

BOOK REVIEWS

EVIDENCE FOR PSI: THIRTEEN EMPIRICAL RESEARCH REPORTS edited by Damien Broderick and Ben Goertzel. Jefferson, NC: McFarland, 2015. Pp. viii + 324. \$49.95 (paperback). ISBN 978-0-7864-7828-6.

Previously, when I had to recommend a book to a friend or colleague interested in psi research, I usually turned to Dean Radin's (1997) *The Conscious Universe*—rightly called “the parapsychologist's manifesto”—to try to win their sympathies. For a popular audience, I still would. But for scientifically trained colleagues, a book has emerged recently that has the potential to pack a much more spirited punch, enclosed in a small volume. That book is *Evidence for Psi: Thirteen Empirical Research Reports*, edited by Benjamin Goertzel and Damien Broderick. A whirlwind tour of psi research, it features simple, proof-oriented experiments, multidisciplinary lines of evidence, and even a few theoretical perspectives—all, however, grounded in solid empiricism. This diversity gives the book a certain robustness, avoiding the narrow lens through which the field is sometimes examined. In my opinion, it comes closer than almost any other (similarly sized) text to explaining what is so compelling about psi research as to drive numerous highly qualified scientists to put up with it, often for the remainder of their careers, despite professional peril and lack of funding. Normally, such a task would be quite daunting—but Broderick and Goertzel embrace it with relish.

Their book has a simple structure; there are 15 chapters, 12 of which are papers by psi researchers, and three of which are written by the authors: an introduction, skeptical overview, and conclusion. This review summarizes each chapter, focusing on the reports I have the most to say about—beginning with the Introduction, which I found to be something of a mixed bag. While Broderick and Goertzel rather ambitiously tried to summarize all of the early history, scientific motivation, available paradigms, research difficulties, and theoretical models for psi in their opening chapter, the effect seemed to be less one of ambition and more one of loss of focus. Their lengthy discussion of the early mediumship studies is exemplary of this; they mention that psychical researchers looked at everything from table rappings to ectoplasm, stating that there *could be* something to these things because contemporary sources (e.g., *Randi's Prize*; McLuhan, 2010) have given “detailed analyses” of the skeptical claims. But they do not provide these analyses. Even though they do express skepticism of the early research, it must seem cheap and unearned to the novice reader, mixing up Broderick and Goertzel's own admirable openness with the rigor of the evidence for psi—a theme repeated in their later discussion of the world-as-a-simulation hypothesis. However, this apart, there is much to commend in the Introduction, and a great deal that is valuable to novices. B&D vitiate parapsychology's *raison d'être*, explaining why scientists study psi in the first place (people from all over the world report psi experiences!), successfully summarizing some of the research paradigms available, and prepping the territory for the coming reports. Additionally, B&D discuss how researchers have tried to fit psi into their understanding of the world, through building models and through attempting to falsify them. All of this makes the Introduction an important foundation for the contributions in the rest of the book.

Among these, the chapter by Jessica Utts, the current president-elect of the American Statistical Association, has to be acknowledged as a powerful accounting of concepts in meta-analysis. In this segment, Utts lays out the pedagogy of statistical power, heterogeneity, effect sizes, confidence intervals, and more, in a side-by-side comparison of a psi meta-analysis and a mainstream one. The effect is to show that psi meta-analyses follow proper statistical guidelines and are no less reliable than standard meta-analyses—by now a classic refrain of psi proponents (for good reason). Although some of the details given, on hypothesis testing, may be tedious to those familiar with them, the sections on meta-analysis present worthwhile insights, such as that psi effects are often larger than mainstream effects, despite being considered “small,” and that small effects can have big consequences.

The next chapter, by Julia Mossbridge, a perception neuroscientist at Northwestern University, is a formidable part of *Evidence for Psi*. Mossbridge studies presentiment; her experiments look for physiological activity that is statistically predictive of randomly determined future events (she calls this phenomenon “Anomalous Anticipatory Activity,” or AAA). In this segment, Mossbridge presents data not only from her own experiments, but also from the meta-analysis of presentiment work she published with Utts and Tressoldi in the Swedish journal, *Frontiers in Psychology*. The gross details of that publication can be found in the book, but what drew my interest in this chapter was Mossbridge’s fortuitous discovery of a rather large, sex-differentiated presentiment effect, because she uses a methodology that conclusively rules out “expectation bias.” That is, she examines only the first trial of every participant, rather than all the trials; this makes it impossible for the averaging artifacts discussed by Dalkvist and Westerlund (2006) to interfere with the results. The effect is fortuitous, too, because in eliminating this bias, it was very possible for Mossbridge to discover a significant reduction in effect size, but what she found instead was that it skyrocketed to $d = 1.41$ —nearly triple the average effect size in psychology experiments ($d = .5$; Bakker, Dijk, & Wicherts, 2012)—but only in males. The reason is simple enough, and is perfectly aligned with the presentiment hypothesis and the nature of the stimulus. Mossbridge asked participants to guess which of two images a computer would (randomly) determine to show next—a psi-guessing task that, by itself, did not return significant results. But when the postfeedback responses of males and females were compared, only males seemed to become more excited when they found out they had answered correctly, versus incorrectly; therefore, only males would be expected to become more excited *prior* to feedback on a correct choice. Most exciting of all, in my opinion (as I found out through personal communications), Mossbridge appears to be replicating this effect, finding one almost twice as large in heart rhythm and replicating the gender differential in her smart-phone research. If this pattern of results holds, Mossbridge’s experiment may become the parapsychologist’s replication paradigm *du jour*, with an effect size that nearly obviates the need for statistical analysis.

Along similar lines, the next chapter in the book, authored by Edwin May, Tamas Paulinyi, and Zoltan Vassy (MPV), discusses possible experimenter effects in presentiment work. The primary portion of this chapter reports a simple replication experiment with audio stimuli, finding results like those of other researchers. But the interesting section focuses on differentiating experimenter versus participant effects. As the authors note, it is often assumed that any presentiment difference is a result of the participant’s physiology responding to a future event, but it could also have something to do with the experimenter’s intuition. The evidence for this is subtle, but interesting; in their experiment, MPV noticed that nonspecific skin conductance responses (ns-SCRs) were relatively infrequent, such that it would be “easier” for an experimenter to intuit (with psi) the starting time for a session so as to *avoid* the maximal number of ns-SCRs prior to silent controls, than so as to *capture* the maximal number of ns-SCRs prior to actual audio startle stimuli (both of which would cause the number of ns-SCRs prior to silent controls versus prior to audio stimuli to differ from each other, leading to significant results). In their analysis, this is exactly what MPV reported: Pre-control ns-SCRs were significantly depressed and pre-audio ns-SCRs were statistically no different from the background rate. So some kind of experimenter effect appears to be a viable and likely explanation. Notably, however, their effect size was very small ($d = .14$), closer to the effect size of forced-choice guessing experiments than to mainstream presentiment experiments ($d = .43$; Mossbridge, Utts, & Tressoldi, 2012). This suggests that the kind of subtle, pattern-sorting intuition discussed by May et al. in this chapter is usually insufficient to produce large effects, but may affect many parapsychology experiments with small effect sizes, which are legion.

Regaining the Utts’ birds-eye-view approach to psi research—and leaving, for the moment, the thickets of experimental methodology—we have a chapter by Bryan J. Williams on a meta-analysis of ganzfeld experiments, one of the most tested, classic paradigms in parapsychology. Williams sets out to verify whether ganzfeld experiments have been replicated, after the famous 1986 *Joint Communique* between proponent, Charles Honorton, and skeptic, Ray Hyman, set out the protocols for a rigorous ganzfeld study. Along the way, Williams treats the reader to a detailed history of ganzfeld research and fairly assesses several of the key debates around this paradigm. He presents useful graphics of experiment success rates that

help the reader get a feel for the variability in ganzfeld results, as well as helpful breakdowns of results by meta-analysis and laboratory. My only criticism of this chapter is that insufficient attention is given to the issue of selected versus unselected participants. The author treats the heterogeneity in experimental results as a natural outcome of design differences, or does not mention it, when there is powerful evidence that almost all of the significant results in the ganzfeld database come from testing participants preselected for psi-enhancing qualities, while other kinds of participants get results only marginally different from chance (Baptista & Derakhshani, 2014)—within the range of “experimenter” effect sizes that MPV found in their chapter.

On the topic of experimenter effects, and experimenters who get exceptional results, probably few compare to the author of the next chapter: Rupert Sheldrake, a biologist, and a fellow of Clare College, Cambridge. This is because Sheldrake’s inventive experiments consistently deliver large, significant results (if they are experimenter effects, they seem to be atypically robust ones). His approach—no doubt molded by his profession—is to set up “ecologically valid” psi experiments that replicate spontaneous psi experiences reported by many people. One of these is “telephone telepathy,” the phenomenon of suddenly thinking about someone before they call. In his chapter, Sheldrake reports on how he tested this with videotaped telephone telepathy experiments in which callers were randomly prompted to call via the toss of a die, and the person to pick up the phone had to guess who called them. He reports a high average hit rate of 45% for these trials, where 25% would be expected by chance. Sheldrake also goes on to discuss a variety of studies he’s done on animal telepathy and telepathy with SMS and email, and even an interestingly designed precognition experiment that gave null results, suggesting to Sheldrake that it really is telepathy—not precognition—that accounts for a person’s knowledge of who is on the other end of the line. Overall, I found Sheldrake’s chapter a nice addition to the book and noticeably different in flavor from the contributions of the other parapsychologists, none of whom are biologists.

In the subsequent chapter, Bryan J. Williams takes the plate again, providing an extensive review of tests done on the psychic Sean Harribance, trying to get at the question of whether he displayed genuine psi. This is probably the most detail-oriented contribution, complete with protocol descriptions, experimenter names, diagrams of the laboratory, tables of results, and arguments and counterarguments—all trying to show that Harribance could not have scored the way he did (on forced choice and micro-PK tasks) by ordinary sensory means or trickery. Indeed, the evidence seems compelling: 10 consecutive experiments conducted at the Psychical Research Foundation, with increasing precautions, showed generally above average, highly statistically significant scoring, with no apparent drop-off in hit rates. Tests done by other laboratories also found highly statistically significant results, suggesting that Harribance really did have psychic powers. Reading this chapter brought me back to J. B. Rhine’s and J. G. Pratt’s experiments with Hubert Pearce in the 30s—this is the classic rule-out-all-the-flaws parapsychology study, and Harribance, a psychic from a small village in the West Indies, is the paradigm case of a selected subject. This makes Williams’ chapter a fine addition to the book, because in modern parapsychology it is sometimes easy to lose sight of the fact that many, many trials have been conducted with talented *individuals* (as opposed to groups) who produced consistent, reliable effects—and that careful investigators instituted numerous precautions to ensure that they did not cheat.

However, the topic of individuals, and individual ability more generally, is not covered only by Williams. Suitbert Ertel, an emeritus psychology professor from Georg-August University, addresses it from a psychometric perspective in his chapter. Ertel points out (correctly) that relatively little work has been done in parapsychology on the issue of estimating psi ability—no one has developed a “PsiQ” test and used commonly accepted psychometric tools to assess its validity and reliability. As a start to that, he offers a simple procedure, wherein participants close their eyes and draw a ball out of a bag. This procedure is not intended to prove the existence of psi—just as an IQ test cannot, by itself, prove that it measures intelligence (or some presumably important aspect of it). In fact, Ertel often lets participants complete the tests at home, without supervision. His goal is to use the results of the “Ball Selection Test” (BST) to explore both the characteristics of psi giftedness, and possibly psi itself, for which purpose he amassed somewhere near 90,000 trials, and a cluster of analyses providing evidence that his results meet two key psychometric

criteria: consistency (they recur in a predictable fashion) and validity (they measure psi). Ertel examines, for instance, the test-retest correlation of his participants (comparing the 1st, 3rd, and 5th trials to the 2nd, 4th, and 6th trials, for a sample of 143 students who each contributed at least 6 runs), finding a Spearman-Brown corrected correlation of $r = .78$ —“medium reliability,” according to psychometric convention. Tempering this result, it should be noted that the test-retest reliability of IQ, which has undergone decades of development by psychometricians, is in the vicinity of .95, for sessions administered *about a year* apart (Brody, 1992). The Big Five Inventory, on the other hand, a well-accepted tool in psychometrics also, is more limited, with a test-retest correlation range of $r = .70$ –.79 (Hampson & Goldberg, 2006 for one-year intervals in adulthood, and .80–.84 for two-week intervals (Gnamb, 2015). Ertel presents a graph in his book that does suggest the strength of his correlation depends on the participant’s average scores, so that mixing in the correlations of people who are just guessing ($r = 0$ by default) with those who seem to display some psi ability will inevitably dilute the correlation. Still, it is clear that the proposed PsiQ test has a way to go.

Ertel’s other analyses are numerous; he looks for earmarks of psi, including hit displacement, psi missing, and hit spreading—a novel property, indicating that higher-scoring participants spread their hits more evenly than would be expected from random distributions. He also tests skeptical hypotheses and brings his higher scorers into the lab, putting them through progressively more controlled tasks. I found Ertel to be a little thin in this part, especially because it was evident that most participants did poorer in these tests; I would have liked to see more details, especially on the computerized Zener cards. Still, there is a lot of good material here. Although effect sizes tended to drop significantly, highly statistically significant results were still observed. For the purpose of convincing investigators who already find psi plausible, I agree with Ertel that these tests are probably sufficient, and validate the BST as a discrimination tool for psi ability. Investigators might ask prospective participants to conduct it at home, to screen for good candidates for psi studies that examine similar tasks. Overall, I greatly enjoyed Ertel’s chapter, and I found myself agreeing much with his perspectives. Continuing a theme, his contribution adds to the diversity of research perspectives in the book, introducing the reader to a data-rich psychometric exploration of psi.

By contrast, Stephan Schwartz offers a look into psi from an archeological angle, detailing the accomplishments of his Mobius remote viewing project (but before he does that, he presents a useful overview of experimental remote viewing experiments). Schwartz claims to have used remote viewers to locate sunken ships, uncover sites of historical interest in the Bay of Alexandria, and discover buried Byzantine ruins in the Egyptian desert. One particularly striking example of one of his successes involved a challenge to find, well, anything, in an area of more than 500 square kilometers, suggested to his team by skeptical archeologists at the University of Alexandria. Although the area had been previously surveyed by researchers at the University of Gelf, without any success, the remote viewers located within several hours what they described as a building foundation and mapped out its contours. After 3 to 4 feet of digging, the archeologists found the structure much as described. Reports like this feature prominently in Schwartz’s chapter, suggesting that the spontaneity and strength of psi outside the lab, with talented participants, can be significant. Although largely qualitative, his “consensus methodology” offers a degree of quantitative evaluation for his studies, especially likely to be useful to those readers familiar with the archeological process and its accuracies and inaccuracies. Mostly, however, the chapter appears to be there to break the reader out of the narrow confines of laboratory work, to see what psi can do in a freer setting.

Physicist York Dobyns’ chapter, up next, reports on work done at the famous Princeton Engineering Anomalies Research (PEAR) laboratory, on both remote viewing and psychokinesis (I will confine my comments to the latter, for brevity’s sake). In the psychokinesis experiments, “operators” attempted to influence the results of random number generators (RNGs), pushing them to produce more 1s, more 0s, or an approximately even distribution of both. The type of trial was determined either by the operator in advance, or assigned to him or her by a random process. The findings collectively indicate that operators were able to shift approximately one bit per 1,000 in their intended direction—or .20 bits in a customary 200 bit trial—and that this yielded highly statistically significant results, with odds against chance of more than seventeen million to one. Lots of details are offered on these experiments: for instance, the curious discovery that among trials where two operators tried to influence the RNG bits in unison, same-sex pairs had

no effect and opposite sex pairs had approximately four times the effect of single operators. Dobyns also overviews difficulties in replicating previous findings and gives refutations of skeptical criticisms, creating an impression of balance. His chapter, as an essential addition to *Evidence for Psi*, discusses one of the most important parapsychological collaborations created to study micro-psychokinesis, out of which arose the Global Consciousness Project (GCP).

This project, headed by Roger Nelson in his home office in Princeton, NJ, began with the FieldReg trials at PEAR; they examined if mental activity could unintentionally induce changes in the bit streams of local RNGs. Experimental successes observed with this method prompted parapsychologists to launch a global distributed network of RNGs, to explore possible data structures caused by psi (e.g., correlations between RNGs, effect of distance on psi, etc.), and whether events of collective global importance would correspond to measurable departures from randomness. After recording more than 400 carefully defined events, including 9/11, the 2004 tsunami, and the election of president Obama, the GCP has amassed a statistical significance of more than seven sigma, or odds against chance of 100 billion to 1; this indicates that RNGs are significantly less random during periods of global import. Nelson, who writes this chapter, provides a detailed breakdown of the equipment, procedures, and results of the GCP. In addition, he imparts interesting insights into the structure of the data during global events—RNG outputs do correlate across different machines, and there is a small distance-time relation—and discusses several sources of competing, non-psi explanations (e.g., electromagnetic disturbances from cell phones, etc.). His chapter reviews a critical, ongoing experiment in parapsychology, setting the reader up for one of the most intriguing articles in the book, which discusses models for the GCP data.

Physicist Peter Bancel, a long-time collaborator with Nelson, writes of how one can formally distinguish between two overarching models for the anomalous GCP effects: one with psi and one with “proto-psi”—a kind of unintentional, residual psi. He uses the mathematical technique of separation of variables to argue that when anomalous perturbations occur without intentionality/engagement (as they presumably do in the GCP), their characteristics should be different than when engagement is present. Also, to form a complete description of any experiment with engagement, that experiment’s results must be included in the description (not necessary for nonengagement). For instance, if I *intend* to select the correct target from among a group of decoys, as a psi task, then what the results of that task will measure is how well I was able to match my intention—there is an inherent circularity in any experiment with engagement, which cannot be avoided. On the other hand, if I merely wish to, say, plot temperature against reaction rate, for a certain chemical mix, my engagement is not a part of the experiment description, and there is no circularity.

As confusing as this is—and it is quite confusing—it illustrates a key difference in how GCP results should theoretically manifest, depending on whether they do measure some latent proto-psi, or are, for instance, an experimenter effect. With this information, Bancel takes the brilliant next step of examining the physical structure of the random event generators (REGs) used in the GCP. What he finds is that one kind of REG has an internal setup that is incompatible with proto-psi as has been formulated, because it uses a deterministic process to randomize the bits after they have been randomly generated from a stochastic physical process—one that could have been affected by proto-psi global consciousness. But if there is truly no engagement in this setup, then any nonrandomness before this deterministic screen would get destroyed; event-correlated nonrandomness emerging after that screen could then *only have been produced* by a contrivance—a prior fitting of the bits so that the (constantly updated) deterministic algorithm would produce nonrandom bits on the other end. This is *de facto* engagement. The catch is that the other kind of REG used in the GCP contains a structural loophole that could, just barely, allow some nonrandomness through. Bancel examines correlations within and between these types of REGs and finds that (a) the non-loophole REGs are not intercorrelated, (b) the loophole REGs are, and (c) the nonloophole and loophole REGs are strangely, strongly intercorrelated. He interprets this last finding as evidence against the proto-psi view, discusses possible alternatives, and then explores other analyses providing evidence that the GCP effect is not well explained by psi selection. His chapter, as a contribution to Broderick and Goertzel’s book, illustrates that psi research can sometimes (and perhaps will increasingly) benefit from the rigor and exactness that pervades the physical sciences, although the conclusions that can be drawn with this constrained approach are still limited.

Coming from a very different angle, but no less intriguing or exact, is the paper by James P. Spottiswoode, a statistician and former consultant to the World Bank, inviting the reader to take a broader perspective on psi by examining a truly unusual correlate: local sidereal time (LST). LST is the time of day as indicated by the rotation of the earth relative to the fixed stars, rather than to the sun. The effect of using this measure is that, across the year, the same sidereal times very accurately correspond to the same regions of the sky, above particular places on the earth's surface, while any given earth time over the course of a year varies across the full celestial circle, and all sidereal times. This allows one to check if something about the position of the earth relative to the sky correlates with psi, and Spottiswoode's data suggest that it does, to a significant extent. His findings document a more than threefold increase in the effect size of free-response psi experiments (mostly ganzfeld and remote viewing trials) around 13.5 h LST, with a sample size of nearly 1,500 trials. Providing further confirmation of this, Spottiswoode crossvalidates his finding with another approximately 1,000 trials, finding a peak at *exactly* the same LST. Together, the two data sets support the hypothesis of an increased effect size in the window of 13.5 ± 1 h LST, for trials conducted (mostly) in the US and Europe, evincing a signal strength of more than five sigma; the physics gold standard. Spottiswoode additionally explores LST as a moderator of the sometimes observed psi-GMF (geomagnetic field fluctuation) correlation, and proposes an intriguing explanation for the LST findings that involves emissions from the center of the Milky Way (neutral hydrogen noise appears to drop off at about the time the psi effect peaks). This chapter truly magnifies the psi research perspective, suggesting moderators of astronomical origin that persist across decades, creating patterns in the psi data that are difficult to explain by any conventional account.

Nevertheless, conventional accounts are briefly looked at in the next chapter, on skeptical responses to psi research. This chapter is written by one of the authors (Ben) and his father (Ted), a sociologist without any connection to psi research, but with experience studying movements, belief structures, and statistical claims. Together they produce an overview of the history of the modern skeptical movement and several of its controversies, trying to give the reader an idea of the state-of-play between professional skepticism and parapsychology today. Not much specific rebutting of criticisms is done in the chapter—much of that was covered by the book's contributors in their various chapters—instead, only a gestalt summary of psi skepticism and a few key examples are given. This brevity and generality of focus are repeated in the final overview chapter, where the authors give their assessments of the evidence, ask some questions, and critique some traditional approaches. They wrap up by speculating on the relationship between psi and physics and psi and biology, giving a couple recommendations for the field.

By now, the reader has hopefully come to terms with the scope and ambition of Broderick and Goertzel's book, which aims to present the best case for psi research, with the aid of the top scientists in the field. In my opinion, *Evidence for Psi* is a rigorously presented, well-curated sample of research papers. The authors are cautious and fair, and although I may disagree with some of their assessments and choices of presentation (e.g., their tacit acceptance of the decline effect as an all-pervasive phenomenon, which is not supported by any of the papers in the book, or their claim that psi experiment *z* scores do not scale positively with sample size, with which I disagree), the vast majority of what they write is in my opinion professional, expository commentary, of the quality that will attract many open minds. The contributions in the book, moreover, are selected from among the best—and it shows—as I had many worthwhile insights while reading it. If someone asks me, therefore, to scientifically justify my “hobby” in psi research, I would feel confident lending them a copy of *Evidence for Psi*, in the knowledge that it will represent me—and my reasons for taking psi seriously—better than I ever could.

References

- Bakker, M., Dijk, A. V., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science*, 7, 543–554.
- Baptista, J., & Derakhshani, M. (2014). Beyond the coin toss: Examining Wiseman's criticisms of parapsychology. *Journal of Parapsychology*, 78, 56–79.
- Brody, N. (1992). *Intelligence*. San Diego, CA: Academic.

- Dalkvist, J., & Westerlund, J. (2006). A bias caused by inappropriate averaging in experiments with randomized stimuli. *Journal of Parapsychology, 70*, 233–254.
- Gnambs, T. (2015). Sociodemographic effects on the test-retest reliability of the big five inventory. *European Journal of Psychological Assessment, 31*, 1–5.
- Hampson, S. E., & Goldberg, L. R. (2006). A first large-cohort study of personality-trait stability over the 40 years between elementary school and midlife. *Journal of Personality and Social Psychology, 91*, 763–779.
- McLuhan, R. (2010). *Randi's prize: What skeptics say about the paranormal, why they are wrong, and why it matters*. London, England: Troubador.
- Mossbridge, J., Tressoldi, P., & Utts, J. (2012). Predictive physiological anticipation preceding seemingly unpredictable stimuli: A meta-analysis. *Frontiers in Psychology, 3*, 390.
- Radin, D. (1997). *The conscious universe: The scientific truth of psychic phenomena*. New York, NY: HarperCollins.

JOHANN BAPTISTA

900 State St.
P.O. Box C267
Salem, OR 97301, USA
baptistajohann@gmail.com