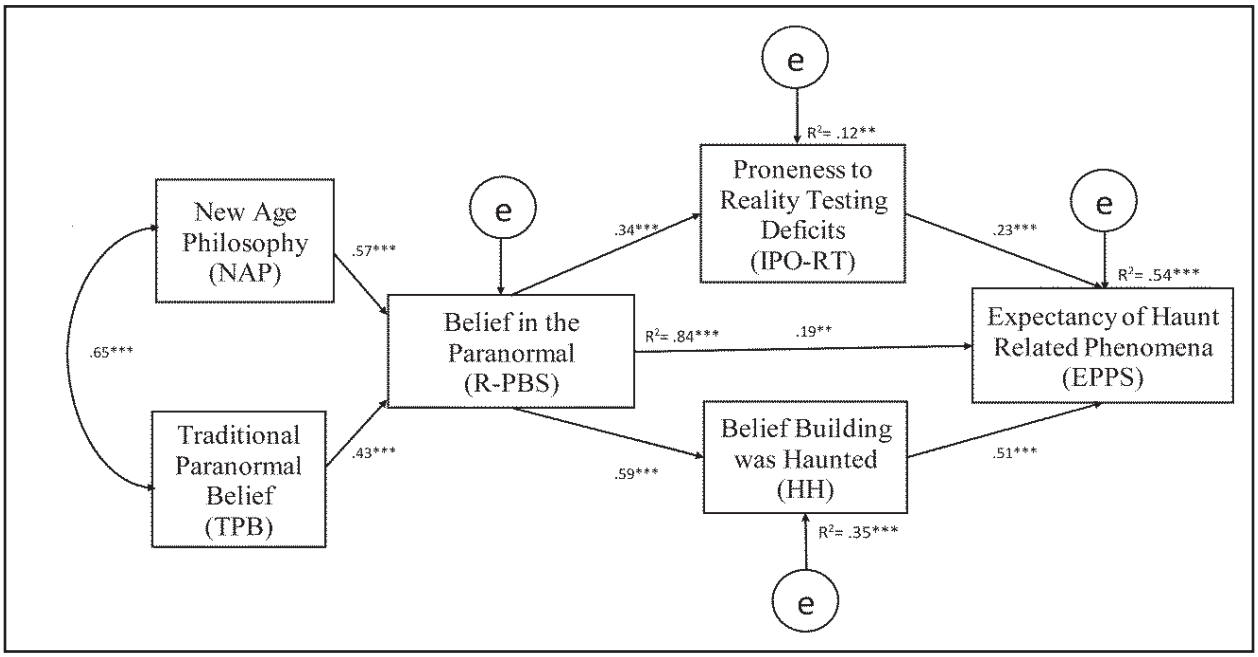
Participants high (vs. low) on R-PBS scored higher on the opinion (ghosts exist) and feeling measures (anxious and fearful). Participants high (vs. low) on IPO-RT scored higher on opinion (ghosts exist). Following application of a Bonferroni correction for multiple comparisons, the feeling measures (anxious and fearful) were not significant. The adjusted alpha level was based upon the notion of family: the smallest set of meaningful inferences within a set of analyses. There were three planned comparisons within each independent variable assessing levels of paranormal belief and proneness to reality testing deficits; hence, the corrected alpha was $p = .017$.

Path Analysis: Paranormal Belief, Reality Testing, and Haunt-Related Phenomena

Within these data, there were no issues with multicollinearity; all correlations were below .9 (see Table 2 for zero-order correlations). All correlations were significant and in the expected direction. A path model examining direct effects (R-PBS on EPPS) and indirect effects (R-PBS on EPPS, mediated by IPO-RT and HH was constructed.

Table 5
Paranormal Belief and Reality Testing Comparisons (Low vs. High) on Haunt-Associated Opinions and Feelings (HAOF)

	Level							
	Below Median		Above Median					
	M	*SD*	*M*	*SD*	*t*	*df*	*p*	*d*
Paranormal Belief								
Opinion								
Ghosts Exist	3.32	2.00	5.09	1.65	-6.49	175.50	<.001	.97
Feelings								
Anxious	3.17	1.96	4.08	1.64	-3.38	175.80	<.001	.50
Fearful	2.96	1.94	3.90	1.79	-3.39	177.96	<.001	.51
Proneness to Reality Testing Deficits								
Opinion								
Ghosts Exist	3.81	2.05	4.55	1.96	-2.46	178	.008	.37
Feelings								
Anxious	3.34	1.90	3.89	1.79	-1.99	178	.025	.30
Fearful	3.13	1.98	3.70	1.82	-1.99	178	.024	.30



$p <$. 05*, $p <$ .01**, $p <$ .001***

Figure 1. The outcome of AMOS 19 path analysis, putative relations between paranormal beliefs subscales (NAP and TPB), level of paranormal belief (R-PBS), proneness to reality testing deficits (IPO-RT), belief the building was haunted (HH) and haunt-related expectations (EPPS). The effects of extraneous variables are depicted by the use of "e" to denote error term.

Standardized estimates, covariance, and regression coefficients (R^2) appear in Figure 1. Kline (2005) recommends several criteria for assessing goodness-of-fit: chi-square, the comparative fit index

(CFI), standardized root mean square residual (SRMR), and root mean square error of approximation (RMSEA). These indices indicated very good overall model fit; chi-square is nonsignificant, $\chi^2(7, N = 180) = 7.79$, $p = .35$; CFI = .99; SRMR = .02; and RMSEA = .02. Hu and Bentler (1999) examined indices cutoffs and suggested that type I and type II errors were best minimised via a combination of relative fit indexes (CFI ≥ 0.95 is indicative of good fit) and absolute fit indices (SRMR, good models < .08; or RMSEA, good models < .06). Path coefficients were significant at the $p < .05$ level. Paranormal belief (R-PBS) had a significant effect on proneness to reality testing deficits (IPO-RT), the belief the building was haunted (HH), and haunt-related expectations (EPPS). IPO-RT and HH also significantly affected EPPS.

Direct and Indirect Effects

To assess whether direct and indirect effects were statistically significant, a mediation analysis using the bias-corrected bootstrap 95% confidence intervals (CI) procedure (Hayes, 2013) was applied with 5,000 bootstrap samples (findings are reported in Table 6). This analysis further examined the specific influence of each proposed mediator; AMOS is unable to examine the unique influence of two or more mediators when they are simultaneously included in a path diagram.

Table 6
Mediation Analysis Direct and Indirect Effects

	Standardized Direct Effect (DE)	Bias-Corrected 95% CI (DE)	Standardized Indirect Effect (IE)	Bias-Corrected 95% CI (IE)
Haunt-Related Expectations (EPPS)				
Paranormal Belief (R-PBS)	.18	.48–.66*	.37*	.28–.47*
Belief Building Haunted (HH)	.50	.19–.46*	None	None
Reality Test Deficits (IPO-RT)	.24	.07–.31*	None	None

* $p < .001$

Significant direct effects were observed on EPPS, R-PBS, HH, and IPO-RT. Paranormal belief (R-PBS) had a significant indirect effect on haunt-related expectations (EPPS). To discern the influence of each proposed mediator (HH and RT) on the relationship between R-PBS and EPPS, Preacher and Hayes' (2008) INDIRECT bootstrapping macro was run. The direct relationship between R-PBS and EPPS was nonsignificant ($b = .12$), indicating that belief the building was haunted (HH) and reality testing deficit (IPO-RT) mediated the relationship between R-PBS and EPPS. Further inspection revealed that IPO-RT was not a significant mediator between R-PBS and EPPS at the 99% confidence level across bias-corrected (99% CI = -.01 to .13) point estimates. HH, however, emerged as significant, and the indirect effect of R-PBS and EPPS through HH was significant at the 99% confidence level across bias corrected (99% CI = .05 to .34) point estimates. The overall model accounts for 53% of the total variance on EPPS.

Discussion

Prior to replication, significant effects within the present study require cautious interpretation. The present paper produced a number of key findings. Within the experimental phase, expectation of haunt-related phenomena as measured by the Environmental Perceptions Phenomena Scale and belief the building was haunted increased as a function of level of paranormal belief and proneness to reality testing deficits. Suggestion, however, had no effect on expectation of haunt-related phenomena. Phase 2 confirmed and expanded upon these findings. Path analysis revealed significant relationships between belief in the paranor-

mal, proneness to reality testing deficits, belief the building was haunted, and expectation of haunt-related phenomena. Further analysis found belief the building was haunted significantly mediated the relationship between paranormal belief and haunt-related expectations. Finally, whilst correlating positively with each other, both belief in the paranormal and proneness to reality testing deficits (information processing style) explained unique variance within haunt-related expectations. The contribution of reality testing deficits, however, in comparison to belief was relatively minor.

The observed association between level of paranormal belief and haunt-related expectations concurs with Wiseman et al. (2002), who found a relationship between belief in ghosts, reporting paranormal experiences, and attributing unusual experiences to a ghost. In the present study, believing in the paranormal predisposed participants to anticipate haunt-related phenomena within the fictitious Administrative Building, especially when participants considered the location haunted. This finding is consistent with previous work on context mediation and psychological haunting research generally (Harte, 2000; Houran & Lange, 1996, 2001; Lange & Houran, 1997). Expectation in the context of real-world locations (disused structures, historical buildings, battlefields, etc.) may arise from physical characteristics (appearance, age, structure, etc.) and social characteristics (e.g., history and reputation). Similarly, proneness to reality testing deficits, and an associated preference for intuitive-experiential information processing, predisposed participants to expect more haunt-related phenomena within the fictitious location. This finding adds to the emergent body of research revealing associations between proneness to reality testing deficits and unconventional beliefs (Drinkwater et al., 2012; Irwin, 2004).

The observation that suggestion had no influence in the current study requires careful elucidation. Previous research has produced mixed results. Generally, research indicates that suggestion affects the perception and recall of paranormal phenomena (French & Wilson, 2006; Wiseman & Greening, 2005; Wiseman et al., 2002), especially in believers, when the suggestion is consistent with a particular belief. Haunting findings, however, are less conclusive. For example, Wiseman et al. (2002) found mixed results. Suggestion had no effect on perception of unusual phenomena in the current study, nor did it influence the degree to which participants attributed unusual events to a ghost. The only prediction to reach significance was that believers (vs. disbelievers) in the positive suggestion condition reported more experiences.

Although the present manipulation failed to produce significant effects, suggestion may still play a role in priming haunt-related experiences. Indeed, previous work has evidenced that suggestion can influence haunt-related perceptions (cf. Houran & Lange, 1996). Clearly, further work is required to identify factors effecting susceptibility to haunting experiences. One obvious potential manipulation is cue salience. The present study used a virtual tour, whilst the significant effect found by Houran and Lange (1996) occurred within a "real" physical setting. Note that Wiseman et al. (2002) failed to find compelling effects in a real-life setting (Hampton Court).

Collectively, haunting research indicates that the effects of suggestion are subtle and less important than belief in the paranormal. More generally, suggestion may influence perception and recall of paranormal events such as spoon bending, fake séances, and psychic readings. Thus, future studies need to explore the conditions under which suggestion influences haunt-type perceptions and recollections. The main factors to consider are whether participants high in haunting belief are more inclined to be affected by haunt-related suggestions (specific beliefs), environmental manipulation (allegedly haunted vs. manufactured setting), and past experience (haunting vs. no haunting). This would produce a set of complex variable manipulations sensitive to potential suggestion effects. Finally, believers gave higher ratings for believing the building was haunted and ghost existence. In addition, believers produced higher anxiety and fear ratings.

This study has some potential limitations. For example, participants completed general questionnaire measures in a single test session. Whilst the order of measures was counterbalanced, previous studies found that personality measures administered in this way may be prone to context effects (Council, 1993). Context effects occur when correlations between measures vary as a function of whether scales were completed in a single testing session (resulting in a high correlation) or in separate/unrelated sessions (resulting in a lower or no correlation). Context effects arise within single session testing situations because participants wrongly infer relationships between scales and become inadvertently motivated to provide consistent responses (Council, Kirsch, & Hafner, 1986). For example, Council (1993) found the typically

cited positive correlation between hypnotisability and absorption only when participants completed both measures in a single testing session. Whilst context effects are a consideration, likelihood of occurrence differs according to the characteristics of specific personality measures and research situations. Pertinent to the present study, Roig, Bridges, Renner, and Jackson (1998), using the Paranormal Belief Scale and the Irrational Beliefs Inventory, found that whilst correlation coefficients between the measures became weaker as a function of context controls, there were no statistically significant effects of context. Similarly, Nadon, Hoyt, Register, and Kihlstrom (1991) failed to find evidence that the repeatedly observed correlations between absorption and hypnosis measures were artifacts of testing context.

Subsequent studies could extend the current study by including a suggestibility measure. Previous research has found associations between the reporting of anomalous experiences and belief in paranormal phenomena (Kumar & Pekala, 2001). Thus, individuals scoring high on suggestibility may be more likely to respond to the verbal suggestion and report higher expectations, especially when the suggestion is consistent with their beliefs. Suggestion may be less likely to affect participants with low suggestibility regardless of whether the suggestion is consistent with their belief or not (cf. McGeown et al., 2012).

Further evaluation of haunt-related expectations could examine the distinction between proneness to anomalous experiences and tendency to draw upon paranormal interpretations. The present study included a range of phenomena associated with, but not exclusive to, haunt-related experiences. Participants imagining entering an unfamiliar environment may anticipate strange sensations and perceptions but not necessarily attribute them to ghosts. Further work could consider the relative contribution of attentional and attributional processes and investigate whether believers and nonbelievers differ on both expectations for unusual and haunt-related phenomena, using, for example, the Survey of Anomalous Experiences (Irwin, Dagnall, & Drinkwater, 2013).

Analysis of direct and indirect effects demonstrated that the belief the building was haunted mediated the relationship between paranormal belief (R-PBS) and expectations of haunt-related phenomena. The emergent haunting model contributes to the work of Houran and Lange (Houran & Lange, 1996; Lange & Houran, 1997). They found that context cues and demand characteristics could induce poltergeist-like perceptions and stimulate paranormal-type experiences. The present study explicates that this is most likely to occur when belief in the paranormal is high and the location considered haunted, especially when the individual believes hauntings are genuine. This finding is an important addition to the literature because, as Laythe and Owen (2012) note, the generally utilised measure of paranormal belief, the R-PBS (Tobacyk, 2004), fails specifically to assess specific haunt-related content and hence underestimates the relationship between belief and perceived ghostly activity. In the present study we used the R-PBS for two reasons. Firstly, the measure is conceptually and psychometrically satisfactory (Tobacyk, 2004). Secondly, the R-PBS is the most widely used instrument of paranormal belief (Goulding & Parker, 2001). Our findings, however, suggest that specific haunt-related beliefs rather than general belief in the paranormal increases haunt-related expectations. Thus, specific beliefs may provide a cognitive framework for structuring haunting-related events and experiences (Houran et al., 2002). This is consistent with Wiseman et al.'s (2002) finding that those who believed in ghosts (vs. nonbelievers) reported more unusual experiences as they walked around Hampton Court Palace. Thus, believers in hauntings, via their beliefs, are inclined to expect haunt-related phenomena in purportedly haunted locations.

As a brief postscript, we would like to contextualize our conclusions. Whilst our results indicate that beliefs may give rise to haunting-like perceptions, they do not imply that all haunting/ghost experiences are explainable via psychological mechanisms. As Broughton (1991) eloquently notes, conventional explanations of hauntings and ghost-related phenomena have yet to provide comprehensive theories that account for the breadth of encounters reported. The present study adds to the body of research demonstrating that psychological processes can explain haunting-like phenomena.

References

Andrich, D. (1988b). *Rasch models for measurement*. Newbury Park, CA: Sage.

Barber, T. X., & Calverey, D. S. (1964). An experimental study of hypnotic (auditory and visual) hallucinations. *Journal*

of Abnormal and Social Psychology, *68*, 13–20.
Broughton, R. S. (1991). *Parapsychology: The controversial science*. New York, NY: Ballantine.
Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.
Council, J. R. (1993). Context effects in personality research. *Current Directions in Psychological Science*, *2*, 31–34.
Council, J. R., Kirsch, I., & Hafner, L. P. (1986). Expectancy versus absorption in the prediction of hypnotic responding. *Journal of Personality and Social Psychology*, *50*, 182–189.
Dafinoiu, I. (1995). Self-attribution, expectancy, suggestibility. *Revue Roumaine de Psychologie, 39*, 139–145.
Drinkwater, K., Dagnall, N., & Parker, A. (2012). Reality testing, conspiracy theories, and paranormal beliefs. *Journal of Parapsychology*, *76*, 57–78.
French, C. C., & Wilson, K. (2006). Incredible memories: How accurate are reports of anomalous events? *European Journal of Parapsychology*, *21*, 166–181.
George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston, MA: Allyn & Bacon.
Goulding, A., & Parker, A. (2001). Finding psi in the paranormal: Psychometric measures used in research on paranormal beliefs/experiences and in research on psi-ability. *European Journal of Parapsychology*, *16*, 73–101.
Harte, T. M. (2000). Contextual mediation of perceptions during hauntings and poltergeist-like experiences: A replication and extension. *Perceptual and Motor Skills*, *91*, 451–459.
Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.
Hergovich, A., & Arendasy, M. (2005). Critical thinking ability and belief in the paranormal. *Personality and Individual Differences*, *38*, 1805–1812.
Hinkle, D. E., Wiersma, W., & Jurs. S. G. (2003). *Applied statistics for the behavioral sciences* (5th ed.). Boston, MA: Houghton Mifflin.
Houran, J., Hughes, L. F., Thalbourne, M. A., & Delin, P. S. (2006). Quasi-experimental study of transliminality, vibrotactile thresholds and performance speed. *Australian Journal of Parapsychology*, *6*, 54–80.
Houran, J., & Lange, R. (1996). Diary of events in a thoroughly unhaunted house. *Perceptual and Motor Skills*, *83*, 499–502.
Houran, J., & Lange, R. (Eds.). (2001). *Hauntings and poltergeists: Multidisciplinary perspectives*. Jefferson, NC: McFarland.
Houran, J., Wiseman, R., & Thalbourne, M. (2002). Perceptual-personality characteristics associated with naturalistic haunt experiences. *European Journal of Parapsychology*, *17*, 17–44.
Hu, L.T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.
Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J., & Popovich, D. L. (2014). Toward a more nuanced understanding of the statistical properties of a median split. *Journal of Consumer Psychology*. Retrieved from http://dx.doi.org/10.1016/j.jcps.2014.12.002
Irwin, H. J. (2003). Reality testing and the formation of paranormal beliefs. *European Journal of Parapsychology*, *18*, 15–27.
Irwin, H. J. (2004). Reality testing and the formation of paranormal beliefs: A constructive replication. *Journal of the Society for Psychical Research*, *68*, 143–152.
Irwin, H. J. (2009). *The psychology of paranormal belief: A researcher's handbook*. Hatfield, England: University of Hertfordshire Press. Amended.
Irwin, H. J., Dagnall, N., & Drinkwater, K. (2012). Paranormal belief and biases in reasoning underlying the formation of delusions. *Australian Journal of Parapsychology*, *12*, 7–21.
Irwin, H. J., Dagnall, N., & Drinkwater, K. (2013). Parapsychological experience as anomalous experience plus paranormal attribution: A questionnaire based on a new approach to measurement. *Journal of Parapsychology*, *77*, 39–53.
Kernberg, O. F. (1996). A psychoanalytic theory of personality disorders. In J. F. Clarkin & M. F. Lenzenweger (Eds.), *Major theories of personality disorder* (pp. 106–140). New York, NY: Guilford Press.
Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.
Kumar, V. K., & Pekala, R. J. (2001). Relation of hypnosis-specific attitudes and behaviours to paranormal beliefs and experiences: A technical review. In J. Houran & R. Lange (Eds.), *Hauntings and poltergeists: Multidisciplinary perspectives* (pp. 260–279). Jefferson, NC: McFarland.
Lange, R., & Houran, J. (1997). Context-induced paranormal experiences: Support for Houran and Lange's model of haunting phenomena. *Perceptual and Motor Skills*, *84*, 1455–1458.

Lange, R., Irwin, H. J., & Houran, J. (2000). Top-down purification of Tobacyk's Revised Paranormal Belief Scale. *Personality and Individual Differences*, *29*, 131–156.

Laythe, B., & Owen, K. (2012). Paranormal belief and the strange case of haunt experiences: Evidence of a neglected population. *Journal of Parapsychology*, *16*, 79–108.

Lenzenweger, M. F., Clarkin, J. F., Kernberg, O. F., & Foelsch, P. A. (2001). The Inventory of Personality Organization: Psychometric properties, factorial composition, and criterion relations with affect, aggressive dyscontrol, psychosis proneness, and self-domains in a nonclinical sample. *Psychological Assessment*, *13*, 577–591.

Lindeman, M. (1998). Motivation, cognition and pseudoscience. *Scandinavian Journal of Psychology*, *39*, 257–265.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*, 19–40.

McConkey, K. M., & Barnier, A. J. (2004). The highly hypnotisable person: Unity and diversity in behaviour and experience. In M. Heap, R. Brown, & D. Oakley (Eds.), *The highly hypnotisable person: Theoretical, experimental and clinical issues* (pp. 61–84). London, England: Brunner-Routledge.

McGeown, W. J., Venneri, A., Kirsch, I., Nocetti, L., Roberts, K., Foan, L., & Mazzoni, G. (2012). Suggested visual hallucination without hypnosis enhances activity in visual areas of the brain. *Consciousness and Cognition*, *21*, 100–116.

Merckelbach, H., & van de Ven, V. (2001). Another White Christmas: Fantasy proneness and reports of "hallucinatory experiences" in undergraduate students. *Journal of Behavior Therapy and Experimental Psychiatry*, *32*, 137–144.

Mintz, S., & Alpert, M. (1972). Imagery vividness, reality testing and schizophrenic hallucinations. *Journal of Abnormal Psychology*, *79*, 310–316.

Moore, D. W. (2005, June 16). Three in four Americans believe in paranormal. *Gallup News Service*, p.161. Retrieved from *www.gallup.com*

Moore, T., Thalbourne, M. A., & Storm, L. (2010). A study on coincidences. *Australian Journal of Parapsychology*, *10*, 154–176.

MORI. (1998). Paranormal survey [Conducted for the Sun newspaper]. Retrieved from http://www.ipsosmori.com/researchpublications/researcharchive/poll.aspx?oItemId=2082

MORI. (2007). Survey on beliefs [Conducted for Schott's Almanac]. Retrieved from http://www.ipsos-mori.com/researchpublications/researcharchive/246/Survey-on-Beliefs.aspx

Nadon, R., Hoyt, I. P., Register, P. A., & Kihlstrom, J. F. (1991). Absorption and hypnotizability: Context effects reexamined. *Journal of Personality and Social Psychology*, *60*, 144–153.

Newport, F., & Strausberg, M. (2001, June 8). Americans' belief in psychic and paranormal phenomena is up over last decade. *Gallup News Service.* Retrieved from http://www.gallup.com/poll/releases.

O'Connor, A. R., Barnier, A. J., & Cox, R. E. (2008). Déjà vu in the laboratory: A behavioral and experiential comparison of posthypnotic amnesia and posthypnotic familiarity. *International Journal of Clinical and Experimental Hypnosis*, *56*, 425–450.

O'Keefe, C., & Parsons, S. (2009). Haunting experiences: An investigative approach. In M. D. Smith (Ed.), *Anomalous experiences: Essays from parapsychological and psychological perspectives* (pp.108–119). Jefferson, NC: McFarland.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891.

Roig, M., Bridges, K. R., Renner, C. H., & Jackson, C. R. (1998). Belief in the paranormal and its association with irrational thinking controlled for context effects. *Personality and Individual Differences*, *24*, 229–236.

Sappington, A. A. (1990). The independent manipulation of intellectually and emotionally based beliefs. *Journal of Research in Personality*, *24*, 487–509.

Terhune, D. B., & Smith, M. D. (2006). The induction of anomalous experiences in a mirror-gazing facility: Suggestion, cognitive perceptual personality traits, and phenomenological state effects. *Journal of Nervous and Mental Disease*, *194*, 415–421.

Terhune, D. B., Ventola, A., & Houran, J. (2007). An analysis of contextual variables and the incidence of photographic anomalies at an alleged haunt and a control site. *Journal of Scientific Exploration*, *21*, 99–120.

Thalbourne, M. A. (2010). Transliminality: A fundamental mechanism in psychology and parapsychology. *Australian Journal of Parapsychology*, *10*, 70–81.

Tobacyk, J. J. (2004). A revised paranormal belief scale. *International Journal of Transpersonal Studies*, *23*, 94–99.

Tobacyk, J. J., & Milford, G. (1983). Belief in paranormal phenomena: Assessment instrument development and implications for personality functioning. *Journal of Personality and Social Psychology*, *44*, 1029–1037.

Williams, B., Ventola, A., & Wilson, M. (2010). *Apparitional experiences: A primer on parapsychological research and perspectives*. Retrieved from http://www.publicparapsychology.org/Public%20Parapsych/Apparitional%20Experiences%20Primer%20Final.pdf.

Wilson, K., & French, C. C. (2008). Misinformation effects for psychic readings and belief in the paranormal. *Imagination, Cognition and Personality*, *28*, 155–171.

Wilson, K., & French, C. C. (2014). Magic and memory: Using conjuring to explore the effects of suggestion, social influence, and paranormal belief on eyewitness testimony for an ostensibly paranormal event. *Frontiers in Psychology*, *5*:1289, 1–9.

Wiseman, R., & Greening, E. (2005). "It's still bending": Verbal suggestion and alleged psychokinetic metal bending. *British Journal of Psychology*, *96*, 115–127.

Wiseman, R., Greening, E., & Smith, M. (2003). Belief in the paranormal and suggestion in the séance room. *British Journal of Psychology*, *94*, 285–297.

Wiseman, R., Watt, C., Greening, E., Stevens, P., & O'Keefe, C. (2002). An investigation into the alleged haunting of Hampton Court Palace. Psychological variables and magnetic fields. *Journal of Parapsychology*, *66*, 387–408.

Acknowledgements

The authors would like to thank the Bial Foundation (Fellowship Programme 39/10) for financial support and anonymous referees for providing helpful suggestions.

Manchester Metropolitan University
Department of Psychology
Faculty of Health, Psychology and Social Care
Birley Building, 53 Bonsall Street
Manchester, UK M15 6GX
n.dagnall@mmu.ac.uk

Abstracts in Other Languages

French

SUGGESTION, CROYANCE AU PARANORMAL, DEFICITS DANS L'INCLINATION
A TESTER LA REALITE ET PERCEPTIONS DANS UN IMMEUBLE REPUTÉ HANTÉ

RÉSUMÉ : La présente étude cherche à savoir si la suggestion, le niveau de croyance au paranormal et une inclination à tester la réalité de manière défectueuse influencent les attentes des participants dans le cas de phénomènes relatifs à la hantise. On présente aux participants un court diaporama sur l'histoire fictive d'un hôpital abandonné. La suggestion est faite lors de la dernière phrase de la présentation et consiste à affirmer que l'immeuble administratif de l'hôpital a soit connu des activités de hantise, soit présenté des problèmes structuraux. Suite au diaporama, pour s'assurer que les participants ont suivi la suggestion, ils lisent une transcription de la présentation. L'expérimentateur informe ensuite les participants qu'ils vont regarder l'intérieur du bâtiment administratif grâce à une vidéo touristique sans son et en noir et blanc. A la fin de la séquence vidéo, les participants remplissaient des échelles évaluant les perceptions et les phénomènes dans l'environnement, les opinions et sensations relatives à la hantise, la croyance au paranormal (Echelle révisée de croyance au paranormal), et l'inclination à tester la réalité de manière défectueuse (Inventaire d'organisation de la personnalité). Lors de la phase expérimentale, seuls les niveaux de croyance au paranormal et l'inclination à tester la réalité de manière défectueuse ont affecté les mesures relatives à la hantise ; la suggestion n'a pas eu d'effet. La seconde phase d'enquête, employant l'analyse des causes, a révélé que l'histoire de hantise (le degré de conviction des participants dans le fait que l'immeuble administratif avait véritablement été hanté) médiatisait la relation entre la croyance au paranormal et l'attente des phénomènes relatifs à la hantise.

German

SUGGESTION, GLAUBE ANS PARANORMALE, NEIGUNG ZU DEFIZITEN IN DER REALITÄTSÜBERPRÜFUNG UND WAHRNEHMUNG EINES ANGEBLICHEN SPUKGEBÄUDES

Die vorliegende Studie untersuchte, ob eine gegebene Suggestion, die Stärke des Glaubens ans Paranormale und die Neigung zu Defiziten in der Realitätsüberprüfung die Erwartung von Teilnehmern in Bezug auf spukähnliche Phänomene beeinflussten. Die Teilnehmer wurde eine kurze Diashow präsentiert, in der die Geschichte eines fiktiven, verlassenen Krankenhauses vorgestellt wurde. Die Suggestion wurde im letzten Satz der Erzählung gegeben, in dem behauptet wurde, dass das Verwaltungsgebäude des Krankenhauses entweder eine Vergangenheit mit Geistererscheinungen oder bauliche Mängel aufwiese. Um sicher zu gehen, dass die Teilnehmer die Suggestion mitbekommen hatten, lasen sie nach der Diashow ein Transkript der vorgestellten Geschichte. Anschliessend informierte der Versuchsleiter die Teilnehmer, dass sie das Innere des Verwaltungsgebäudes mittels einer tonlosen Videotour in Schwarzweiss zu sehen bekämen. Nach Abschluss der Filmsequenz füllten die Teilnehmer Fragebögen zur Erfassung ihrer jeweiligen Umgebungswahrnehmungen und Phänomenen, ihrer Meinungen und Gefühle in Bezug auf Spukhaftes, ihren Glauben ans Paranormale (Revised Paranormal Belief Scale) und ihrer Neigung zu Defiziten in der Realitätsüberprüfung (Inventory of Personality Organization) aus. Innerhalb der Experimentalphase beeinflussten nur die Stärke des Glaubens ans Paranormale und die Neigung zu Defiziten in der Realitätsüberprüfung die auf Spuk bezogenen Einschätzungen; die Suggestion selbst blieb wirkungslos. Eine Auswertung in der zweiten Phase mittels einer Pfadanalyse ergab, dass die Spukgeschichte (das Ausmaß, zu dem die Teilnehmer glaubten, dass das Verwaltungsgebäude im Ruf stünde, dass es in ihm spuke) die Beziehung zwischen dem Glauben ans Paranormale und der Erwartung von auf Spuk bezogenen Phänomenen vermittelte.

Spanish

SUGESTIÓN, CREENCIA EN LO PARANORMAL, PROPENSIÓN A DÉFICITS EN LA PRUEBA DE REALIDAD Y PERCEPCIÓN DE UN EDIFICIO SUPUESTAMENTE ENCANTADO

RESUMEN: Este estudio investigó si la sugestión, el nivel de creencia en lo paranormal, y la propensión a déficits en la prueba de realidad influyeron en las expectativas de participantes en fenómenos relacionados con casas encantadas. Los participantes vieron una breve presentación de diapositivas que resumía la historia ficticia de un hospital abandonado. La sugestión ocurrió en la última frase de la narración en donde se mencionó que el edificio administrativo del hospital tenía una historia de actividad fantasmal o problemas estructurales. Después de la presentación de diapositivas, para asegurar que los participantes habían prestado atención a la sugestión, leyeron una transcripción de la presentación. El experimentador le dijo entonces a los participantes que iban a ver las características internas del edificio administrativo a través de una visita guiada en video en blanco y negro y sin sonido. Al concluir el video los participantes contestaron cuestionarios para medir percepciones y fenómenos ambientales, opiniones y sentimientos relacionados con las casas encantadas, creencia en lo paranormal (Revised Paranormal Belief Scale), y propensión a déficits en la prueba de realidad (Inventory of Personality Organization). Dentro de la fase experimental, sólo el nivel de creencia paranormal y la propensión al déficit de prueba de realidad afectaron las puntuaciones relacionadas a la casa encantada; la sugestión no tuvo ningún efecto. Una segunda fase de evaluación, utilizando análisis de rutas, reveló que la historia sobre la casa encantada (el grado en que los participantes creyeron que el edificio administrativo tenía una historia de ser casa encantada) medió la relación entre la creencia paranormal y la expectativa de los fenómenos relacionados con el encantamiento.

Journal of Parapsychology, 79(1), 105–107

RESEARCH NOTE

POSTSCRIPT TO WATT (2014) ON PRECOGNITIVE DREAMING: INVESTIGATING ANOMALOUS COGNITION AND PSYCHOLOGICAL FACTORS

By Caroline Watt and Milan Valášek

ABSTRACT. Mörck and Hansen have questioned whether the results of Watt (2014) could be affected by unreported data from participants who dropped out of the study. We welcome this observation and present the results of these unreported trials. We also compare the outcome of the unreported trials with those reported in Watt (2014) and find a significant difference suggesting that the hit rate in Watt (2014) is inflated due to the omission of the dropouts' data.

Keywords: precognitive dreaming, filedrawer effect, precognition

Our online dream precognition study recently reported in the *JP* (Watt, 2014) obtained overall significant results on the planned analysis: 64 hits in 200 trials (32% hit rate where MCE = 25%, exact binomial $p = .015$, one-tailed, effect size (ES) = $z/N^{1/2} = 0.16$. However, as was pointed out in the Discussion, post hoc analyses did not support an interpretation of these results in terms of dream precognition on the part of our participants. For example, the independent judges' ratings of targets were not significantly higher than their ratings of decoys. So the excess hits in the study did not appear to be attributable to the participants' dreams resembling the targets more than they resembled the decoys.

Recently in a Letter to the Editor in the *JP*, Mörck (2015) helpfully raised an issue that George Hansen brought to his attention. This point may help to disambiguate the original study findings. As planned, the end of the study was defined as the first 50 participants to complete four trials (= 200 trials), and the data from any participants who did not complete four trials were discarded. There were 10 participants in the latter group (another 11 did not complete any trials). Statistically, this is not necessarily a problem if participants' departure from the study is random. However, Mörck and Hansen raised the question of whether there could be some pattern to their departure that might have influenced the hit rate on the remaining 200 trials. This would be a form of filedrawer effect that could either raise or lower the published hit rate, contributing to a type I or type II error, respectively. Although participants might depart a study for numerous reasons unrelated to the study itself, one factor that *could* influence participants' decision to depart or stay with the study is their perception of their performance on the precognition task.

Although participants did not receive trial-by-trial feedback on the judges' ratings of their trials, participants *were* given feedback in the form of viewing the target video clip after they had submitted their dream reports. This is because in a precognition study participants' task is to dream about the target clip they will later view. It is therefore possible that participants' continuation in the study might have been influenced by their perception of whether or not their dream predicted the target clip. This could occur in two ways.

First, in line with Tart's (1984) "fear of psi" concept, some participants may have been alarmed if they perceived a strong correspondence between their dreams and the target clip. After all, unlike a telepathy study, in a precognition experiment there is no sender with whom to share responsibility for any seemingly psychic ability. If so, frightened participants might have dropped out of the study, and the independent judges might have judged these trials to be hits. This "frightened by early success" hypothesis would have the effect of reducing the overall hit rate of the study because hits are being discarded, contributing to type II error.

Alternatively, it is possible that some participants may have dropped out because they were disappointed that their dreams did not seem to resemble the target clip. Let us assume on this "discouraged by early failure" hypothesis that independent judges scored these trials as misses. This trend would have the effect of artifactually inflating the overall hit rate of the study because misses are being discarded, contributing to type I error. Since we wish to understand the significant overall hit rate, it makes sense to test this discouraged by early failure hypothesis. We therefore present a post hoc analysis of the scoring of those 10 participants who dropped out of the study before completing four trials, to see whether it is consistent with this discouraged by early failure hypothesis.

The 10 incomplete data participants scored just three hits in 19 trials (15.8% hit rate; exact binomial $p = .53$, two-tailed). Combining their data with that of the 50 participants who each completed four trials gives 67 hits out of 219 trials (30.6% hit rate, exact binomial $p = .04$, one-tailed, ES = 0.12). (A one-tailed p value is reported for this combined analysis because the original planned analysis was one-tailed). The difference between the hit rates (number of hits/number of trials completed per participant) of the 10 incomplete data participants and the 50 completed is significant (Mann-Whitney $U = 149.5$, $p = .04$, two-tailed). These analyses therefore support the discouraged by early failure hypothesis. The combined analysis remains significant by a one-tailed test, so by our planned analysis the psi hypothesis is still supported. But the outcome is no longer significant by a two-tailed test, and the effect size is lower, indicating that the originally reported hit rate is inflated by the omission of the dropouts' data. These further analyses do not change our original assessment that it is difficult to account for the study results in terms of our participants dreaming precognitively.

Seven participants in our experiment completed four trials after the planned criterion for ending the study was reached. Their data were therefore not included in the analyses reported in Watt (2014). Although their data are not relevant for the disappointed by early failure hypothesis, we report them here to bring them out of the file drawer. They obtained five hits out of 28 trials (17.9% hit rate, exact binomial $p = .51$, two-tailed).

Our original *JP* manuscript was submitted in August 2012 and went through four rounds of reviewing and revision before it was accepted for publication in April 2014. This account shows that even with such scrutiny, potential artefacts can be missed, and parapsychologists need to be continually on their guard against both type I and type II errors.

References

Mörck, N. C. (2015). [Letter to the Editor]. *Journal of Parapsychology*, *79*(1)**,** 127.

Tart, C. T. (1984). Acknowledging and dealing with the fear of psi. *Journal of the American Society for Psychical Research, 78*, 133–143.

Watt, C. (2014). Precognitive dreaming: Investigating anomalous cognition and psychological factors. *Journal of Parapsychology, 78*, 115–125.

Psychology Department
University of Edinburgh
7 George Square
Edinburgh EH8 9JZ, Scotland
caroline.watt@ed.ac.uk

Abstracts in Other Languages

French

POSTCRIPT À WATT (2014) SUR LA REVERIE PRECOGNITIVE :
ETUDE DE LA COGNITION ANOMALE ET DES FACTEURS PSYCHOLOGIQUES

RÉSUMÉ : Mörck et Hansen se sont demandés si les résultats de Watt (2014) pouvaient être affectés par

des données non-reportées par les participants qui ont quitté l'étude en cours de route. Nous recevons cette observation et nous présentons les résultats de ces essais non-reportés. Nous comparons aussi les essais non-reportés avec ceux reportés dans Watt (2014) et trouvons une différence significative suggérant que le taux de succès dans Watt (2014) est exagéré du fait de l'omission des données mises de côté.

German

POSTSKRIPT ÜBER PRÄKOGNITIVE TRÄUME (WATT 2014):
ZUR UNTERSUCHUNG ANOMALER KOGNITION UND PSYCHOLOGISCHER FAKTOREN

ZUSAMMENFASSUNG: Mörck und Hansen haben die Frage aufgeworfen, ob die Ergebnisse von Watt (2014) nicht durch Daten von Teilnehmern hätten beeinflusst werden können, die nicht berichtet wurden, weil sie aus der Studie vorzeitig ausgeschieden sind. Wir begrüssen diese Beobachtung und präsentieren die Ergebnisse dieser nicht berichteten Versuche. Wir vergleichen auch die Ergebnisse der nicht berichteten mit denjenigen bei Watt (2014) und finden eine signifikante Differenz, die darauf hindeutet, dass die Trefferrate bei Watt (2014) als überhöht angegeben wurde, weil die Daten der vorzeitig ausgeschiedenen Teilnehmern nicht berücksichtigt wurden.

Spanish

POSDATA A WATT (2014) SOBRE SUEÑOS PREMONITORIOS:
INVESTIGANDO A LA COGNICIÓN ANÓMALA Y LOS FACTORES PSICOLÓGICOS

RESUMEN: Mörck y Hansen han cuestionado si los resultados de Watt (2014) podrían haber sido afectados por los datos no reportados de los participantes que dejaron prematuramente el estudio. Agradecemos esta observación y presentamos los resultados de esos datos no reportados. También comparamos los resultados de los datos no comunicados con los mencionados en Watt (2014) y encontramos una diferencia significativa que sugiere que la tasa de éxito en Watt (2014) se infló debido a la omisión de los datos de abandono.

Journal of Parapsychology, 79(1), 108–118

BOOK REVIEWS

ANOMALOUS COGNITION: REMOTE VIEWING RESEARCH AND THEORY by Edwin C. May & Sonali B. Marwaha. Jefferson, NC: McFarland, 2014. Pp. vi + 434. $45.00 (paperback). ISBN 978-864-9458-3.

Background

Spying is one of the oldest professions and if psi exists it could be an invaluable tool: If assassination by PK (or anomalous perturbation—AP) proves impractical, then ESP (or anomalous cognition—AC) can be used indirectly to discover enemy secrets (and *then* clobber them)—or to predict future events. This idea was likely first implemented by shamans and later by an impressive array of warlords, right through classical antiquity up to modern times.

Some within the CIA clearly followed parapsychology with Argus eyes and the notorious MKUltra project, initiated in 1953, included parapsychology experiments. Later J. B. Rhine (1957) wrote: ". . . modern war, even more than ever before, hinges mainly upon secrecy and surprise. If these two dangers could be banished, a long lease on world peace could be gained . . ." (p. 249). Harold Puthoff and Russell Targ apparently agreed and in 1972 initiated what is now known as the Star Gate program, which focused on remote viewing (RV): Edwin May joined them in 1975.

Contents

Now 40 years later May, together with colleague Sonali Marwaha, has put together a compendium of articles which he authored or coauthored. This anthology is decidedly not another glossy "salon table" book on the wonders of psi but is for the most part previously published technical work on RV. As Broughton notes in his introduction, May's is a decidedly "no-nonsense" approach: The reader will look in vain for metaphysical disquisitions. The going is quite tough, with statistics, equations and all. It comprises 26 papers, divided into six sections: Methods, (Electro-)Physiology, Decision Augmentation Theory, Entropy, Miscellaneous, and Challenges.

The seven papers in the Methods section comprise a self-contained guide to RV protocols, target construction, and analysis techniques all conveniently gathered together in one place. The book is worth having for this alone; but it is pointless to try to condense this technical material into a few words. Particularly worthy of attention is the novel "figure of merit," based on fuzzy sets. This is basically an attempt to "semi-automate" the tedious judging process involved in rating and ranking.

The last paper in this section is of more than purely methodological interest: This is an attempt to determine whether feedback to the remote viewer (RVer) is important for success. Feedback was given over a range of tachistoscopic intensities: Unfortunately only two of the four RVers were able to score under these conditions, so the experiment effectively collapsed to case studies. The results were disparate: For one of them there was a positive but nonsignificant relationship between feedback intensity and AC success, while for the other RVer this was significantly *negative*. We are told (p. 110) that "All receivers believed strongly that feedback was not necessary for success, but they were uncertain about the degree to which feedback might contribute to success." What is not reported is the kind of thing which would have been examined in a planned case study, for example, did this *particular* RVer strongly disbelieve in the relevance of feedback? I have proposed elsewhere (Millar, in press) that a simplified form of this theoretically important experiment should be performed on a large scale by experimenters and participants with diverse expectations.

As an old electrode buff, the six papers in the Physiology section draw my particular attention. May's team has looked rather widely (though thinly) at physiological responses instead of the usual con-

scious responses: They used measures ranging from skin conductance to EEG alpha power to state-of-the-art magneto-encephalography. Some promising results were registered. The best (p. 152), in a presentiment experiment with auditory stimuli, yielded an over five standard deviation difference between experimental and control conditions. However, while the skin response measure went up for experimental trials, it went down for the controls. This is of little utility to the participant but does serve the experimenter's goal of getting a significant difference. I have proposed elsewhere (Millar, 2012) that such "mirroring" is a "fingerprint" for experimenter psi (epsi). In the following study (pp. 158–171) more direct evidence was found for an epsi mechanism: Here the difference was primarily due to a negative deviation in the control trials. It seems that the (so-called) participant's only function was that of random skin conductance generator!

Review Plan

If parapsychology were normal science, these first two sections, which account for more than half of the book, would contain most of the "meat," definite facts learned about the nature of psi. Investigators of simple physical systems usually have little difficulty in replicating initial results and gradually refine them to more and more decimal places. Parapsychology is just not like that: Typically one experimenter gets result A, while another reports B and never the twain shall meet.

From this point the order of the book is followed less closely: It is approached instead via a number of themes. Because of its nature as an anthology of May's work, the book systematically lacks context. Here it is examined within the background of parapsychology in general. First on the order are theories: Decision augmentation theory and entropy are examined within a historical framework. With these under the belt, the underlying question is tackled: Does RV really work and, if so, is it better than other methods? A critical topic, endemic to parapsychology as a whole, is then taken up: Where or from whom does the psi registered come? Some words are reserved for the promotional aspects of the volume.

Decision Augmentation Theory (DAT)

The "jewel in the lotus" of this book is undoubtedly decision augmentation theory, a phenomenological/mathematical account of psi, which May describes as the earlier psychological psi-mediated instrumental response theory of Stanford (1974) with a math topping: The core is that psi is usually quite unconscious and goes on automatically all the time to satisfy a "need" (or "disposition"). Phenomenal representation, such as having a precognitive dream or seeing an apparition, is quite unusual. DAT is based on precognition. Psi habitually acts on all the trivial choices of everyday life so as to result in a "good" future and avoid the "bad." This frequently involves unconscious precognition-guided timing. For example, Alice wakes up late one morning and misses her usual train, which subsequently crashes into a chemical tanker, crisply toasting all aboard.

The psi failure of the prematurely cremated passengers is *not* accounted for by DAT: The obvious suggestion, though, is that the unfortunate majority just didn't have enough psi—and this is surely May's answer (though it may not be welcomed by Stanford). In any case, it is likely true that most psi is unconscious and little under the control of the percipient/agent—in Rhine's terminology psi is only marginally "dirigible."

DAT applies, in principle, to *any* study in which statistical inference is used and arbitrary (timing) decisions are made by subject or experimenter: Psi enters at just these points. Of importance here is how DAT applies to RNG-PK experiments. In this theory there is no such thing as PK, only precognition: the RNG machine is switched on at just the right moment to catch a chance deviation which is "good" (nominally significant). DAT pseudo-PK is actually unconscious precognition-based selection of the "good seed" with every timing decision.

But is it true? If psi selects out the good seed from a purely chance set, it follows that the "bad seed" is left behind. DAT theorists have looked at the experimental data (good seed) but have not examined whether the positive deviation found is matched by an equal negative deviation in the "control" data. In

fact, these residue/control data have often been potential rather than real. Needed are experiments in which the total set (say 10,000 randomly pre-generated sessions) is actually pregenerated and recorded before separating the data into experimental and control: in this way the *full* prediction of DAT (for both experimental *and control*) can be put to the test.

A compensating negative deviation in nonobserved data may be expected rather generally for *any* theory that involves *selection* of data sets from a larger "random" set, by whatever mechanism, rather than being specifically characteristic for DAT.

To this reviewer it seems a chicken and egg question whether the basic phenomenon is precognition or (retro-)PK because both involve an influence of the relative future on the relative past. DAT fails to give a clear account of this coupling: Precognition for May seems to be regarded as something basic and unanalyzable. Physically, however, it must be nonlocal in nature, because a *signal* cannot propagate backwards in time. May asserts (personal communication, 2015) "there is NO evidence of retro-PK in physics": (If this were true) it impacts DAT with equal force because future/past coupling is the same irrespective of the verbal nicety of whether it is called retro-PK or precognition.

DAT pseudo-PK superficially looks like "ordinary" PK. A little bit of math, though, reveals that the DAT version results in a *constant z score*, independent of the number of trials: *an experiment with a few trials (and/or participants, etc.) is as good as a big one!* By contrast, it is well known in psychology that for any trial-by-trial effect, the *z* score increases with the square root (Sqrt) of the number (*N*) of trials. The startling counter-intuitive DAT result comes about because the session (or experiment) is influenced *as a whole*: Only if psi functions on a trial-by-trial level is Sqrt(*N*) dependence to be expected. The apparent paradox is expected also from any other model whatsoever in which the data are psi-influenced together as a single composite unit.

Singled out is the global consciousness project (GCP; pp. 268–277): It is instructive to compare Nelson's (2015) "world-mind friendly" account with the treatment by Bancel (2015) in the same volume. GCP seemed to follow a constant *z* rather than Sqrt(*N*) dependence, and this was interpreted by May et al. as likely epsi (Nelson-psi). However, Bancel calculated that the expected regression slope is so small that the empirical data are statistically consistent with both models. DAT has not (yet) notched up a decisive victory on Sqrt(*N*) grounds. Other evidence May et al. advance, however, seems more persuasive: The GCP events proposed by Nelson himself have significantly higher scoring than those contributed by others: Nelson himself was forced to suggest he may just know by experience what kind of events give the best scoring. Influentially, Bancel (2015) has recanted his earlier support for a real effect of global consciousness and comes out (guardedly) in favor of epsi.

A major limitation of the presentation here is that it gives the impression DAT is entirely de novo: a good theory-mother just knows his own brain-child is unique. The historical context and relationship to other theories is not explored. DAT has much (but by no means all) in common with the observational theories (OTs) and they have both been treated (Millar, in press) as non-local-in-time theories (NLTs).

In the quantum world systems can exist in several states *at the same time* (superposition), quite unlike the macroscopic world of experience. In the early stages of quantum theory it was even less clear than now where the quantum description ends and classical physics takes over. To dramatize this, Schrödinger imagined a cat in a box together with an apparatus based on radioactive decay, which either killed it or not, with a 50% chance of each: He reasoned that the cat should then exist in a superposition, both living and dead at the same time, until someone opened the box and "collapsed" the state vector to one definite event or the other.

Walker seized on this for one of the several (incompatible?) mechanisms for psi he proposed, "Walker's Alice", if you like. On observing a random generator output and wishing for "1", Alice enters a superposition in which she "sees" both 0 and 1 simultaneously: She goes around in this superposition (which naturally passes unnoticed) until she (or her density matrix) is later collapsed back to ordinary billiard-ball reality, but with the desired *1 in her memory:* All the rest of the world "plays along" as if a real (classical) 1 had been generated in the first instance.

If this sounds like Alice in Wonderland – it is. The majority of (European) parapsychologists in-

volved with OT rapidly chucked this notion like a hand grenade: Only Bierman (perhaps) followed Walker here. Schmidt, ever the gentleman, was silent about Walker's Alice in public. In private he dismissed it as nonsense: Elementary calculation shows that an Alice superposition could not exist for more than about a millionth of a microsecond. However, May took Walker's mechanism seriously enough to subject it to experiment (pp. 339–350), which lent it no support: He could have saved himself a great deal of trouble.

If this was May's idea of what the OTs are all about, it is no wonder he developed a theory of his own, first the intuitive data sorting (IDS) model, which was succeeded by DAT. In the continuing course of development they have recently (Marwaha & May, 2015) started sketching an updated version—the multiphasic model of precognition (MMPC).

Entropy

Another intriguing section deals with the relation between entropy and psi. There are two kinds of entropy: physical entropy and informational (Shannon) entropy. Physical entropy was first defined purely in terms of heat and temperature, progeny of the steam engine. It later turned out that it is a general measure of organizational disorder: For any (macroscopic) closed physical system entropy always increases (or stays the same).

Shannon entropy was inspired by the telegraph. A communication system must be designed to cater for not just a single message but the whole set of *possible* messages. Shannon characterized an information source by a measure of the available choice, Shannon entropy. This has the same mathematical form as physical entropy but is *not* restricted to any particular physical system. Information (or communication) theory (IT) is not a *substantive* theory: It is a mathematical method. IT may be described as a theory of correlation for categorical data. The beauty of IT is how it expresses things in neat (logarithmic) terms that have some connection to intuitive ideas about information. This correspondence is far from complete: In particular, there is no (direct) reference to "meaning." IT may prove to be a fruitful method for parapsychology: Already various measures of ES are in use, which are actually (implicit) measures of information.

The first mention of physical entropy in parapsychology was long ago when it was still psychical research. In the context of hauntings and poltergeists, it was conjectured that maybe the energy for the manifestation was drawn from the surrounding air, leaving "cold breezes" behind, in violation of the second law of thermodynamics. Critics found this a "load of hot air." Donald and Martin (1975) much later presented a formal observational theory (OT) based on statistical thermodynamics. These ideas are conceptual forerunners of DAT's psi-separation of the good and bad seed and the antientropic nature of psi.

In the course of their RV experiments May and colleagues garnered the impression that target events that involve conspicuous physical-entropy change, such as rocket launching, are relatively successful. They then turned to informational entropy (a different beast) in target pictures and found that RV scoring related to what they call the gradient of the Shannon entropy: This is a measure of the diversity within each picture. How is this entropy gradient effect to be explained? Perhaps it has nothing to do with entropy as such. If high entropy-gradient pictures stand out visually then the experiment becomes a (no tachistoscope) test of the importance of sensory feedback "in disguise" (see "Contents" section above).

Importantly they suggest an ingenious equation that relates maximum possible psi scoring to (informational) entropy change: This really warrants a separate review too extensive to incorporate here.

The authors must be lauded for drawing attention to thermodynamic and informational entropies, topics which have largely been neglected in parapsychology; however, it seems to this reviewer that here once more, their own contribution is presented as unique and they fail to set it adequately within the historical background.

Does RV Really "Wash Whiter"?

Although the title "Anomalous Cognition" is likely to attract only those with at least a modicum of curiosity about the field, the readership targeted is primarily the mainstream technical scientist. The work is pitched in a way that is familiar to the researcher: A few may be intrigued enough to try it out for themselves.

Will those scientists who attempt RV experiments get similar overwhelmingly positive results? My own crystal ball on RV tells me that only a few will get convincing outcomes: I very much hope I am wrong. Few current parapsychologists, least of all May, would be so naive as to maintain that psi-scoring can be obtained by simply following a cookbook. Time alone will tell how many experimenters can produce a gourmet meal with his book as guide.

Evaluation of published evidence is influenced to an important (and usually unacknowledged) degree by one's own experience: This is true both for experimenters who obtain substantially positive results and those who do not. Near the end of his long (near uniformly negative) career in experimental parapsychology, Beloff (1994) summarized the field as a "succession of false dawns and frustrated hopes" (p. 7). Like sentiments have been expressed by pillars of parapsychology from Pratt to Rao.

It seems that psi-missing, declines, displacements and like effects that have historically given psi its characteristically elusive character have *not* manifested in May's experience. He emphasizes (p. 21) that no decline is to be seen in his RV data and he has never yet seen psi-missing. However, in one paper addressing the effect of sidereal time on AC scoring (pp. 377–386), highly significant effects were found, which in later work (Spottiswoode, 2015) have (practically) disappeared. Other RV researchers, too, have not experienced the same "plain sailing" as May. For example, when Targ attempted to develop a practical application of RV, psi-hitting capriciously switched to psi-missing (Targ, Braud, Stanford, Schlitz, & Honorton, 1991, pp. 76–77).

Has May discovered the long sought secret of banishing decline, or is it waiting just around the corner? According to Kennedy (2003), a new psi line typically starts with a bang (Beloff's false dawn): Thereafter, scoring of selected experimenters may remain stable for some time but becomes unreliable when they attempt to clinch it, either by formal studies or to apply the effects. Finally it fades away to a limbo, which yields neither convincing evidence for psi nor its absence. The practical application that would provide a compelling common-sense case for psi still remains a "frustrated hope," though persistent attempts continue to be made.

RV is basically "guided imagination," an association exercise steered by the demand characteristics of the experiment. While handbooks have been written on how to do RV and commercial training programs are offered, it is by no means clear that any of this is of real value (Utts, 1996). While opinions in the RV community are mixed, Targ considers that one can learn all one needs to know about RV in 15 minutes.

There does not seem to be anything very specific that distinguishes RV from a plethora of other methods, and it is interesting to explore what elements might account for its success (see also the interesting psychological speculations by McMoneagle and May. pp. 368–376). The most visible difference from more standard psi experiments is that RVers are systematically selected on the basis of past success. Many alternative methods and measures were tried in an effort to forecast RV performance, but none was found to be useful in practice. This is singular, given the manifold claims for substantial correlation of psi with inter alia assorted personality measures.

The best available predictor of psi-scoring, according to May et al., is previous psi scoring: Dumas Père long ago noted "Rien ne réussit comme le succès" (Nothing succeeds like success). RV contrasts strongly with the standard approach, which currently almost eschews subject selection. If the standard recipe begins "Take 20 psychology students," then May agrees with Mrs. Beaton: "First catch your chicken"—a strategy earlier promoted by Honorton.

The conventional approach only makes sense if psi is as widely distributed as most psychological variables, and this was (and still is) the standard assumption. An alternative is the traditional (ecological) belief in the special powers of saints and witches versus the rest of humanity (HAVEs and (nearly) have-NOTs). May himself pays lip service to ". . . the normal continuum of human perceptual abilities . . ." but adds ". . . there are some who are AC gifted, much like giftedness in music, sports . . ." (p. 4), and he makes clear his preference that parapsychology should go for these "biggies." One estimate (Utts, 1996) is that 1% of the general population is fairly successful at RV.

It is to be hoped that May's approach will stimulate the renaissance of empirical investigations into the actual distribution of psi, a topic which has been nearly taboo since the '80s (Millar, 1979). A hopeful bode is that Dobyns (2015, p. 233) has analyzed the PEAR REG database and found the best fit indicates

15% constitute a "talented subpopulation" that was responsible for the psi. Apparently, however, this does not fit their RV data.

Does RV actually offer a higher yield of psi than other methods? The luminous impression from reading selected excerpts is that there is something much more powerful going on here than in the typical psi experiment. There is reason to doubt this: Utts (1996) compared the performance of RV (at SRI) to ganzfeld (at PRL): The effect sizes (ESs) for novices are nearly identical (0.16 vs. 0.17), as are the ESs of experienced participants (0.39 vs. 0.35). It looks as if ping-pong balls may be as effective as RV. The (rather poorly specified) selection procedures employed seem to be about equally good in both cases: ES is near doubled, which corresponds to a major (4-fold) increase of information per trial.

Free response trials typically take much more time than card guessing, so the standard ES (per trial) is inappropriate for comparing the two. Nelson (2006) suggests instead ES per unit time, *z*/Sqrt(hours). In a rapid search, I was unable to find any formal comparison of RV studies versus forced-response using his measure. However, if an RV trial takes 30 min and a pack of 25 cards 5 min, the time ratio per trial is about 150 to 1 and the corresponding Nelson-ES ratio is the square root of that, roughly 12 to 1. The empirical ratio of ESs (RV/cards) is perhaps 0.16/0.015 or 11 to 1—nearly the same. This "back of the envelope" estimate suggests that RV may not, in fact, be superior to card-guessing in terms of invested person-hours.

Whodunnit?

The enormous experimenter involvement in RV fairly jumps off the page. One can hardly speak of separate experimenter and participant: It is more like a symbiotic pair. This is very unlike the usual situation in psi experiments, where the participant is "on his own" and the experimenter has a relatively aloof supervisory role. The RV setup seems likely to maximize both psychological and psi contributions due to the experimenter. The authors warn that experimenters should not run too frequent sessions (even with different RVers) or results deteriorate: This is a "warning" sign that the experimenter may be an essential part of the psi system. Already (see above) there are clear suggestions of epsi in the electrophysiology experiments.

This is fine if all that is wanted is to demonstrate that "sometimes psi happens"; but what if RV success is dependent on the (exceptional) person of May and his results are mere "selfies"? He recognizes the problem, which is inherent in his cherished theory (DAT), but discussion is largely relegated to a short section "Ideologically-driven Experimenter Expectancy Effects" (p. 394), in a paper on "Research Challenges" tucked away near the end of the book. It is almost as if May suspects he is an essential part of his experiments but promotes RV anyway because it is his own particular working "ritual." He explicitly proposes (see below) that another successful researcher, Nelson, plays just this kind of role in his own global consciousness project (GCP). *Is parapsychology an "experimenter-specific science"?* May may be right in thinking the first priority is to get a psi effect any which way and only later to narrow down where it is coming from.

The million dollar question: Is it the RV method that works or is it the experimenter/lab? May (and Targ too) has a track record of successful experiments other than RV—strike one against the method. But some of the RVers involved, notably McMoneagle, have scored independently (one all). *I must admit to an intense curiosity whether such RV stars would continue to shine with me as experimenter.*

What have other people made of RV? Baptista, Derakhshani, and Tressoldi (in press) provide a valuable comprehensive summary, though approached exclusively from the traditional "participants dunnit" viewpoint. Schwartz (2015) fills in much fascinating background and his much more positive evaluation (partly based on his own experiments) provides a counterpoint to my own. Apart from the SRI International (and SAIC) work, the major investigation of RV was done by members of the PEAR group, Dunne (and nominally Jahn). This was marred by a pair of severe methodological errors. Hansen, Utts, and Markwick (1992) remark: "The research departs from criteria usually expected in formal scientific experimentation" and they continue: "Many of the issues of remote-viewing methodology were identified by Stokes and Kennedy over 10 years ago" (p. 97). Baptista et al. agree, though they express it more diplomatically. Evidence for RV consequently depends quite heavily on May's lab. The odds, I wager, remain even.

Polemics and Politics

The banner headline of May's introduction screams: "Science Has Demonstrated That ESP Is Real!" If "science" has demonstrated psi "in spades" to May, she has, in contrast, dealt me a distinctly "thrifty" hand. The role of the scientist is not that of evangelist, to push his own overwhelming experience down the throats of benighted souls: This leaves a sour taste. So long as gross and persistent experimenter differences exist, doubt about the reality of psi remains eminently reasonable.

Concluding Remarks

This book is the result of the personal dedication and determination of Ed May, who took the greatest advantage possible of a temporary window of public funding. Despite conditions not always conducive to science, his decades of application have produced a lasting monument. This is an exciting book with fresh ideas and it deserves a place on the bookshelves of every serious parapsychologist. *Anomalous Cognition* is a "must have" for the professional who intends to do RV experiments. It should, however, be borne in mind that the book is a "one man band" and May's results may be idiosyncratic. It is very much a preliminary sketch of diverse work that should be vigorously followed up by many others. *Anomalous Cognition* can also be recommended for the contemporary scientist interested in psi, though it should be read in conjunction with a more general account of parapsychology.

References

Bancel, P. A. (2015). An analysis of the global consciousness project. In D. Broderick & B. Goertzel (Eds.), *Evidence for psi* (pp. 255–277). Jefferson, NC: McFarland.

Baptista, J., Derakhshani, M., & Tressoldi, P. (in press). Explicit anomalous cognition: A review of the best evidence in ganzfeld, forced-choice, remote viewing and dream studies. In E. Cardena, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century.* Jefferson, NC: McFarland.

Beloff, J. (1994). Lessons of history. *Journal of the American Society for Psychical Research, 88*, 7–22.

Dobyns, Y. (2015). The PEAR laboratory: Explorations and observations. In D. Broderick. & B. Goertzel (Eds.), *Evidence for psi* (pp. 213–236). Jefferson, NC: McFarland.

Donald, J. A., & Martin, B. (1975). Time-symmetric thermodynamics and causality violation. *European Journal of Parapsychology, 1,* 17–36.

Hansen, G. P., Utts, J., & Markwick, B. (1992). Critique of the PEAR remote-viewing experiments, *Journal of Parapsychology, 56,* 97–113.

Kennedy, J. E. (2003). The capricious, actively evasive, unsustainable nature of psi: A summary and hypotheses. *Journal of Parapsychology, 67*, 53–74.

Marwaha, S. B., & May, E. C. (2015). Rethinking extrasensory perception: Toward a multiphasic model of precognition. *Sage Open*, Jan/Mar, 1–17. Retrieved from *http://sgo.sagepub.com/content/spsgo/5/1/2158244015576056.full.pdf*

Millar, B. (1979). The distribution of psi. *European Journal of Parapsychology, 3*, 78–110.

Millar, B. (2012). Towards a forensic parapsychology in the OT paradigm. *Journal of Nonlocality*, *1*. Retrieved from *http://journals.sfu.ca/jnonlocality/index.php/jnonlocality/article/view/16*

Millar, B. (in press). Quantum theory and parapsychology. In E. Cardena, J. Palmer, & D. Marcusson–Clavertz (Eds.), *Parapsychology: A handbook for the 21st century.* Jefferson, NC: McFarland.

Nelson, R. D. (2006). Time-normalized yield: A natural unit for effect size in anomalies experiments. *Journal of Scientific Exploration, 20,* 177–199.

Nelson, R. D. (2015). The global consciousness project: Subtle interconnections and correlations in random data. In D. Broderick. & B. Goertzel (Eds.), *Evidence for psi* (pp. 237–254). Jefferson, NC: McFarland.

Rhine, J. B. (1957). Why national defense overlooks parapsychology. *Journal of Parapsychology, 21*, 245–258.

Schwartz, S. A. (2015). Through time and space: The evidence for remote viewing. In D. Broderick. & B. Goertzel (Eds.), *Evidence for psi* (pp. 168–212). Jefferson, NC: McFarland.

Spottiswoode, J. (2015). Psi and the environment: Local sidereal time and geomagnetic effects. In D. Broderick & B. Goertzel (Eds.), *Evidence for psi* (pp. 278–290). Jefferson, NC: McFarland.

Stanford, R. G. (1974). An experimentally testable model for spontaneous psi events: I. Extrasensory events. *Journal of the American Society for Psychical Research, 68*, 34–57.

Targ, R., Braud, W. G., Stanford, R. G., Schlitz, M. J., & Honorton, C. (1991). Increasing psychic reliability. *Journal of Parapsychology, 55,* 59–83.

Utts, J. (1996). An assessment of the evidence for psychic functioning. *Journal of Scientific Exploration, 10*, 3–30.

Brian Millar

Denpassar, Bali
bmillar@ymail.com

ESP WARS by Edwin C. May, Victor Rubel, & Loyd Auerbach. Palo Alto, CA: Laboratories for Fundamental Research, 2014. Pp. xii + 323. $20 (paperback). ISBN 978-1500743000.

The literature concerning the psychic arms race between America and the Soviet Union is an intriguing mixture of fact and fiction. Many of the claims are bound to exceed boggle thresholds, but some seem more plausible. *ESP Wars* allegedly contains the true stories as narrated by some of the key players.

The book starts with a wide-ranging historical survey with accounts from ancient times, including stories about shamans, yogis, and saints. Alongside well-known mediums, there are also fascinating individuals such as Blavatsky and Rasputin who make an appearance. Messing does not appear, but Hanussen, "Prophet of the Third Reich," does, in a section concerning Nazi occultism, a subject which, like the named individuals, is itself surrounded by myths. When Hitler came to power he began to persecute those involved in the occult, and something similar occurred under Stalin's regime: "Books on occultism were removed from shelves all across the country, members of esoteric groups were sent to camps and shot" (pp. 44–45). Officially, in the late 1930s occultism and psi research in the Soviet Union ceased to exist.

The history of psi research in the West and the East is also outlined. Oddly, neither the declassified reports nor the accounts by parapsychologists who had been in the Soviet Union are cited or even mentioned. Project MKULTRA, initiated by the CIA in 1953, is briefly covered. According to the authors, the CIA's interest in psi was due to its ". . . potential both as a means of mental manipulation and as a method of covertly sending and receiving information" (p. 53).

The well-known story about the telepathy experiment involving the U.S. submarine Nautilus, which allegedly took place in 1959, is naturally also covered. The authors appear oblivious to Martin Ebon's (1983) investigation: in short, this suggests it to have been a hoax by the author Jacques Bergier, who may have been deliberately fed disinformation. The story did function as an ignition spark for the psychic arms race—Soviet parapsychologists successfully used it to argue that psi research should start anew. With minor exceptions no psi research has officially been carried out there since the late 1930s. It could thus be argued that the U.S. had a head start.

The Soviets' attempt to catch up, combined with lack of reliable information and the sensational book *Psychic Discoveries Behind the Iron Curtain* (Ostrander & Schroeder, 1970), created the need for a threat assessment. The zeitgeist in the 1970s presumably also contributed to the U.S. intelligence agencies' increased interest. Later, in addition to Soviet émigrés' claims (Starr & McQuaid, 1985), popular books and articles concerning a psychic arms race emerged (McRae, 1984; White, 1988)—the impact of all this is not discussed.

The limits of psi are not known, and this worried the U.S. intelligence agencies. The need for a threat assessment resulted in what is now known as the Star Gate program, which focused on remote viewing. The program can in hindsight be said to have been initiated in 1972 by the CIA, which for a few years supported the research (Kress, 1977/1999; Richelson, 2001). Some in the CIA were clearly impressed by the initial research, but they also became aware of the still unsolved problem concerning how to separate fact from fiction in the remote viewing data. In addition, there was some unwanted publicity about the intelligence agencies' interest in psi (Wilhelm, 1976, 1977).

Given some stunning results and the lack of reliable information about the development in the Soviet

Union, a military program was nevertheless initiated in 1978 at Ft. Meade.

The authors claim that approximately 3,000 intelligence personnel worldwide were screened with regard to their potential participation in the program. Declassified documents, however, reveal that 251 personnel were considered, 117 initially interviewed, and 6 individuals eventually selected (CIA-RDP96-00788R001100020001-8); one of them was Joseph McMoneagle. Both Edwin May and McMoneagle contribute autobiographical sections; although interesting, they seem somewhat redundant given the book's focus. More interesting are the accounts about search tasks involving remote viewers. One account is unverifiable, but the others can be scrutinized.

One search task (Project 8916) was initiated in order to locate Drug Enforcement Agency agent Charles Jordan. The authors mention only remote viewer Angela Dellafiora Ford's participation, yet 17 sessions focused on finding Jordan and several viewers were involved—their impressions differed from one another. The authors also fail to note that Ford's preferred method is referred to as written remote viewing—like a trance medium, she channeled information from entities (Smith, 2005). In addition, it is far from clear if her impressions really led to the capture of Jordan as the authors claim (Graff, 1998, 2000; Nickell, 2004).

Another case of selective reporting concerns the remote viewers' involvement in the search for the kidnapped Brigadier General James Dozier. He was allegedly impressed by some of the remote viewing data, but it should be noted that it was a Red Brigades member, not McMoneagle's impressions, that provided the clues that led to his location (Graff, 2000). Also briefly covered is the remote viewers' involvement in the Iran hostage crisis—more than 200 sessions focused on this. Andrew Endersby's (2014) examination reveals that there was a lot of erroneous information in the remote viewing data.

The authors also claim that McMoneagle was involved in the attempt to locate a Soviet aircraft which had crashed in Zaire (this claim turns up elsewhere as well). The plane, however, crashed in March 1979, before he had completed his remote viewing training. That said, data from two remote viewers, Rosemary Smith and Gary Langford, made the search team shift their focus, which caused them to encounter natives who had found the plane (Graff, 2000; Smith, 2005).

Relatively little is said about the U.S. research on remote viewing. The reader is reminded of the researchers' inability to find out why some individuals are good remote viewers and others are not. In passing, the remote viewing experiment involving the submarine Taurus is also covered. Of most interest are the comments about the psychic Ingo Swann, who was tasked with developing a method to train remote viewers. According to May, he had a brilliant mind and worked 12 to 14 hours each day for years, but it is stressed that he was not a scientist and that the method he developed was fundamentally flawed. The main problem was that Swann was not blind to the target and provided feedback to the remote viewers during training sessions: "Assuming no psychic ability whatsoever, a person could arrive at the correct site via clever responses, conscious or unconscious" (pp. 140–141). May also claims that Swann instilled an anti-science attitude in his trainees, which contributed to the creation of a chasm between the researchers and the remote viewers at Ft. Meade.

Naturally the eventual closure of the Star Gate program in 1995 is also covered. May has expressed his views about this before (May, 1996), but he provides some additional information about what occurred behind the scenes. For more than a decade after the closure he and McMoneagle tried to start a new program—all their efforts were in vain. Nevertheless, it is still rumored that an active remote viewing program exists (e.g., Margolis, 2013, pp. 108–110).

So far, this review has barely touched on psi research in the Soviet Union, despite the fact that the sections concerning this are presumably of most interest. The claims that these sections contain are much harder to assess and are bound to exceed boggle thresholds, although the authors for their part seem to take most of the claims at face value. Due to the frustrating lack of details it is difficult to grasp the scope of psi research after the disintegration of the Soviet Union, but it seems as if in the 1990s, the focus was on applied psi rather than research.

Some of the stranger claims that reached the U.S. concered psychotronic generators—hardware devices, often with nothing under the casing that ostensibly utilized psychic energy. In the Soviet Union dozens were constructed but most of them never worked. That said, according to Major General Nikolai Sham, the few that did ". . . were, and still remain, utterly unique designs that were frequently ahead of

their time and which created the foundation for future technologies" (p. x). May, however, claims that Sham told him that although he had funded 40 different institutes specifically to develop psychotronic weapons they just could not get them to work. When the Cold War ended development of psychotronic weapons was discontinued.

Psychics were, however, still used, but according to the Russians their intelligence agencies only sporadically consulted psychics and primarily in crime investigations or counter-espionage operations. The authors note: "The KGB held two opposing views of psychics at the same time: Officially, the KGB prosecuted them. Unofficially, they secretly used them" (p. 281). Several examples are given, including some recollections by the psychic Tofik Dadashev (concerning him, see Gris & Dick, 1978). The Russians also, allegedly, had psychics on Dickson Island in the Kara Sea who ". . . studied the American military satellites by means of remote viewing, even to the details of their designs" (p. 170).

Major General Boris Ratnikov and Major General Georgii Rogozin consulted psychics concerning threats to Boris Yeltsin. On some occasions Ratnikov even changed Yeltsin's schedule based on information from psychics. They also used psychics to protect important politicians from psychic influence and mind reading, but no details are given. Ratnikov relates that a psychic informed him that Yuri Skokov, Secretary of the State Security Council, was subject to psychic influence during his and Yeltsin's visit in the U.S. in 1992: "Naturally, we tried to protect Skokov from this psychic influence and to block the leak of information through extrasensory methods" (p. 189). Ratnikov clearly believed such things could occur and has previously claimed that psychics read former U.S. Secretary of State Madeleine Albright's subconscious mind in 1999 (Smolchenko, 2007).

Equally mind-boggling is Lt. General Alexei Savin's account about military unit 10003. Savin comes across as an exceptional man who after three near-death experiences in childhood became psychic. Friends in high places ensured that his extensive program, initiated in 1989, was well-supported. Initially his staff consisted of 10 people, but by the year 2000, it included more than 50 people. Groups of psychics were trained (unfortunately no details are given about their training) and applied their abilities in a variety of ways. My impression is however that Savin would agree with Major General Nikolai Sham: "But most importantly, unique techniques of developing extraordinary human abilities and qualitatively increasing intellectual and spiritual levels were developed and carefully tested in practice" (p. xi). The program remained active until 2004.

In conclusion, it should be admitted that the book lacks an index, is somewhat poorly organized, and references are rarely given. Parapsychologists know where to search for more detailed information, but the book is clearly aimed at laymen. They would likely also appreciate a timeline with key events. Whether the true stories have really been provided can be disputed, but it is an interesting book. Most of the information concerning the Star Gate program is already in the open literature (e.g., May, 2014), but especially May's recollections provide some new glimpses behind the scenes. The authors also managed to get the Russians to share some new thought-provoking information that would fit well in a science fiction novel.

References

Ebon, M. (1983). *Psychic warfare: Threat or illusion?* New York, NY: McGraw-Hill.

Endersby, A. (2014). *America's imaginary hostage crisis*. North Charleston, SC: CreateSpace.

Graff, D. E. (1998). *Tracks in the psychic wildness*. Boston, MA: Element Books.

Graff, D. E. (2000). *River dreams*. Boston, MA: Element Books.

Gris, H., & Dick, W. (1978). *The new Soviet psychic discoveries*. Englewood Cliffs, NJ: Prentice-Hall.

Kress, K. A. (1999). Parapsychology in intelligence: A personal review and conclusions. *Journal of Scientific Exploration, 13*, 69–85. Original work published 1977.

Margolis, J. (2013). *The secret life of Uri Geller*. London, England: Watkins.

May, E. C. (1996). The American Institutes for Research review of the Department of Defense's STAR GATE program: A commentary. *Journal of Scientific Exploration, 10*, 89–107.

May, E. C. (2014). Star Gate: The U.S. government's psychic spying program. *Journal of Parapsychology, 78*, 5–18.

McRae, R. (1984). *Mind wars.* New York, NY: St. Martin's Press.

Nickell, J. (2004). *The mystery chronicles*. Lexington, KY: University Press of Kentucky.
Ostrander, S., & Schroeder, L. (1970). *Psychic discoveries behind the iron curtain*. Englewood Cliffs, NJ: Prentice-Hall.
Richelson, J. T. (2001). *The wizards of Langley*. Boulder, CO: Westview.
Smith, P. H. (2005). *Reading the enemy's mind*. New York, NY: Forge Books.
Smolchenko, A. (2007, November 7). Putting words in Albright's mouth. *Moscow Times*. Retrieved from http://www.themoscowtimes.com
Starr, D., & McQuaid, E. P. (1985). Psi soldiers in the Kremlin. *Omni, 7*, 80–82, 104–108.
Wilhelm, J. L. (1976). *The search for superman*. New York, NY: Pocket Books.
Wilhelm, J. L. (1977, August 7). Psychic spying? *Outlook/Washington Post Sunday Magazine*, B1, B5.
White, J. (1988). *Psychic warfare: Fact or fiction?* Wellingborough, England: Aquarian Press.

Nemo C. Mörck

Torggatan 6A lgh 1203
282 31 Tyringe, Sweden
nemomorck@hotmail.com

Journal of Parapsychology, 79(1), 119–123

CORRESPONDENCE

To the Editor:

I wish to thank professor Caroline Watt for taking the time to respond to my letter (in the Fall 2014 *JP*). Needless to say, opinions on the value of specific studies are bound to differ. Partly due to a misunderstanding Watt implied that I had not bothered to read her report carefully. The misunderstanding occurred due to my use of the expression "specific study" with which I did not mean to refer to meta-analyses or research overviews. It is admitted that Watt (2014) in her introduction wrote that "target feedback was rapidly given" (p. 116), yet to me it remained unclear what this meant in practice. In her response, Watt states that feedback was usually given within hours.

I appreciate her clarification, yet after having written my letter George P. Hansen made me aware of a potential problem. According to Watt's report, her study consisted of 200 trials (preplanned as four trials each from 50 participants) and there were 64 direct hits … giving a 32% hit rate. Using an exact binomial test, this result is significant, $z = 2.21$, $p = .015$, *one-tailed*, ES ($z/N^{1/2}$) = 0.16" (emphasis added, p. 120). This result is however not necessarily just due to psi, because 21 participants did not complete all four trials and the data from their trials were therefore discarded. The study did thus, in fact, consist of between 221 and 263 trials. The participants were not informed of the judges' ratings after each trial but target feedback was given after each trial, and presumably participants who felt that they were not doing well were more prone to drop out.

If there are few hits in the discarded data, the actual hit rate, when all trials are included, is probably not significant, and then Watt's reported hit rate was significant just due to a procedural oversight. This seems consistent with some of her puzzling results: "... there was no significant difference between the judges' ratings of targets and decoys. Furthermore the ratings for targets that scored a hit were on average no more similar to the dream reports than the ratings for those that did not ..." (Watt, 2014, p. 123).

Reference

Watt, C. (2014). Precognitive dreaming: Investigating anomalous cognition and psychological factors. *Journal of Parapsychology, 78*, 115–125.

Nemo C. Mörck

Torggatan 6A lgh 1203
282 31 Tyringe, Sweden
nemomorck@hotmail.com

To the Editor:

I am writing in response to Chris Carter's review of my book *Reimagining the Soul: Afterlife in the Age of Matter* in the Fall 2014 issue of this *Journal*.

Carter asserts that I suffer from intellectual timidity by kowtowing to the materialist belief system of mainstream science and neglecting the evidence for the survival of personality elements based on mediumship, near-death experiences (NDEs), and the rest of the usual suspects. I have been involved in parapsychology for over four decades. In the first three of those decades I largely devoted my time to critiques of parapsychological methodology, theories, and findings. These critiques address virtually all of the phenomena that Carter presents as supporting his own world view. I did not avoid these topics out of

timidity, but out of boredom and the desire to make progress in the field. I believe that my critiques over the decades have sufficiently demonstrated the weakness of the evidence for these "fringe" phenomena. I was not, as Carter implies, afraid of addressing these ostensible phenomena. I merely saw no point in beating a horse that in my view was already not only dead but had already made its pilgrimage to the glue factory. I refer the reader to my two previous books on parapsychology as well as my 750-page essay "Consciousness and the Physical World," which has been freely available on the Internet since 2006. Carter states that he has written a 300-page book on NDEs and 369-page book on mediumship, as if this somehow validates his views. If length of verbiage is to be the arbiter of truth, I have beaten Carter by 81 pages.

Mr. Carter accuses me of intellectual cowardice. In 1974, with a newly minted Ph.D. in experimental psychology from one of the nation's premier universities in hand, I walked away from the scientific establishment to join J. B. Rhine's Institute of Parapsychology, knowing that there could be no return voyage and that I was leaving the cushy life of a nonthinking academic conformist behind me forever. Does that sound like the act of an intellectual coward? I have refused to enter the womb of organized pseudoskeptics because I have no respect for them. I have basically expelled myself from the world of traditional academic parapsychology, as in my view this is a field that has largely dried up in its present form. I wished to explore the nature of the afterlife in a way that is unfettered from formal religion and the traditional preconceptions of psychical researchers. Are these the acts of someone too timid to confront the orthodox intellectual establishment, as Carter charges?

Two of Carter's favorite classes of phenomena suggestive of survival are out-of-body experiences (OBEs) and NDEs. However, reports of these phenomena have no direct bearing on the issue of survival, as they are provided by people with living brains. He also offers up terminal lucidity, in which dying brain-compromised patients become cognitively lucid just prior to death. I am intrigued by this phenomenon as well and am eager to see further research in this area. However, once again these reports involve subjects with living brains.

Carter also approvingly cites work by Rupert Sheldrake and Gary Schwartz as supporting his position. I have repeatedly pointed out systematic flaws in these authors' methodologies and theories over the years. A critique of Schwartz's work appears in *Reimagining the Soul* and critiques of both Sheldrake's and Schwartz's work are provided in my previous book *The Conscious Mind and the Material World*. Mr. Carter has even coauthored a book with Sheldrake. It appears that by "intellectual timidity," Carter means the reluctance to endorse each and every crackpot theory and bogus experimental result that comes down the pike.

Publishers are reluctant to publish long treatises, such as my above-mentioned 750-page essay. In *Reimagining the Soul*, I chose not to devote much space to these topics as the evidence for these phenomena is weak and I wished to use the available space to present a new and different approach to the problem of survival, rather than to reiterate my criticisms of most of the evidence for the survival of personality, which have already made their way into the academic literature on numerous occasions. This has been the new focus of my efforts over the past decade.

Given the overwhelming rejection of parapsychological phenomena by the scientific community, which is reasonable given the unconvincing evidence presented on their behalf, I thought it would be more interesting and productive to consider what forms of survival would be consistent with the materialistic worldview of modern science. Of course most modern scientists and parapsychologists would say "none." But they are generally reacting to naive religious conceptions of the soul and the afterlife, in which the personality is assumed to survive death in an essentially intact condition. Incidentally, my rejection of personal survival is not a priori (before consideration of the evidence), but a posteriori (after consideration of the evidence). I could become convinced of the survival of personality elements if the evidence were stronger. Right now, the strongest evidence in my view is the reported memories of past lives, to which I devote an entire 28-page chapter.

Strangely, none of the reviewers of *Reimagining* have mentioned the mathematical argument I present showing that the results of meta-analyses purportedly demonstrating the existence of psi are in fact about what one would expect based on the assumption that psi does not exist and known rates of fraud and

data selection in the general research community. In my view, this undermines the experimental support for psi phenomena. I view this as one of the main contributions of my book. Perhaps the reviewers did not recognize the significance of this argument because of its mathematical nature.

It seems that Carter took the book review as an opportunity to present his own ideas and to neglect completely the main thrust of my book. Thus, the reader can gain no inkling of the main ideas presented in my book from his review. He does not even mention the fact that I have proposed a model of survival involving a hierarchy of centers of pure consciousness that does not rest on the assumption that psi exists. It is quite possible that this model could be investigated empirically (by someone younger than me). If Carter thinks my ideas are based on cowardice, we will see how quickly the academic community embraces my model. I am sure that many if not most of them would, after a nanosecond of thought (actually pretty long for them), immediately dismiss my model as homuncular. However, in the model I propose the apparatus of thought is ascribed to the brain system as a whole rather than being located in a singular center of consciousness, or homunculus.

Incidentally, Carter states that I believe that only "indeterministic systems are associated with consciousness, as conscious minds would be of no use to a mechanistic system." This quotation is from David Hodgson, not me, as clearly cited in the text. It is hard to see how this view would be consistent with Carter's statement that I am a panpsychist. (Perhaps this works if one doesn't think about it at all.) In fact, my model of the soul and consciousness is quite compatible with billiard ball Newtonian determinism and does not depend on the existence of psi or indeterminism in any way.

Carter's own world view is largely the same as that of most orthodox psychical researchers working in the area of survival. Perhaps he is the timid one, and is afraid to leave the nest of orthodox parapsychology, whereas I jumped out of that nest a decade ago.

Finally, *Reimagining* is written in a somewhat playful, nonlinear manner, with micropoetry and off-the-wall comments scattered throughout the text. Strangely, none of the reviewers have commented on that. (Perhaps just as well.)

Douglas M. Stokes

424 Little Lake Drive, #3
Ann Arbor, MI 48103, USA
Dstokes48103@yahoo.com

To the Editor:

I wrote a highly critical review of Stokes' book *Reimaging the Soul: Afterlife in the Age of Matter*, to which the author has responded. Here is my rejoinder.

In my review of this book I criticized what I believe to be the book's superficial treatment and cavalier dismissal of the evidence—from near-death experiences, mediumship, and terminal lucidity—that contradict the author's repeatedly stated view that personality and memory depend upon a functioning brain and so cannot survive the death of the brain.

In the introduction to his book Stokes clearly stated that the central arguments of his book "will not depend on parapsychological evidence such as hauntings, claimed memories of previous lives, and ostensible messages from the dead provided by mediums or in dreams, *as these findings are not accepted by mainstream scientists*" (p. 5, emphasis added). Claims similar to this—"most scientists are skeptical," "rejected by mainstream science," etc.—are made repeatedly throughout this book, and that was what I meant by a "curious conformist timidity" of the author's writing.

However, in the second paragraph of his reply, Stokes writes that "I believe that my critiques over the decades have sufficiently demonstrated the weakness of the evidence for these 'fringe' phenomena. I merely saw no point in beating a horse that was . . . already dead." In my review I quoted his criticisms of the evidence from NDEs and mediumship. If these remarks dismissing the evidence from NDEs and me-

diumship are an accurate summary of his views, then he has most certainly *not* "sufficiently demonstrated the weakness of the evidence."

Consider his treatment of the NDE: At the end of the 2½ pages Stokes devoted to the NDE, he casually concluded that "Various neurophysiological causes for such *hallucinations* have been proposed, including seizures to the temporal lobes of the brain, lack of oxygen to the brain, the release of endorphins in the brain, and the random firing of cells in the visual cortex" (pp. 120–121, emphasis added).

With regard to hallucinations as an explanation of the NDE—whether caused by temporal lobe seizures, anoxia, or drugs—neuropsychiatrist Bruce Greyson, editor of the *Journal of Near-Death Studies*, wrote: "Without exception, every report of a large study of NDEs published in a mainstream medical journal has concluded that these phenomena cannot be explained as hallucinations. Such unanimity among scientific researchers is unusual and should tell us something. Why is it that scientists who have done the most near-death research believe the mind is not exclusively housed in the brain, whereas those who regard NDEs as hallucinations by and large have not conducted any studies of the phenomena at all?" (Greyson, 2007, p.140, emphasis added). "The major advantage of the hallucination model is its compatibility with the materialistic worldview favored by a majority of neuroscientists (though not by a majority of physicists). The major disadvantage of the hallucination model is that it fails to account for the phenomenon [accurate perceptions], and is plausible only if we discredit or discount much of our data. As astronomer and spectroscopy pioneer Paul Merrill quipped, 'If you eliminate the data that do not agree, the remaining data agree very well.' But disregarding disagreeable data is the hallmark of pseudoscience, not science" (Greyson, 2007, pp. 141–142).

Stokes added: "Reports of these phenomena [NDEs] have no direct bearing on the issue of survival, as they are provided by people with living brains." While I do agree that the NDE provides at most only suggestive evidence of survival, that was not the point of mentioning NDEs in my review. They were mentioned only as providing "evidence that the mind, complete with memories, may exist in the absence of a properly functioning brain." Reported memories of lucid experiences which seem to have occurred during periods in which there is every medical reason to believe that the subjects' brains were either severely impaired or entirely nonfunctioning (that is, during a state of clinical death) most certainly are evidence contrary to Stokes's stated belief in the dependence of memory and personal identity on a properly functioning brain.

Stokes then mentions that I approvingly cite work by Rupert Sheldrake and Gary Schwartz. In my review Schwartz's name was only mentioned once in passing, and I quoted only a purely logical argument by Sheldrake with regard to loss of memories from brain damage, and what that does or does not imply for that hypothesis that memories are stored in the brain. I also did not "coauthor a book with Sheldrake"—he wrote the Foreword to one of my books. Sheldrake's theories are nonorthodox yet testable, insofar as the criterion of the scientific status of a theory is its testability. A list and results of his published papers and empirical findings can be found online at http://www.sheldrake.org/research.

Stokes then writes: "Given the overwhelming rejection of parapsychological phenomena by the scientific community . . . I thought it would be more interesting and productive to consider what forms of survival would be consistent with the materialistic worldview of modern science."

However, if Stokes had read the book that I "coauthored" with Rupert Sheldrake, he would have learned that it is a myth that most scientists reject the existence of psychic phenomena, and also that the worldview of modern science is no longer materialistic. Two surveys of over 500 scientists in one case and over a thousand in another were made in the 1970s. Both surveys found that the majority of respondents considered ESP "an established fact" or "a likely possibility": 56% in one (Evans, 1973) and 67% in the other (Wagner & Monet, 1979). In the study by Evans, 53% of the "ESP is an impossibility" responses came from psychologists, although psychologists made up only 6% of the total sample. Only 3% of natural scientists considered ESP "an impossibility," compared to 34% of psychologists. Since Stokes is a psychologist, I suspect that his experience with "the scientific community" may be heavily skewed.

As for the "materialistic worldview of modern science," this is also a misconception, and one that is common among psychologists. The great psychologist Gardner Murphy (1969), president of the Ameri-

can Psychological Association and later of the American Society for Psychical Research, urged his fellow psychologists to become better acquainted with modern physics.

> . . . the difficulty is at the level of physics, not at the level of psychology. Psychologists may be a little bewildered when they encounter modern physicists who take these phenomena in stride, in fact, take them much more seriously than psychologists do, saying, as physicists, that they are no longer bound by the types of Newtonian energy distribution, inverse square laws, etc., with which scientists used to regard themselves as tightly bound. (p. 527)

Matter is not even considered in remotely the same terms in modern physics as it was in classical physics. Atoms are no longer thought of as "solid, massy, hard, impenetrable moveable particles," as Newton described them, but rather as potentialities, possibilities with a wavelike structure that can interfere like waves. Possibilities that become fully real only when observed are more like ideas than like tiny, observer-independent billiard balls. The classical idea of substance—self-sufficient, unchanging, with definite location, motion, and extension in space—has been replaced by the idea that physical reality is not made out of any material substance, but rather out of events and possibilities for those events to occur. These possibilities, or potentials, for events to occur have a wavelike structure and can interfere with each other. As the celebrated philosopher of science Karl Popper frequently remarked, materialism has transcended itself.

In conclusion, I stand by my review of Stokes's book. I did admire parts of it, although I disagreed with most of the conclusions.

References

Evans, C. (1973, January). Parapsychology—What the questionnaire revealed. *New Scientist*, *25*, 209.

Greyson, B. (2007). Commentary on "Psychophysiological and cultural correlates undermining a survivalist interpretation of near-death experiences." *Journal of Near-Death Studies 26*, 127–145.

Murphy, G. (1969). Psychology in the year 2000. In L. B. Murphy (Ed.), *There is more beyond: Selected papers of Gardner Murphy*. Jefferson,NC: McFarland.

Wagner, M., & Monet, M. (1979). Attitudes of college professors toward extra-sensory perception. *Zetetic Scholar*, *5*, 7–16.

Chris Carter

2 Halkirk Bay
Winnipeg, MB, Canada R2K 2V7
webslinger_999@yahoo.com