Journal of Parapsychology, 81(2), 143-159

PA Presidential Address 2017: Withering Skepticism

Chris A. Roe[1]

University of Northampton

Abstract: In this address I consider the objections to parapsychology offered by high profile representatives of the skeptical community in their contributions to Krippner and Friedman's (2010) book, *Debating psychic experience*, comparing them with previous critical commentary published in 1993 as part of a special issue of *Scienza & Paranormale* devoted to parapsychology. Many of the claims made in that earlier collection had already been challenged by Charles Honorton in his classic paper, "Rhetoric over substance: The impoverished state of skepticism", and I revisit some of those arguments to gauge the extent to which skepticism has responded to them. I find that despite the passage of 25 years and the concomitant advances in the approaches, methods, and accumulated data of parapsychology, skepticism shows little evidence of progress and in fact may have regressed to more rudimentary rhetorical practices that rely on vague aspersions and show very little familiarity with actual activities and findings of contemporary parapsychology.

Keywords: skepticism, rhetoric, replication. parapsychology

Charles (Chuck) Honorton died on November 4 1992, aged only 46, as a consequence of the congenital disease osteogenesis imperfecta. Despite his premature demise, Chuck was arguably the most influential parapsychologist of the second half of the 20th century. He was a major contributor to two of the most fruitful lines of research into ESP — ganzfeld stimulation and dream states— and was a great driver of methodological innovation, refinement, and automation, particularly as a means of improving research quality. But perhaps his main contribution was in setting a standard for how to engage with skeptics. Where he regarded their criticisms or counter-explanations as frivolous or wilfully misrepresentative — as he found in the scenarios proposed by Mark Hansel (1966, 1989) to account for outcomes in a set of classic parapsychological studies — his criticism was suitably scathing, but where he felt that such interactions could be constructive — such as in his initial debate with Ray Hyman concerning the early ganzfeld work— he was a willing and respectful adversary (cf. Honorton, 1985). Perhaps the most notable consequence of the latter was the joint communiqué he co-wrote with Hyman in which they agreed:

> There is an overall significant effect in this data base that cannot reasonably be explained by selective reporting or multiple analysis. We continue to differ over the degree to which the effect constitutes evidence for psi, but we agree that the final verdict awaits the outcome of future experiments conducted by a broader range of investigators and according to more stringent standards. (Hyman & Honorton, 1986, p. 351)

1 Correspondence about this article should be sent to: Chris Roe, Centre for the Study of Anomalous Psychological Processes, The University of Northampton, Park Campus, Northampton NN2 7AL, UK, chris.roe@northampton.ac.uk

They outlined what those more stringent standards should be, and Honorton himself took up the challenge of producing a gold standard protocol that addressed the various methodological concerns that Hyman had raised; if the significant hit rates were still observed then it would cast doubt on the likelihood that these “flaws” could explain earlier successes, illustrating how criticism could be constructive in specifying the conditions under which the case for psi would be more compelling.

Shortly before he died, Chuck was invited by the Italian Skeptics’ Society CICAP to contribute to a special issue of their journal *Scienza & Paranormale* that would be devoted to parapsychology. They had three position papers from parapsychologists Richard Broughton, Bob Morris, and John Palmer for which Susan Blackmore would provide a response, and three critiques of parapsychology from James Alcock, Ray Hyman, and James Randi, on which Honorton would provide a commentary. In his contribution, “Rhetoric over substance: The impoverished state of skepticism”, he provided an astute and damning rejoinder to the criticisms offered by the principal counter-advocates of the day. Rather than take issue with particular points these commentators had made, he took a more holistic perspective that focused on the rhetorical strategies they had adopted, comparing them with previous objections to parapsychology so as to see how (or indeed if) those arguments had moved on to keep pace with developments in empirical work. This usefully allowed him to identify ways in which the skeptical position had shifted, and drew attention to what they no longer said. For example, the CICAP papers did not claim that the results of the major lines of experimental psi research could be explained in terms of the null hypothesis. “This concession is important”, he noted (p. 191; all page references are to the reproduction in Rao, 1994), “because it shifts the focus of the debate from the *existence* of effects to their *interpretation*” (emphasis in original). And similarly, “they no longer claim to have demonstrated a relationship between methodological flaws and study outcomes”, so restricting the range of legitimate counter explanations.

Chuck also reflected on the exceptional situation in parapsychology in which most counter-advocates were not empirical researchers engaged in psi research, so that their counter-explanations tended to be evaluated on the basis of *plausibility* rather than on the basis of *evidence derived from direct empirical tests*. He wrote,

> Controversies in science normally occur between groups of researchers who formulate hypotheses, develop research methods, and collect empirical data to test their hypotheses. When disputes arise over the interpretation of experimental findings, or when critics suspect the findings were caused by artifacts, they design new experiments to test alternative explanations or the impact of suspected artifacts. It is through this process that scientific controversies are resolved. In contrast, the psi controversy is largely characterized by disputes between a group of researchers, the parapsychologists, and a group of critics who do not do experimental research to test psi claims or the viability of their counterhypotheses. Psi critics argue the plausibility of various alternative hypotheses (or the implausibility of the psi hypothesis) but they rarely feel obliged to test them.

According to Honorton, this lack of empirical engagement has produced a cycle of criticism in response to new claims that begins with *statistical criticisms* intended to demonstrate that the claimed effects are not really significant, *methodological criticisms* that are intended to account for observed effects in terms of procedural flaws, and finally *speculative criticisms* based on a priori and ad hominem arguments. Ironically, in showing how this cycle could be applied to both the

ESP card guessing studies of the 1930s to 1950s and also to the ganzfeld studies of the 1970s and 1980s, Honorton was able to demonstrate a stagnation in parapsychological criticism very similar to the "lack of cumulativeness" that has been regarded by counter-advocates as a principal weakness in the case for parapsychology.

The claim that parapsychology is not progressive is epitomised in a weary mantra that forms part of the skeptic's standard-issue armoury. Honorton (p. 192) describes it thus, "at the core of the critics' current arguments is the rhetoric that 100 years of research has failed to provide convincing evidence for parapsychological phenomena". The refutation of this assertion lies in the data, and Honorton cites the ganzfeld paradigm as providing evidence of cumulativeness, particularly drawing attention to the effects of the joint communiqué on the subsequent standard protocol, and the negligible effect this increased methodological rigour has had on study outcomes. But fundamentally Honorton sees the 100 years mantra as a rhetorical device, reflecting a double standard when judging parapsychology against orthodox research because the implication is that other areas with similar resources have made tremendous advances. He notes, "If we were to apply the 'century of failure' arguments of Hyman and Alcock to academic psychology, we might well conclude that psychology has failed in its mission: after a hundred years of relatively well funded research, vigorous controversies continue over such basic phenomena as memory, learning, and perception" (p 193). Although I think that Honorton's objection is still valid today (cf. Kelly, Kelly, Crabtree, Gauld, Grosso, & Greyson, 2006), I have cause to return to both the use of this mantra and the strategy of applying a double standard later in this paper.

I sense that as he worked through these critical reviews Honorton grew increasingly disgruntled at skeptics' reluctance to engage with the nitty gritty of actual outcomes from actual experiments, perhaps because they could uncover the kind of thorny anomaly that would pose a real challenge to conventional explanation. Instead, Honorton found that they preferred to "offer a caricature of the history of parapsychology and present polemical arguments designed to convince us that there is really nothing in parapsychology that warrants scientific interest" (p. 192).

Modern Skepticism

It is now 25 years since Charles Honorton's untimely death, and in marking that anniversary it seems appropriate to review modern criticism of parapsychology to see to what extent his portrait of skepticism in "Rhetoric over substance" still holds true. During this period, parapsychology has witnessed quite a dramatic transformation in its preferred methodologies and the particular expressions of psi they elicit (see, e.g., Broderick & Goertzel, 2015; Cardeña, Palmer, & Marcusson-Clavertz, 2015; May & Marwaha, 2015), and it has benefited from the availability of more standardised approaches to making sense of accumulated evidence (e.g., Baptista, Derakhshani, & Tressoldi, 2015; Bem, Tressoldi, Rabeyron, & Duggan, 2016; Mossbridge, Tressoldi, & Utts, 2012; Storm, Tressoldi, & Di Risio, 2010) such that one might expect to see similar advances in the nature and scope of skepticism. In the remainder of this paper I should like to test this expectation by focusing on the contributions to Krippner and Friedman's (2010) book, *Debating psychic experience*. This well-regarded collection of essays and rejoinders represents a coming together of the advocate and counter-advocate communities and has much in common with the earlier CICAP volume. It includes critical appraisals of the state of parapsychology from James Alcock, Ray Hyman, Chris French, Michael Shermer, and Richard Wiseman, so I think is a fair representation of the skeptical position in the 21st century. The advocate position is represented by Dean Radin, Chris Carter, and Stephan Schwartz who, although fewer in number, are very capable of giving an informed response to the criticisms raised. In reviewing this material I found a number of recurrent themes that I would like to explore before reflecting on how this might be similar to or different from the state of affairs a quarter of a century ago.

Statistical Problems

Honorton noted that the cycle of objections to parapsychology begins with statistical concerns. In *Debating psychic experience* this focuses especially on a suspicion that meta-analysis is being misused. Hyman (p. 44) argues that "a meta analysis is basically an exploratory rather than confirmatory procedure … [such that] parapsychologists who try to justify the replicability of psi results with meta-analysis are using a retrospective notion". This seems unnecessarily restrictive. For example, it seems clear that some of the variables incorporated into the design of the PRL autoganzfeld trials (most notably the effects on performance of using dynamic versus static targets and sender-receiver pairs who were or were not emotionally close) were derived from the earlier ganzfeld debate and joint communiqué (Bem & Honorton, 1994) and so justifiably any analysis of these variables in the latter meta-analysis could be classed as prospective rather than retrospective.

As in previous publications (e.g., Hyman, 2009), Hyman's deep-rooted suspicions concerning meta-analytic reviews that followed the ganzfeld joint communiqué leads him to eschew them in favour of a single prospective study. He invariably chooses Broughton and Alexander's (1997) series as his case in point. This study failed to replicate the autoganzfeld outcomes, with an overall hit rate of 25.8% (albeit with an emotionally close series that gave a significant hit rate of 37.3%), and he interprets this as compelling evidence that psi effects cannot be replicated. But, as I pointed out to Hyman in an exchange published in the UK's *Skeptic* magazine (Roe, 2009; that is, over a year prior to the publication of *Debating psychic experience*), one could just as easily have chosen as one's prospective study Parker's (2000) automated ganzfeld database, consisting of 150 trials that gave a significant hit rate of 36%, or Dalton's (1997) series of 128 trials with creative participants that gave an even more impressive hit rate of 47%. Which of these should we prefer as our definitive test? None. Although it is tempting to focus on the outcomes of individual studies, particularly when so few people are professionally engaged in parapsychological research and it takes so long to build up anything like a reliable database, we must accept that individual studies are susceptible to giving outcomes that reflect sampling error or are affected by idiosyncratic features of the experimental environment — not obscure or mystical features, but ordinary factors such as experimenter-participant rapport or differences in recruitment strategy. For this reason I believe we must prefer summary reviews when making judgments about the robustness of effects, whatever the shortcomings of this approach.

Incidentally, with respect to sampling error, Chris Carter in his essay in *Debating psychic experience* offers a secondary analysis of the Broughton and Alexander failed replication, stating (p. 160):

> If the true hit rate were 33 percent with 25 percent expected by chance alone [as suggested by Bem and Honorton], then the probability that a sample size of 151 will fail to yield results significant at the 5 percent level is 28 percent. In other words, Broughton's failure to replicate with a sample that small is even less remarkable than flipping a coin twice and getting heads both times.

This does seem to undermine Hyman's argument for making definitive statements based on the outcome of single studies.

Replication Problems

The value that Hyman places on prospective studies clearly derives from expectations he has concerning independent replication. On the first page of his review he asserts that "science can deal only with data and evidence that are objective, lawful and independently replicable" (p. 43) and he evaluates the claims of parapsychology in relation to whether they meet these criteria. One does not need to be a psychic to predict that his conclusion is negative. In his judgement, "parapsychological claims depend upon evidence that is inconsistent and apparently nonreplicable" (p. 47), and later he complains, "somehow parapsychology has managed to continue to pursue its claims even after a long history of inconsistent and nonreplicable evidence" (p. 50). Now, he is characteristically vague when it comes to which lines of research he has in mind when making this claim for inconsistency, perhaps because it is difficult to identify specific examples. Recent mathematical reviews of research on dream ESP (Storm, Sherwood, Roe, Tressoldi, Rock & Di Risio, in press), the ganzfeld (Storm, Tressoldi, & Di Risio, 2010), predictive physiological anticipation (Mossbridge, Tressoldi & Utts, 2012), distant intentionality and the feeling of being stared at (Schmidt, Schneider, Utts & Walach, 2004), and "feeling the future" via implicit precognition studies (Bem, Tressoldi, Rabeyron, & Duggan, 2016) all suggest that effects can be much more robust and consistent than is portrayed here, giving cumulative outcomes that differ from chance expectation to a highly significant degree (I discuss replication in more detail in Roe, 2016a). It seems clear to me that the onus is on counter-advocates to demonstrate that their concerns apply to particular datasets, and where they fail to do so then the criticism should be regarded as a rather hollow rhetorical device.

Interestingly, in his contribution to this volume, Richard Wiseman is more specific, noting that "replicating Rhine's work proved highly problematic, with researchers struggling to find individuals who could reliably produce above-chance scoring" (p 170-1), and this gives us something we can empirically verify. However, his description is at odds with Honorton's (1975) summary of the same material, where he shows that 27 of 33 such experiments were independently significant, when perhaps two might be expected by chance. More of these studies came from outside Rhine's laboratory than from within it, contrary to popular perception and implying independent replication. It seems, then, that where claims are specific enough to be tested they turn out to be untrue.

Another way in which this criticism often remains vague is in failing to prescribe the level of replication that would be regarded as acceptable. Hyman sidesteps this issue by referring to phenomena from psychology that he believes are essentially repeatable on demand. He notes (p. 137)

> in psychology … introductory students can be assigned *hundreds of paradigm experiments* in perception, memory, learning, decision making, and the like with full confidence that the students will obtain the expected outcome. Parapsychology stands alone [he says] as the only discipline claiming to be a science that has not one such experiment. [emphasis added]

Notwithstanding that Hyman is very selective in the areas of psychology from which he draws these paradigmatic experiments (it would have been harder to identify examples from social psychology, health psychology, and personality psychology, let alone clinical or depth psychology), even within this restricted range his claim is simply wishful thinking. It is clearly a very long time since Hyman last ran a research methods workshop in psychology! Those workshops do cherry pick some of the most robust effects so as to be able to focus on learning about the methods of data collection and analysis. Nevertheless, in my 25 years of teaching research methods almost every workshop group includes maverick outcomes, and these are never treated as suggesting the effects are delicate, capricious or suspect.

In his contribution to *Debating psychic experience*, Chris French acknowledges that

> direct replications are quite rare in all areas of science, including psychology. Furthermore, many psychological effects reported in the literature have turned out to be difficult, sometimes impossible, to replicate. This often comes as a complete revelation to new postgraduate students who … are shocked to discover that often they cannot even replicate the basic effects that they intended to investigate in their own research. (pp. 57-58).

This has been my experience. I had one colleague whose cognitive psychology PhD consisted of a suite of 10 experiments of which only 2 gave data that were consistent with the prevailing theory within which she was working. There are no prizes for guessing which of these experiments have been published and which languish in a rarely consulted PhD thesis. Another PhD candidate at my university withdrew from his studies (also in cognitive psychology) because he was unable to reproduce the established effect that he needed to demonstrate before he could go on to look at the effects of varying conditions. It seems clear, then, that replication on demand is an unrealistic expectation for many social science phenomena.

Hyman also portrays psi phenomena as susceptible to decline effects and other inconsistencies that suggest the effects are inherently capricious, citing parapsychologists such as Jim Kennedy (2001, 2003) and Dick Bierman (2001). However, when I considered the issue of replication for a recent issue of *Mindfield* (Roe, 2016a) I actually found that when analysed by Baptista, Derakshani and Tressoldi (2015) replication rates in parapsychology— at least for the ganzfeld — closely matched what they expected given the effect sizes reported and the study power typically involved. I am not persuaded, then, that parapsychological phenomena are so inconsistent or unlawful as to require some special property to explain them.

Indeed, replication rates seem at least comparable to many of those found in more mainstream areas of psychology. The Open Science Collaboration (2015), a coalition of 270 research psychologists, was designed to address the dearth of exact replications published in psychology journals. They agreed to conduct exact replications of published studies to see what proportion of findings could be confirmed independently. To avoid selection bias they chose as their sampling frame articles published in 2008 in three journals that represented premier outlets for psychological research (*Psychological Science*, *Journal of Personality and Social Psychology*, and *Journal of Experimental Psychology: Learning, Memory, and Cognition*). Outcomes were deemed significant if the probability of the observed result was less than or equal to $p = .05$ given the null hypothesis. Unsurprisingly, given publication bias towards positive outcomes, 97 of the 100 original effects were significant, with other results falling in the "suggestive" range, with p values between .05 - .06. Given the statistical power of the replication attempts, they predicted that 89 would be significant, but in fact only 35 met this threshold. Most of the replications (83%) produced a smaller effect size than did the original. This mini decline effect could reflect regression to the mean, or the capitalisation upon chance variation in outcomes due to sampling error exacerbated by a pronounced publication bias towards significant results. (There is no publication bias in the replication studies because all results are reported whatever the outcome.) Baptista and Derakhshani (2014) make a persuasive argument that the replication rate found in the Open Science Collaboration is what one would expect for a database that consists of studies with an average study power of about 35% so that only 35% of the original studies represent true effects, and these have been replicated with high fidelity. Ominously, of course, this would imply that 65% of reported successful outcomes reflect Type 1 errors, and this interpretation has been typical of more mainstream coverage (e.g., Baker, 2015). In terms of how to judge replication levels, this stochastic approach provides, I think, a more realistic context within which to assess parapsychological outcomes than the paradigmatic one suggested by Hyman.

Insinuation of Methodological Problems as Responsible

Nevertheless, in attempting to account for a supposed inconsistency in outcome, skeptics still allude to methodological weaknesses. Sometimes this is expressed quite subtly, so French (p. 56) finds that:

> the best research in parapsychology appears to be at least as scientific as that in other areas of social science, including psychology. Of course, *there are many examples of shoddy research* in all disciplines but it is only fair to judge a discipline on the basis of its highest quality outputs. (emphasis added)

Others speak more baldly. Hyman still refers (p. 143) to the original ganzfeld database as "riddled with flaws", and "as a result of the debate between Honorton and me, the quality of ganzfeld psi experiments has improved greatly. Along with this improved quality, the average effect sizes have also declined" (p. 144). It is indeed true, according to Storm Tressoldi and Di Risio that there is a significant negative correlation between year of study and ES for free response studies ($r_s = .21$, $p = .049$), though they claim that the data are better accounted for by a quadratic model that reflects a rebound effect ($R^2 = .12$, $p = .004$), with the most recent studies showing larger effect sizes, presumably despite continuing improvements in quality.

More tellingly, when discussing his original coding of ganzfeld studies and subsequent debate with Honorton, Hyman offers us a rare insight into the purpose of flaw analysis for him. He admits that,

> By devoting my effort and time towards defending my assignment of flaws and showing how they could have accounted for the effects, I allowed the parapsychologists to switch the burden of proof from their side to my side. It was now up to me to prove that the flaws made a difference. *This is impossible to do on the basis of a single database* (p. 148, emphasis added)

This is a puzzling comment. It is unclear to me what the point would be of identifying some design feature as a *flaw* if it is not possible to demonstrate that this may be at least partly responsible for the supposedly spurious effect. Identifying a flaw as a flaw in some abstract inconsequential sense would relegate Hyman's diligent efforts to the category of Mark Hansel's fantastical suppositions rather than offering a cogent alternative explanation for actual outcomes.

Experimenter/Skeptic Effects

Wrapped up in the issue of replication difficulties and methodological artefacts is a concern about experimenter effects. In summarising his contribution, French asks for a reliable demonstration of psi with a medium to large effect size that can operate in the presence of skeptics. Although of course this request seems eminently reasonable, it simultaneously subtly perpetuates the myth that psi phenomena are coy in the presence of disbelievers, at best suggesting a whimsical or capricious character that sets them apart from other phenomena of the natural world, at worst suggesting that effects arise from methodological errors that are not repeated by more sophisticated skeptical researchers. For example, Alcock claims (p. 31) that "there have been a number of scientists over the years who have taken an interest in parapsychological matters, to the extent that they have conducted their own research but then later abandoned it when their data showed no evidence of paranormality". As a case in point, he refers to two studies by Stanley Jeffers (2003) that failed to find evidence of PK. Embarrassingly, Carter is able to point out (p. 89) that, rather than abandoning psi as a lost cause, Jeffers went on to publish a significant PK effect in his next study (Freedman, Jeffers, Saeger, Binns, & Black, 2003). Conspicuous by their absence from a list of neutral scientists are figures such as Robert Jahn and Daryl Bem, who were initially very skeptical of the psi hypothesis but were persuaded to conduct their own research into psi after being impressed by the results of others and then found out that they could produce their own effects. I suspect that these individuals would be dismissed as "parapsychologists" rather than as initially neutral scientists, but of course that defence would reduce the claim to a tautology. It also seems untenable to me to claim, for example, that the replications of Bem's "feeling the future" suite of experiments from 33 laboratories in 14 countries were all conducted by believers or sympathisers (Bem et al., 2016).

Nevertheless, although I do not believe that skeptics will inevitably fail to find evidence of psi in any experiment they conduct, it also seems obvious that the experimenter can have an effect upon the beliefs, expectations, and behaviour of the participant in ways that affect the outcome, such that it is rather surprising that commentators are suspicious of experimenter effects. I discussed this in some depth in a recent *Mindfield* article (Roe, 2016b) and found that such effects should be expected given the social dynamic interaction between experimenter and participant that affords an opportunity for all sorts of implicit information to be communicated. I noted there the very extensive literature that documents such effects in conventional areas of psychology. Rosenthal (1994, p. 176) refers to 464 studies that demonstrate the effect in a variety of contexts, including studies of reaction time, interpretation of inkblots, animal learning, person perception, and skill development. Expectancy effects have been demonstrated not just in the laboratory, but in classrooms, courtrooms, nursing homes, management settings, and even swimming pools (Rosenthal, 1994, p. 178). It would be astonishing if we were not to find evidence of expectancy effects in parapsychological studies.

Even so, Wiseman writes, "perhaps the most far reaching version of [the] 'get out of the null effect free' card involves an appeal to the 'experimenter effect', wherein any negative findings are attributed to the psi inhibitory nature of the researchers running the study" (cited by Carter, p. 98). This is especially baffling since one of the most compelling demonstrations of the experimenter effect in parapsychology comes from Wiseman himself. He and Marilyn Schlitz had been independently conducting staring detection studies, with Schlitz reporting evidence of psi (Schlitz & LaBerge, 1994) and Wiseman reporting chance outcomes (Wiseman & Smith, 1994). Interestingly they agreed to conduct a series of joint studies in which they would both use the same experimental set-up and draw participants from the same pool to see whether they could replicate their differential performance. The first experiment (Wiseman & Schlitz, 1997) was conducted at Wiseman's Hertfordshire laboratory and involved 16 participants each. Conditions were identical for the two groups except for the identity of the experimenter/starer. The EDA of Schlitz's participants was significantly higher in stare than in no-stare trials, as predicted, whereas the EDA of Wiseman's participants showed no effect, thus confirming their earlier, separate findings. A second experiment (Wiseman & Schlitz, 1999) took place at Schlitz's laboratory at the Institute of Noetic Sciences (IONS) with both experimenters again employing the same procedures, equipment, and participant pool (n = 35 for each experimenter). Schlitz's participants again showed a statistically significant effect, whereas Wiseman's did not, although in this case Schlitz's participants were significantly *less* activated during the stare than nonstare periods, contrary to study one. In a third study (Schlitz, Wiseman, Watt, & Radin, 2006), the design was more complicated so as to tease out the roles of experimenter and starer. The study was again run at IONS and the 100 participants were staff members or local volunteers. The mean effect was somewhat larger when Schlitz was greeter rather than Wiseman, and when Wiseman was starer rather than Schlitz, but none of the effects deviated meaningfully from chance. This is a shame given that the study promised to tease apart two important roles, but with no evidence of psi there was no possibility to explore experimenter effects. In my *Mindfield* article I reflect on interpersonal and procedural differences between Wiseman's and Schlitz's approaches when interacting with participants. Based on Rosenthal's findings it seems clear to me that Schlitz would give rise to positive results and Wiseman to negative ones.

Lack of Interest in Normal Explanations

Alcock complains (p. 38) "there is an obvious lack of interest within parapsychology to explore … non-paranormal explanations. This is regrettable for it adds weight to my contention that parapsychology represents beliefs in search of data, rather than data in search of explanation". This seems an odd claim to make when so many parapsychologists have contributed to the literature concerned with non-paranormal explanations for paranormal belief and experience (such as Carlos Alvarado, Etzel Cardeña, James Houran Harvey Irwin, Rens Lange, Michael Persinger, and Michael Thalbourne). In contrast, there is a surprising dearth of empirical testing of normal explanations by counter-advocates. Honorton's earlier characterisation of the skeptic as an armchair theorist rather than an empiricist comes to mind. Let us take for example cold reading as an explanation for impressive communications from psychics and mediums. The cold reading model describes how a number of techniques can be used either to persuade people that generally true information says something especially or even uniquely true about themselves or to glean information from people nonverbally and verbally that can be fed back later as if from a paranormal source. The model is quite elegant and plausible in some circumstances. It also makes testable assumptions about client behaviour, including their tendency to recall only the hits and forget the misses during a reading, and to elaborate on given material in ways that make the recalled version more specific to them. Surprisingly, to my knowledge the only attempts to test these assumptions have been conducted by a parapsychologist (Roe, 1994, 1995). Rather than seeking experimental evidence for cold reading, skeptical researchers have been content to apply the method after the fact to given "real-world" data in a manner that would be scorned if done by a parapsychologist (e.g., Greasley, 2000; Underdown, 2003; for a more detailed overview see Roe & Roxburgh, 2013).

Rhetorical Ploys 1: Eschewing the Data

My main response to the contributions to *Debating psychic experience* from the most influential and knowledgeable members of the counter-advocate community is surprise at how superficial is their coverage of the parapsychological literature. Admittedly, in essay reviews it is difficult to go into great detail, but their reticence to deal with specifics (in terms of reanalysis of datasets, scrutiny of experimental methods, or elucidation of the effects of particular suspected flaws) seems to me a conscious strategy designed to make it difficult to refute any of the broad claims that they do make. Worryingly, as a consequence the critics show very little familiarity with particular contemporary studies or outcomes.

More commonly, the authors simply make sweeping generalisations with no attempt to evidence or justify them. Thus Hyman claims (p. 195) that every science but parapsychology builds upon its previous data, "the data base continually expands with each new generation, but the original investigations are still included. In parapsychology, the data base expands very little because previous experiments are continually discarded and new ones take their place". And Alcock (p. 32) claims that phenomena have been abandoned when "no solid evidence was forthcoming" only to return at later times with a new generation of enthusiasts. These assertions are much too vague and insubstantial to support any kind of evidence-based response.

Shermer is especially guilty of wilful ignorance of the literature. He asserts (p. 155) "under controlled conditions remote viewers have never succeeded in finding a hidden target with greater accuracy than random guessing" — even a cursory perusal of the literature provides numerous examples of statistically significant remote viewing experiments (cf. Dunne & Jahn, 2003; Roe & Hickinbotham, 2015; Schlitz & Haight, 1984; Targ, 1994). He goes on, "the occasional successes you hear about are due either to chance or suspect experimental conditions, such as when the person who subjectively assesses whether the remote viewer's narrative description seems to match the target already knows the target location and its characteristics. When both the experimenter and remote viewer are blinded to the target, *my analysis of the literature* indicates that psychic powers vanish" (p. 155, emphasis added). Of course, no particular studies are cited in this parodic review since it is unthinkable that any of the studies by the researchers I have just listed would be so naively designed as to not be double blinded. In the circumstances, the idea that Shermer's conclusions are based on a thorough review of evidence is unpersuasive. This type of shallow scholarship cannot be allowed to pass unchallenged.

In this context I can return to the skeptic's mantra I mentioned earlier. Richard Wiseman begins his essay with "after over one hundred years of parapsychological research there exists no consensus on the most fundamental question facing the field — does psi exist?" (p. 169). In an earlier exchange with Hyman, who had also opened with this gambit, I pointed out that clearly the reader is expected to infer that extension over such a long time (in research terms) equates to extensive and intensive activity, so raising their expectations as to the degree of progress that should have been made by a bona fide science tackling real phenomena. Against this, any actual progress seems meagre, and the reader is moved to conclude that the phenomena are non-existent rather than elusive. But of course this inference is unwarranted. Sybo Schouten (1993) demonstrated how unlike other social sciences parapsychology is in terms of the very limited human resources it has available to it, such that the "over one hundred years" amounts to just 53 days of activity in North American psychology during the period that Schouten was writing. I recently discussed this in more detail and offered an update (Roe, 2017), and agreed with Schouten (1993) that, given available resources, progress is at least comparable with that found in psychology.

However, a legitimate issue raised by Wiseman's reference to a lack of consensus is the ambivalence of the parapsychological community with respect to the quantity or calibre of evidence for various anomalies collected under the label *psi*. It is a straightforward matter to distinguish between any effect and its interpretation so we need not be mired in ontological debates and I am beginning to think that the PA has done the field a disservice in not making a position statement with respect to the evidence, one that draws attention to recently published summary evaluations and makes observations regarding the size and consistency of effects in relation to phenomena that have been generally accepted in other areas of social science.

Wiseman also makes the sweeping and unfalsifiable generalisation that parapsychologists conduct very many studies that fail and therefore are never written up for conference presentation and journal publication. This observation is apparently based on personal experience and hearsay. In fact the only "evidence" comes from a study by Watt (2006) of undergraduate projects that unsurprisingly found a strong positive publication bias, with projects that "worked out" being presented at conference and/or submitted for journal publication. Having spent the last 25 years teaching in various university psychology departments, I know this is common practice across all areas of psychology; indeed, it is regarded as a necessary way of boosting research output for overworked teaching staff, who typically co-author these studies. This is not to excuse the practice, which remains highly problematic given how it can serve to skew the database, but it does show that this is not a problem that is peculiar to parapsychology. The description of the issue by Wiseman implies that accepted practice in parapsychology is different from that in psychology and other social sciences when plainly it is not.

Rhetorical Ploys 2: Assuming an Inappropriate or Unachievable Comparator

This implication that parapsychological practice is somehow deficient or substandard takes a number of forms. I found that a common strategy adopted by the skeptics in this volume was to use examples from physics and other natural sciences as exemplars for how normal science is practised, rather than examples from the social sciences. For instance, when explaining independent replication, Alcock writes (p. 36), "scientists may not fully understand what underlies gravitational attraction, but they can predict its effects with great accuracy. All competent researchers will obtain similar results when studying gravitational influences". Likewise, when Hyman claims (p. 48) that the term *anomaly* is used in parapsychology "in ways that differ in important ways from the manner these terms are used by the scientific community", the examples he cites for the appropriate usage are from physics, such as the orbital anomalies that led to the discovery of Neptune, and a disconfirmed anomaly that suggested a planet closer to the sun than Mercury! Later, in thinking of examples of controversial claims that came to be accepted due to the accumulation of confirming evidence Hyman lists "relativity, quantum mechanics, evolution by natural selection, N-rays (sic), and continental drift…" (p 135), while Alcock's rebuttal refers to relativity theory as exemplar of the exactitude needed for a testable theory.

It is not at all clear to me why we should expect parapsychological practice and the behaviour of its phenomena to operate in a similar manner to physical science and its phenomena, given that the former deals with a sentient subject matter and highly complex systems with multiple layers of interacting variables (I allude to this when I discuss standards of replication in Roe, 2016a). Indeed, French concedes that "if psychological research were to be subject to the same level of critical scrutiny, doubt would probably be cast upon many (but not all) of the effects reported". It is disappointing that Hyman and Alcock seem quite incapable of providing even one instance from the history of social science of normative practices or accepted phenomena against which psi research might reasonably be compared.

Conclusion

So what can we conclude about the state of skepticism from this comparison of essays separated by a quarter of a century? My enduring impression is *plus ça change, plus c'est la même chose* (the more things change the more they stay the same). The criticisms laid out against parapsychology in *Debating psychic experience* would not, on the whole, have been out of place in the *Scienza & Paranormale* special issue and, I suspect, would have been treated with similar disdain by Honorton. Indeed, the skeptical positions presented here seem lazily anachronistic given the great changes that have taken place in parapsychology in terms of the research approaches adopted and the bodies of data they have produced. It is quite disturbing to see virtually no mention of presentiment, staring detection, modern mediumship, or implicit psi research. Even for that favourite target of criticism, the ganzfeld (which has 41 index entries), time seems to have frozen with Bem and Honorton's (1994) meta-analysis and responses thereto despite some of the concerns raised in that debate having been addressed in later contributions (e.g., Bem, Palmer & Broughton, 2001; Storm & Ertel, 2001).

Instead we have concerns raised about the precarious nature of replication in parapsychology. But these survive scrutiny only insofar as they remain vague and general, and can exploit an expectation for a kind of replication on demand that is only possible in the natural sciences, if at all (a lesson that psychology is learning as it matures and embraces more stochastic approaches to replication).

On this evidence, it is tempting to conclude that engaging with dyed in the wool skeptics is futile, since there seems no real prospect of constructive dialogue, as seemed possible in the mid 1980s. But just as a government in office needs a discerning opposition to call it to task to ensure good governance, so parapsychology needs a strong and capable counter-advocate movement to ensure that our methods are fit for purpose and our findings valid and meaningful. We need to demand more of our skeptical colleagues otherwise we are partly culpable. As Chuck Honorton warned in his PA Presidential Address (cited by Palmer, 1993, pp. 177-8), "we should not continue to play the game that eventually, after all, science is objective and our findings will eventually become accepted on their merit. I do not believe this. We have been struggling against irrational prejudice for a long time ... if our work is faulty, it should be criticised, but the criticism must be substantive, not a priori".

References

Baker, M. (2015). Over half of psychology studies fail reproducibility test: Largest replication study to date casts doubt on many published positive results. *Nature News*. doi:10.1038/nature.2015.18248

Baptista, J., & Derakhshani, M. (2014). Beyond the coin toss: Examining Wiseman's criticisms of parapsychology. *Journal of Parapsychology*, *78*, 56-79.

Baptista, J., Derakhshani, M., & Tressoldi, P. (2015). Explicit anomalous cognition. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.). *Parapsychology: A handbook for the 21st century*. Jefferson, NC: McFarland & Co. (192-214).

Bem, D. J. & Honorton, C. (1994). Does psi exist? *Psychological Bulletin*, *115*, 4-18.

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

Bem, D., Tressoldi, P., Rabeyron, T., & Duggan, M. (2016). Feeling the future: A meta-analysis of 90 experiments on the anomalous anticipation of random future events. *F1000Research 2016*, *4*, 1188 (doi:10.12688/f1000research.7177.2)

Bierman, D.J. (2001). On the nature of anomalous phenomena: Another reality between the world of subjective consciousness and the objective world of physics? In P. van Locke (Ed.), *The physical nature of consciousness* (pp. 269-292). New York: Benjamins.

Broderick, D., & Goertzel, B. (2015) (Eds.). *Evidence for psi: Thirteen empirical research reports*. Jefferson, NC: McFarland & Co.

Broughton, R.S., & Alexander, C.H. (1997). Autoganzfeld II: An attempted replication of the PRL Ganzfeld research. *Journal of Parapsychology*, *61*, 209-226.

Cardeña, E., Palmer, J., & Marcusson-Clavertz, D. (2015) (Eds.). *Parapsychology: A handbook for the 21st Century*. Jefferson, NC: McFarland & Co.

Dalton, K. (1997). Exploring the links: Creativity and psi in the Ganzfeld. *Proceedings of Presented Papers: The Parapsychological Association 40th Annual Convention*, 119-134.

Dunne, B.J., & Jahn, R.G. (2003). Information and uncertainty in remote perception research. *Journal of Scientific Exploration*, *17*, 207–241.

Freedman, M., Jeffers, S., Saeger, K., Binns, M., & Black, S. (2003). Effects of frontal lobe lesions on intentionality and random physical phenomena. *Journal of Scientific Exploration*, *17*, 651-668.

Greasley, P. (2000). Management of positive and negative responses in a Spiritualist medium consultation. *The Skeptical Inquirer*, *24*(5), 45-49.

Hansel, C. E. M. (1966). *ESP: A scientific evaluation*. New York: Scribners.

Hansel, C.E.M. (1989). *The search for psychic power: ESP & parapsychology revisited*. Buffalo, NY: Prometheus Books.

Honorton, C. (1975). Error some place! *Journal of Communication*, *25*, 103-116.

Honorton, C. (1985). Meta-analysis of psi ganzfeld research: A response to Hyman. *Journal of Parapsychology*, *49*, 51-91.

Honorton, C. (1993). Rhetoric over substance: The impoverished state of skepticism. *Scienza & Paranormale*, 1(3), reproduced in K.R. Rao (Ed.) *Charles Honorton and the impoverished state of skepticism: Essays on a parapsychological pioneer* (pp. 191-214). Jefferson, NC: McFarland & Co.

Hyman, R. (2009). The demise of parapsychology, 1850-2009. *The Skeptic*, *22*(2), 17-20.

Hyman, R., & Honorton, C. (1986). A joint communiqué: The psi ganzfeld controversy. *Journal* *of Parapsychology*, *50*, 351-364.

Jeffers, S. (2003). Physics and claims for anomalous effects related to consciousness. *Journal of Consciousness Studies*, *10*, 135-152.

Kelly, E., Kelly, E.W., Crabtree, A., Gauld, A. Grosso, M., & Greyson, B. (2006) (Eds.). *Irreducible mind: toward a psychology for the 21st century*. Plymouth: Rowan & Littlefield.

Kennedy, J.E. (2001). Why is psi so elusive? *Journal of Parapsychology*, *65*, 219-246.

Kennedy, J.E. (2003). The capricious, actively evasive unsustainable nature of psi. *Journal of Parapsychology*, *67*, 53-74.

Krippner, S., & Friedman, H. L. (2010) (Eds.). *Debating psychic experience*. Santa Barbara, CA: Praeger.

May, E.C., & Marwaha, S.B. (2015) (Eds.). *Extrasensory perception: Support, skepticism and science*. Santa Barbara, CA: Praeger.

Mossbridge, J., Tressoldi, P., & Utts, J. (2012). Predictive physiological anticipation preceding seemingly unpredictable stimuli: a meta-analysis. *Frontiers in Psychology*, *3*, Article 390, doi: 10.3389/fpsyg.2012.00390

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), 943, aac4716-1-8.

Palmer, J. (1993). The psi controversy. In K.R. Rao (Ed.), *Charles Honorton and the impoverished state of skepticism*. Jefferson, NC: McFarland & Co. (177-189).

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

Rao, K.R. (1993) (Ed.), *Charles Honorton and the impoverished state of skepticism*. Jefferson, NC: McFarland & Co.

Roe, C.A. (1994). Subjects' evaluations of a Tarot reading. *Proceedings of Presented Papers: The Parapsychological Association 37th Annual Convention*, 323-334.

Roe, C.A. (1995). Pseudopsychics & the Barnum effect. *European Journal of Parapsychology*, *11*, 76-91.

Roe, C.A. (2009). When is evidence sufficient? *The Skeptic*, *22*(2), 23-25.

Roe, C.A. (2016a). Is inconsistency our only consistent outcome? *Mindfield*, *8*(2), 70-75.

Roe, C.A. (2016b). Experimenter as subject: What can we learn from the experimenter effect? *Mindfield*, *8*(3), 89-97.

Roe, C.A. (2017). Has parapsychology made progress? *Mindfield*, *9*(2), 42-47.

Roe, C.A., & Hickinbotham, L. (2015). Performance at a precognitive Remote viewing task, with and without ganzfeld stimulation. *Abstracts of presented papers: Parapsychological Association 58th Annual Convention & 39th SPR International Conference*, pp. 31-32.

Roe, C.A., & Roxburgh, E. (2013). An overview of cold reading strategies. In C. Moreman (Ed.) *The Spiritualist Movement: Speaking with the Dead in America and around the World: Volume 2, Belief, Practice, and Evidence for Life after Death*. (pp. 177-203). Santa Barbara, CA: Praeger.

Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions* *in Psychological Science*, *3*, 176-179.

Schlitz, M., & Haight, J. (1984). Remote viewing revisited: An intrasubject replication. *Journal of Parapsychology*, *48*, 39-49.

Schlitz, M.J., & LaBerge, S. (1994). Autonomic detection of remote observation: Two conceptual replications. *Proceedings of presented papers: The Parapsychological Association 37th annual convention*, 352–360.

Schlitz, M., Wiseman, R., Watt, C., & Radin, D. (2006). Of two minds: Skeptic-proponent collaboration within parapsychology. *British Journal of Psychology*, *97*, 313-322.

Schmidt, S., Schneider, R., Utts, J., & Walach, H. (2004). Distant intentionality and the feeling of being stared at: Two meta-analyses. *British Journal of Psychology*, *95*, 235-247.

Schouten, S.A. (1993). Are we making progress? In L. Coly & J.D.S. McMahon (Eds.), *Psi research methodology: A re-examination* (295-322). New York: Parapsychology Foundation.

Storm, L. & Ertel, S. (2001). Does psi exist? Comments on Milton and Wiseman's (1999) Meta-analysis of ganzfeld research. *Psychological Bulletin*, *127*, 424-433.

Storm, L., Sherwood, S.J., Roe, C.A., Tressoldi, P.E., Rock, A.J., & Di Risio, L. (in press). On the correspondence between dream content and target material under laboratory conditions: A meta-analysis of dream-ESP studies, 1966-2016. *International Journal of Dream Research*.

Storm, L., Tressoldi, P. E., & Di Risio, L. (2010). Meta-analysis of free-response studies, 1992–2008: Assessing the noise reduction model in parapsychology. *Psychological Bulletin*, *136*, 471–485.

Targ, R. (1994). Remote-viewing replication: Evaluated by concept analysis. *Journal of Parapsychology*, *58*, 271-284.

Underdown, J. (2003). They see dead people – or do they? An instigation of television mediums. *The Skeptical Enquirer*, *27*(5), Accessed from http://www.csicop.org/si/show/ they_see_dead_people_-_or_do_they_an_investigation_of_ television_mediums/

Watt, C. A. (2006). Research assistants or budding scientists? A review of 96 undergraduate student projects at the Koestler Parapsychology Unit. *Journal of Parapsychology*, *70*, 335-356.

Wiseman, R. & Schlitz, M. (1997). Experimenter effects and the remote detection of staring. *Journal of Parapsychology*, *61*, 197-208.

Wiseman, R., & Schlitz, M. (1999). Replication of experimenter effect and the remote detection of staring. *Proceedings of presented papers: The Parapsychological Association 42nd annual convention*, 471–479.

Wiseman, R., & Smith, M. D. (1994). A further look at the detection of unseen gaze. *Proceedings of presented papers: The Parapsychological Association 37th annual convention*, 465–478.

Abstracts in other Languages

French

Adresse Présidentielle PA 2017: Le scepticisme flétrissant

Dans cette adresse, je passe en revue les objections à la parapsychologie développées par des représentants de haut niveau de la communauté sceptique, à travers leurs contributions dans le livre de Krippner et Friedman (2010), Debating psychic experience, en les comparant avec de précédents commentaires critiques publiés en 1993 lors d'un numéro spécial de Scienza & Paranormal consacré à la parapsychologie. Plusieurs des affirmations faites dans ce précédent numéro avaient déjà été discutées par Charles Honorton, dans son article devenu classique : "Plus de rhétorique que de matière : l'état appauvri du scepticisme", et je revisite certains de ses arguments pour évaluer dans quelle mesure le sceptique y a répondus. Je constate qu'en dépit de 25 années écoulées, et les avancées concomittantes de la parapsychologie quant à ses approches, ses méthodes et les données accumulées, le scepticisme montre peu de preuves d'un progrès et pourrait même avoir regressé vers des pratiques rhétoriques encore plus rudimentaires, qui s'appuient sur des aspersions vagues n'ayant que peu de familiarité avec les activités réelles et les découvertes de la parapsychologie contemporaine.

German

PA-Präsidentenansprache 2017: Vernichtender Skeptizismus

In dieser Ansprache untersuche ich die Einwände gegen die Parapsychologie, die von führenden Vertretern der skeptischen Gemeinschaft in ihren Beiträgen zum Buch von Krippner und Friedman Debating psychic experience (2010) geäußert wurden, und vergleiche sie mit früheren kritischen Kommentaren, die 1993 in dem der Parapsychologie gewidmeten Teil des Sonderheftes der Zeitschrift Scienza & Paranormale erschienen sind. Viele der Behauptungen in dieser früheren Sammlung wurden bereits von Charles Honorton in seiner klassischen Erwiderung "Rhetoric over substance: The impoverished state of skepticism" [„Mehr Rhetorik als Substanz: Der dürftige Zustand des Skeptizismus"] herausgefordert, und ich würdige einige dieser Argumente erneut, um zu prüfen, inwieweit der Skeptizismus darauf eine Antwort gefunden hat. Ich bin der Meinung, dass ungeachtet der mittlerweile vergangenen 25 Jahre und der damit einhergehenden Fortschritte, was Zugänge, Methoden und akkumulierte Daten der Parapsychologie betrifft, der Skeptizismus nur geringe Zeichen von Fortschritt zeigt und in Wirklichkeit eher zu verkümmerten rhetorischen Praktiken zurückgekehrt ist, die eher auf vagen Anschuldigungen beruhen und kaum Vertrautheit mit den aktuellen Aktivitäten und Ergebnissen zeitgenössischer Parapsychologie erkennen lassen.

Spanish

Ponencia Presidencial de la PA, 2017: Escepticismo marchito

En esta ponencia considero las objeciones a la parapsicología de representantes de alto calibre de la comunidad escéptica en sus contribuciones en el libro de Krippner y Friedman (2010), Debating psychic experience, y las comparo con el comentario crítico publicado en 1993 como parte de un número especial de Scienza & Paranormale dedicado a la parapsicología. Muchas de las afirmaciones hechas en esa colección anterior ya habían sido cuestionadas por Charles Honorton en su clásico ensayo titulado " Rhetoric over substance: The impoverished state of skepticism". Reviso algunos de esos argumentos para medir hasta qué punto el escepticismo ha respondido a ellos. Me parece que a pesar del paso de 25 años y de los avances concomitantes en los enfoques, métodos, y datos acumulados en la parapsicología, el escepticismo muestra poca evidencia de progreso y de hecho puede haber regresado a prácticas retóricas más rudimentarias basadas en aspersiones vagas y muestra muy poco familiaridad con las actividades y hallazgos reales de la parapsicología contemporánea.