Journal of Parapsychology, 80(1), 118–132

CORRESPONDENCE

To the Editor:

The papers (Parker, 2015a, 2015b) to which Dr. Cardeña refers in his letter published in the Fall 2015 *JP* (Cardeña, 2015) were intended to be constructive critiques of all the recent work on the hypnotic state and on hypnotic psi. It is rather against the spirit of these papers that Cardeña dismisses the review as being "slanted," having "innuendos," and being "inaccurate." I also regret that Dr. Cardeña refers to a debate over the loss of the Freiburg chair and its funding for psi research, but he does not inform the reader of my detailed reply to this rather different issue (Parker, 2013).

It is for me not meaningful to spend time on snide remarks and diversions to other issues. Nor is it so meaningful to spend time on who first coined specific terms within the hypnosis area, but it has to be said categorical statements are not in order. For instance, concerning the coining of the word "hypnotism," it is true that "D'Henin de Cuvillers used various terms beginning with 'hypno,' as for example hypnobate, hypnoscopie, hypnotique, hypnologie but while he may have been getting close, none of his terms caught on" (A. Gauld, personal communication, October 21, 2015). I should however thank Dr. Cardeña for spending valuable time on finding amongst 30 pages of text that "dissociation" in DES should have actually been written "dissociative" and what was actually a "he" should indeed have been written as "she." I also thank Dr. Cardeña for making clear his use of "wolverine" which I had misunderstood.

Nevertheless, I take the above tone of animosity to indicate Dr. Cardeña thinks he was unfairly treated. I wish therefore to look at the relevant complaints and see if any of them have substance and can be resolved fairly. I did say I thought the study had a "handful of participants," and it was "overloaded with variables and hypotheses" that were tested on small subgroups. The "handful of participants" refers of course not to the starting number, which was 26, but rather to the fact that some of the main hypotheses were evaluated using five to nine participants. Frankly, I would find it difficult to get any support from statisticians for drawing conclusions on *F* tests with cell frequencies as low as five. I confess that I cannot understand why Dr. Cardeña does not agree with me. Instead, Dr. Cardeña countered by saying that our ganzfeld studies had only a few more participants than his starting quota of 26. He will find on reading the relevant studies concerning psychological variables that these were based on 50 participants (Parker, 2000; Parker, Grams, & Pettersson, 1998).

The overall outcome of the Marcusson-Clavertz and Cardeña (2011) ganzfeld study with the 26 participants gave 27% as the hit rate, which is of course very near mean chance expectancy. It is never stated (beyond saying it was process research) why significant overall scoring was not a main hypothesis and why there was a failure to replicate previous findings. Instead of this, the psi scores were analyzed in relationship to numerous variables: hypnotizability, dissociation, eleven dimensions of the PCI, and the dimensions of belief in psi. The description "overloaded with variables and hypotheses" is then surely not too "inaccurate."

In his defense, Dr. Cardeña advertises that he has one study that used even more variables than were "accepted into the very rigorous *Journal of Experimental Psychology*." A recent evaluation of its replication rate (Open Science Collaboration, 2015) suggests it might not be so rigorous. Whatever the case, surely the malpractice of others is not a good argument—and certainly not something for parapsychology to emulate.

In the review of the Marcusson-Clavertz and Cardeña study, it was stated that the "sheep-goat effect" was the one positive finding. This is conventional shorthand for what Dr. Cardeña means when he says two of their hypotheses were confirmed: Hypotheses 1 and 2, namely, those concerning "believing that one would be 'successful' in the experiment" and "having previous psi experiences." If so we are in agreement. Since we also agree Hypothesis 3 (concerning psi scores of those with high versus low hypnotizability) was not confirmed, this leaves Hypothesis 4.

Hypothesis 4 was that dissociation would modify the effect of hypnotizablity on psi scores. Cardeña still claims this was confirmed. Other than a negative relationship between hypnotizability taken alone and the psi score, I cannot find anything to support this claim.

Hypothesis 5 concerned predicting a relationship between psi scores and the reporting of altered states (in the ganzfeld). However, the only significant finding, as interesting as it is, occurred only *within the highly hypnotizables* and is clearly post hoc. Moreover, the questionnaire used to evaluate the altered state in the ganzfeld session was assessed in another (mind-wandering) experiment, which seems a little problematic.

Dr. Cardeña further maintains the choices of statistics used for testing of all the hypotheses were all specified in advance. Maybe they were, but where? What is written on page 244 does not appear to have been prespecified. Likewise, when it concerns the nonparametric tests, Cardeña refers us again to page 244 of his original text and such tests are certainly mentioned there, but here again, where are the results of these tests? The only exception I can find concerns the above finding that I do not dispute, namely the sheep-goat test.

Dr. Cardeña rejects my suggestion that they did not identify which were the post hoc findings: He replies this is "false since we described our hypotheses in the paper and the reader can then precisely determine which are the post-hoc findings." However given the vagueness of some of the hypotheses and the multiplicity of the tests used, to insist on this surely presupposes some form of extraordinary ability on the part of the reader in order to discern what is what. Magnus Fontes, professor of statistics at Lund University, totally dismissed the whole study, expressing a radically opposite opinion on the claims that were then made and which Dr. Cardeña still continues to make (Fontes, 2013). The same was true of Georg Lindgren, the statistician included amongst the nine professors who published a critical article titled "Pseudoscience is Spread Uncritically" (Halle et al., 2012), which might easily have led to the loss of the Lund chair.

It is of course true, as Dr. Cardeña says, that it is now possible to register the means of testing hypotheses in advance via the Edinburgh KPU study registry, but the debated study was carried out prior to the registry and is not to be found there.

Finally, Dr. Cardeña asks for a reference concerning the possible bimodality of the Dissociative Experiences Scale (DES). The term bimodality can be, on reflection, misleading here, but I was wanting to summarize the Wright and Loftus (1999) review, which concluded there are two different populations with their own means—those with dissociative disorders and those without. Given that Cardeña's reply suggests that we agree that the DES is a highly skewed scale and is aimed at identifying pathology, then why use it on students when there are plenty of other good tests of dissociation aimed at measuring healthy experiences?

At the end of his reply, Dr. Cardeña thinks it unfortunate that I had not discussed "integrative, multifactorial models," but perhaps there is room for agreement here too. Although the model he seems to prefer, namely Shor's model (and I do apologize for once making a typical Swedish typo in spelling Shor's name), is of early 1970s vintage, it should be said that some of the more skeptical contemporary theorists are indeed now moving slowly towards this type of theory of hypnosis (see, e.g., Lynn, Laurence, & Kirsch, 2015). I had hoped that the very titles and themes of my two papers would communicate just that: Hypnosis is a jungle of variables—or if you will, multifactorial—which makes it exceedingly difficult to combine with psi research, as I believe Dr. Cardeña's work illustrates. The papers I wrote focused finally on the specific ways in which research in this complex multifactorial area can go further.

References

Cardeña, E. [Letter to the Editor]. *Journal of Parapsychology*, *79*, 252–255.

Fontes, M. (2013, April) *A parapsychological experiment in Lund. Debate presentation: Pseudoscience—an innocent game or a serious parasite*. Paper presented at the Chalmers University Science Festival, Gothenburg, Sweden.

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

Lynn, S. J., Laurence, J-R., & Kirsch, I. (2015). Hypnosis, suggestion, and suggestibility: An integrative model. *American Journal of Clinical Hypnosis, 57*, 314–329.

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

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science, 349*, 1095–9203. Retrieved from: http://www.sciencemag.org/content/349/6251/aac4716
Parker, A. (2000). A review of the ganzfeld work at Gothenburg. *Journal of the Society for Psychical Research, 64,* 1–15.
Parker, A. (2013). Betraying the future by distorting the present: A reply to Bauer et al. *Journal of the Society for Psychical Research, 77*, 125–127.
Parker, A. (2015a). The jungle of hypnotic psi: Part 1. Research on hypnosis relevant to psi. *Journal of Parapsychology, 79*, 20–36.
Parker, A. (2015b). The jungle of hypnotic psi: Part 2. Research on relationships between psi and hypnosis. *Journal of Parapsychology, 79*, 37–52.
Parker, A., Grams, D., & Pettersson, C. (1998). Some further variables relating to psi in the ganzfeld. *Journal of Parapsychology, 62*, 27–45.
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.

Adrian Parker

Department of Psychology
Gothenburg University
Box 500
40550 Gothenburg, Sweden
adrian.parker@psy.gu.se

To the Editor:

Dr. Parker continues to attack our 5-year-old publication (Marcusson-Clavertz & Cardeña, 2011) and also my reply (Cardeña, 2016) to his criticisms. Regarding his remark that his papers were "intended to be constructive critiques," anyone who is aware of his more than decade-long public and private history of attacks against me will be able to evaluate his likely intention.

Now to more substantial issues:

1. Dr. Parker maintains that it is not "so meaningful" to spend time pointing out the various typos, misinterpretations, and inaccuracies in his original papers, but besides correcting the record they provide a potential indicator of how careful and accurate may be the rest of the content.
2. His statement that "some of the main hypotheses were evaluated using five to nine participants," is highly misleading. We had five main hypotheses (Marcusson-Clavertz & Cardeña, 2012). Hypothesis 1, belief in self-success in the psi task will be positively related to psi scoring, was confirmed with all 26 participants. The same was true of Hypothesis 2, which postulated that previous psi experiences (not the same as the "sheep-goat" effect, contrary to what Dr. Parker states) would also correlate with psi scoring. For Hypothesis 3, we conducted a *t* test between high and low hypnotizables (a correlational analysis would be inappropriate because the medium hypnotizables were excluded), comparing cells of 14 versus 12. For Hypothesis 4, contrary to Dr. Parker's claim that it "was confirmed," we actually wrote on page 246 that "Neither was there a significant interaction between hypnotizability and dissociation." This is the *only* hypothesis-testing analysis for which we had cells of between five and nine. Incidentally, Table 6 of one of his papers shows that Dr. Parker (Parker, Grams, & Pettersson, 1998, p. 330) analyzed cells of $n = 8$ and $n = 10$, besides carrying out many analyses on multiple variables. Finally, for Hypothesis 5, which proposed that greater alterations in consciousness would relate to high psi performance and hypnotizability, the significant result with a very strong size effect was with the group of 14 high hypnotizables, not a cell of "five to nine."
3. As for not predicting that participants overall would significantly psi hit, we wrote on p. 244, "Evaluating overall psi was not a target of this process-oriented research." Choosing participants

that we thought would be psi-hitters along with those we thought would be chance-scorers or even psi-missers made it oxymoronic to predict overall success. As Bem, Palmer, and Broughton (2001, p. 215) wrote "genuine progress in understanding psi rests on investigators' being willing to risk replication failures by modifying the procedure in any way that seems best suited for exploring new domains or answering new questions."

Regarding other misleading statements by Dr. Parker:

1. He mentions that we did not specify "which were the post-hoc findings," but on page 247, we stated that analyses with other PCI dimensions than the *altered state* one "can be considered exploratory."
2. As for preregistering our study, at that time it was not common practice and the PKU registry was not even started until the fall of 2012, so if Dr. Parker wants to blame us for that lack of precognition, perhaps he could start listing his own previous studies with preregistered analyses. And, as I mentioned in my previous reply, the samples for each one of his studies (Parker et al., 1998) was 30. Whether he combined his data later for some analyses is irrelevant to my statement.
3. Contrary to what Dr. Parker seems to hint, an indication that we were aware of the strengths and limitations of the DES and other dissociation instruments is that the American Psychiatric Association tasked me some years ago to write a review chapter of dissociation measures (Cardeña, 2008).
4. Dr. Parker questions that the *Journal of Experimental Psychology: Learning, Memory, and Cognition* is a rigorous journal because a recent Open Science Collaboration paper found problems with replicability of some studies (incidentally, these findings have also been disputed, see Gilbert, King, Pettigrew, & Wilson, 2016), despite its 78% rejection rate (http://www.apa.org/pubs/journals/features/2013-statistics.pdf) and 2.86 impact factor. Sadly Dr. Parker offers no references to journals in which he has published with better statistics than these.
5. Finally, Dr. Parker might cite and agree as much as he wants with the antipsi people here in Sweden, but his statement that their criticism "might easily have led to the loss of the Lund chair" goes against the public statements of support I received in the media by Lund University's then-President Per Eriksson (Stiernstedt, 2012), the current Dean of my College Ann-Katrin Bäcklund (Anonymous, 2012), and the then-Chair of the Department, Per Johnsson (Fagerström, 2012; see also Cardeña, 2015).

Is there anything I can agree with Dr. Parker on? Well, of course our study had a number of limitations, which we listed on page 252 and which Dr. Parker repeats, including a small *N* and using the measure of alteration of consciousness with the same people under ganzfeld but not from the psi session. As with most research in any field, it should not be considered definitive but as evidence supporting some hypotheses and not others, and offering some promising new leads. So I hope that he will agree with me that it will be more constructive for everyone concerned to spend time conducting their own preregistered new research with large *N*s rather than pursuing this debate.

References

Anonymous (2012, October 31). Experiment om tankeöverföring [Experiment on telepathy]. *Studio Ett*. Retrieved from http://sverigesradio.se/sida/artikel.aspx?programid=1637&artikel=5330277

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.

Cardeña, E. (2008). Dissociative disorders measures. In A. J. Rush, M. B. First, & D. Becker (Eds.), *Handbook of psychiatric measures* (2nd ed.) (pp. 587–599). Washington, DC: American Psychiatric Press.

Cardeña, E. (2015). The unbearable fear of psi: On scientific censorship in the 21st century. *Journal of Scientific Exploration, 29,* 601–620.

Cardeña, E. (2016). [Letter to the Editor]. *Journal of Parapsychology, 79,* 242–245.

Fagerström, E. (2012, April 19). Ett decenium i vetenskapens gransland. [A decade on the borders of science]. *Sydsvenskan.* Retrieved from http://www.sydsvenskan.se/2015-04-19/ett-decennium-i-vetenskapens-gransland

Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on "Estimating the reproducibility of psychological science." *Science, 351,* 1037–1038. doi:10.1126/science.aad7243

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

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

Stiernstedt, J. (2012, July 14). "Vi studerar tomtar och troll också." ["We also study brownies and dwarves"]. *Svenska Dagbladet.* Retrieved from http://www.svd.se/vi-studerar-tomtar-och-troll-ocksa

Etzel Cardeña

Department of Psychology
Lund University
Paradisgatten 5P
Lund 223 50, Sweden
etzel.cardena@psy.lu.se

To the Editor:

Few recent parapsychological experiments have given rise to as much acrimony as a study in this journal by Dr. David Marcusson-Clavertz and Prof. Etzel Cardeña (2011; hereafter DM/EC). They reported that, for high hypnotizables *only*, there was a significant correlation between psi *z* scores and being in an altered state of consciousness. Later in an interview, EC was quoted as having said that this correlation is ". . . a very strong indication that telepathy has really occurred" (Oredsson, 2012, p. 17). Vociferous exception was registered by a number of Swedish academics (Halle et al., 2012). They emphasized that the overall hit rate was in fact nonsignificant (27% direct hits, chance = 25%) and considered the correlation an artifact of multiple analyses: The study was profiled as pseudoscience. For most English-speaking readers this controversy lies largely behind the (Swedish) language barrier (but see Cardeña, 2013a, 2013b).

More recently, in the course of a major review of psi and hypnosis research in this journal Prof. Adrian Parker (2015) characterized the DM/EC study as "overloaded with variables and hypotheses" (p. 41). Cardeña (2015) has disputed aspects of Parker's criticism but did not really settle the underlying issue. Is the DM/EC inference from the data really justified?

The recent statement by the American Statistical Association (Wasserstein & Lazar, in press) highlights the gulf between "statisticians' statistics" and "researcher's statistics" and brings the vexing problem of multiple analyses into the limelight. Every additional analysis is basically an extra shot at the significance jackpot and the strict logic of statistical significance testing requires that a "family-wise" (or similar) analysis is performed per study; only then is the probability value calculated correct for the *set* of analyses, rather than for the individual analysis. This does not mean that only a single variable should be investigated per study; rather the *set* of *p* values must be corrected appropriately (e.g., Bretz, Hothorn, & Westfall, 2011).

Most researchers do not formally correct probabilities for multiple analyses. If a *large number* of analyses are performed this has unfortunate consequences: (Virtually) every such study contains sufficient (spurious) "significances" to be published as "evidence of some effect." It is unlikely that such a study will come up totally empty with "no evidence for an effect." Karl Popper (1959 /1992) proposed that the essential characteristic of scientific theories is that they are *falsifiable* (demarcation criterion). Excessive uncorrected analyses "immunize" against the possibility of falsification.

With 100 analyses per study the (binomial) probability of some "significances" is as high as 99.4% and the corresponding empty hands (no "sigs") is a negligible .6%. For 14 independent analyses per experiment just more than half of the studies are expected to be "publishable" merely by capitalizing on chance. Stacking the odds in the researcher's favour by multiplying analyses is decidedly not "playing the game."

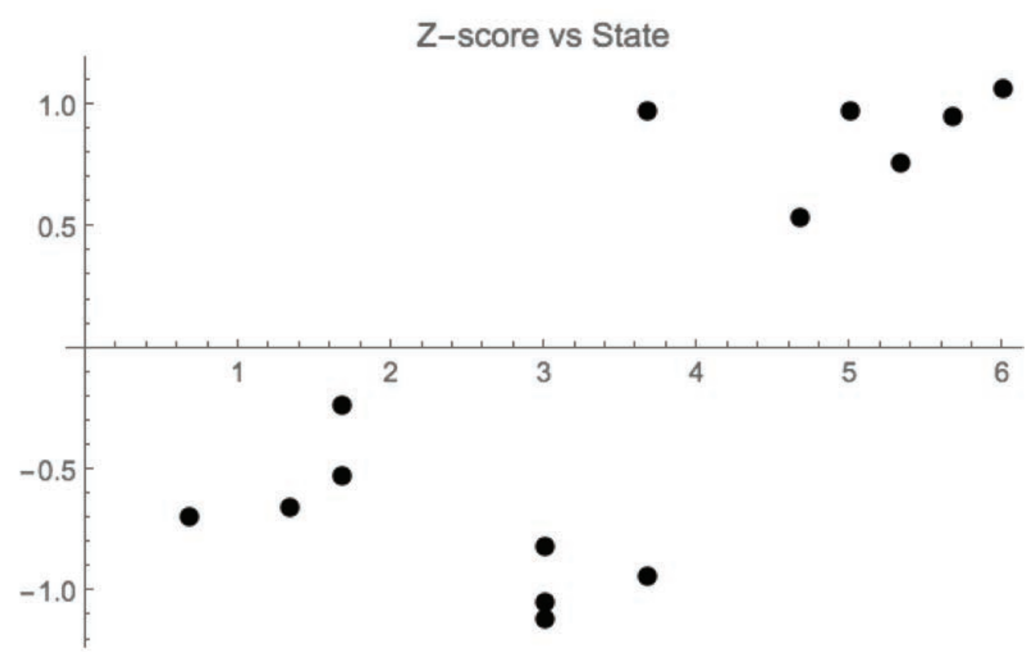
In lieu of formal correction for multiplicity, researchers early devised a rough and ready workaround. Two broad classes are distinguished—confirmatory and exploratory studies. For confirmatory studies the number of analyses is typically restricted to a few and significance testing can be meaningful (if taken with the appropriate amount of salt). In exploratory research, the net is spread wide, over every variable that might be relevant. Exploratory studies do not *prove* anything: the idea is simply to get some idea of what's what. For such studies statistical probabilities are essentially meaningless (unless the total number of analyses is known). For exploratory work "significant" findings are, at best, hints which have to be followed up by rigorous confirmatory studies. Some researchers, however, speciously regard "significance" ($p < .05$) as a "magic marker" for a real effect, *regardless* of how the analysis came to be carried out. And the mandatory publication ritual is to report such nugatory "significances" even for exploratory studies.

In the DM/EC paper there is no discussion of the distinction between exploratory and confirmatory studies. Is the study exploratory or confirmatory, or perhaps some mixture of the two? Were sufficient analyses actually performed to make cherry picking a plausible explanation for most (or all of) the results of the DM/EC experiment? In the following I prefer the term "selection" (with scare quotes) to avoid unintended negative connotations.

In reply to Parker (2015), Cardeña (2015) states: "As far as 'overloading' of hypotheses, we had a big total of five" (p. 253). There is, however, no mention of the statistics to be employed and each can be tested in many ways, limited only by the ingenuity of the analyst. Unless an *exact* specification is made, the researcher is free to select (unwittingly) whichever "equivalent" analysis is most significant. EC maintains that "... the reader can ... precisely determine which are the post-hoc findings" (p. 253). But this reader is not alone in being unable to identify them unambiguously. It would have saved a great deal of puzzling if each analysis had been clearly labelled (confirmatory or exploratory). Were there, as EC implies, only five preplanned analyses, with the remainder (about 95%) post hoc (exploratory)? In Tables 1 and 3 alone there are as many as 78 analyses reported and there are many others in the text. A minimum of a hundred analyses must have been carried out in all—substantially more (one hopes) than in a typical parapsychological study. Analyses outnumber participants (26) by something like 4 to 1. Although the study seems to be very much of the exploratory type, DM/EC do not so much as mention the impact of multiple analyses: A single Bonferroni correction is reported in passing (p. 244).

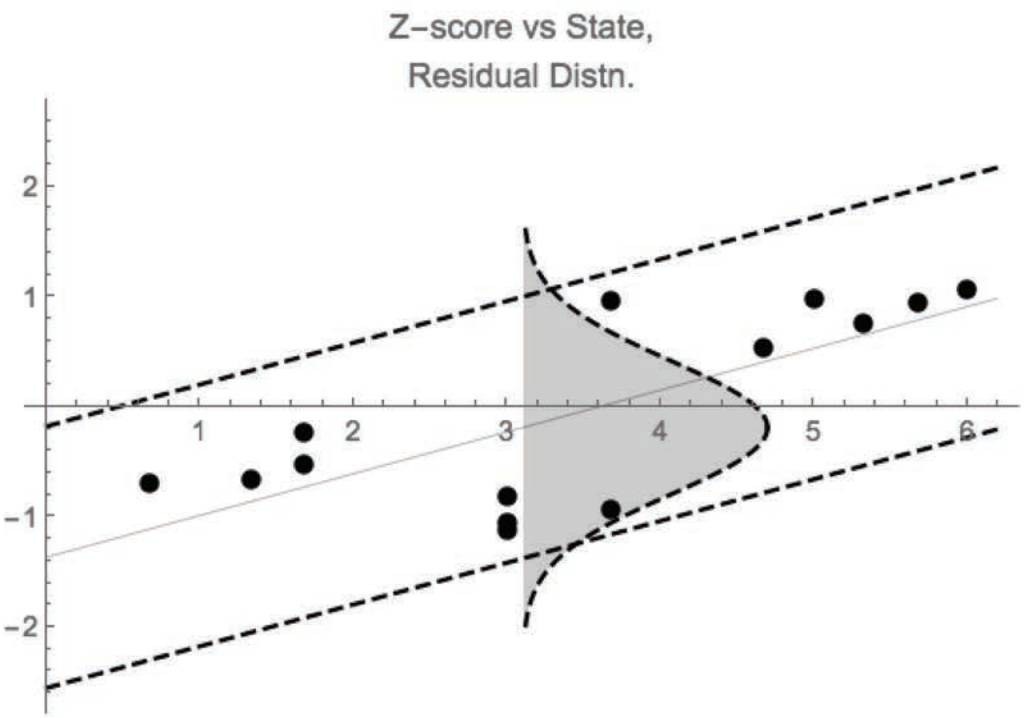
In the first column of Table 1 are a total of 15 correlations with psi z scores, six accompanied by "significance stars." However, when corrected for multiple analyses using the Holm method (R package "p.adjust") only one remains significant. This is the altered state/psi correlation, which DM/EC regard as their prime finding: "Generally, our results may be the clearest evidence until now of a relationship between experiencing an altered state of consciousness and psi-hitting . . ." (Marcusson-Clavertz & Cardeña, 2011, p. 251).

The raw data for high hypnotizables are regraphed as Figure 1.



Inspection reveals that "psi-hitting" palpably does NOT increase with increasing state change. The *z* score actually changes from *missing* for low state change to *hitting* for larger changes. This is difficult to interpret in terms of participant state change. Here the slope is significant (p = .002), whereas the mean *z* score for the DM/EC psi data is -.05, very near zero.

1. The most negative (regression) *z* score is (perversely) at zero state change: This intercept is significantly negative (p = .003). The most positive *z* score at maximum state change cancels this out (overall chance scoring).
2. A peculiarity less obvious to the eye is that many of the data points seem to be just "too close" to the regression line. This residual variance is 0.37, only about a third of the expected value for a *z* score (1.0): This is significantly low (p = .025).



The pattern is very similar to what may arise due to experimenter effects. In the simplest case, the experimenter has a hypothesis that a given experimental (E) condition is associated with the occurrence of psi and compares this with a control (C) condition. A significant E/C difference is interpreted as supporting the hypothesis, but closer examination reveals C is about as much *below* chance as E is above and together they cancel out. For a continuous independent variable, as here, there is a significant correlation without any overall psi. There is actually a little evidence in parapsychology suggestive of an experimenter effect with hypnotized people. In one experiment (Fahler & Osis, 1966) confidence calls produced a highly significant positive deviation, whereas the remainder gave an almost equal negative deviation, the classic mirror pattern.

Although an experimenter effect is theoretically a possible explanation for the results of the DM/EC study, much more plausible is run-of-the-mill "selection." To make things concrete, imagine that with everything else being the same as in DM/EC, random sampling from a standard normal replaces the empirical psi *z* score measures. A linear regression is done for each of a very large number of these thought experiments and those with nominally significant slopes are selected. The properties of these "selected" experiments will be systematically different from the great mass of pure random regressions and not just in that they have significant slopes.

It is easily seen that such "selection" accounts for the zero-straddling (change from missing to hitting) nature of the claimed DM/EC regression, because "selection" for slope does not change the (zero) mean. The total variance too is not influenced by selection for slope; but this is the sum of regression and residual components, and since those with large regression are selected, there is systematically less residual variance. The "selection" hypothesis accounts neatly for all of the observations above, in particular the very peculiar deficit of residual variance.

I have implemented the above thought experiment for "selection" as a numerical computer simulation. Although details lie outside the scope proper for a Letter, I note here that the mean simulation regression line is virtually the same as that from the DM/EC study.

Many (perhaps most) parapsychologists "want to believe." For a confirmatory experiment, statistical testing imposes some restraint. In an exploratory study, this check on "fooling oneself" is missing and uncorrected statistical probabilities actually seem to support belief. Did DM/EC deceive themselves into believing one thing they wanted to see in the extensive array of analyses they laboriously carried out, like tea-leaf reading?

Uncorrected multiple analyses is the one "questionable research practice" (QRP) to be found, to some degree, in nearly *every* published study. Parapsychology too has its fair share of such studies (a single putative example is Honorton, Davidson, & Bindler, 1971). The DM/EC study is no worse than many others but *no worse is not good enough*.

About 70% of psychological studies published in the premium journals have recently proven unrepeatable in other hands (Open Science Collaboration, 2015). While long suspected, it is still a blow to find that two thirds of psychology findings now appear to be spurious. For too long researchers have relaxed the rigour theoretically required for experiments and some QRPs have even become the norm: The chickens have come home to roost. Current methodological standards in the behavioral sciences are woefully inadequate for experimental psychology itself. Tighter procedural and statistical standards of evidence are required for more contentious areas, such as parapsychology.

Dr. John Palmer (2013) has set out his editorial policy with respect to multiple analyses and whether statistical corrections should be made. Like most journals this is regarded as the responsibility of the individual author. The general problem, however, cannot effectively be tackled on that level: Preregistration of experiments and planned analyses are the way to go. Will this make positive (parapsychological) findings vanish? This does seem to be the case for clinical trials testing heart disease treatments (Kaplan & Irvin, 2015).

In a forensic context, eyewitness testimony has been extensively studied (Loftus, 1996). In general, testimony is surprisingly unreliable: Factors which play an important role include reconstructive memory as well as confirmation bias. In theory, all analyses should be preplanned in full detail before an experiment is ever begun. But in practice (for other than very simple designs) this does not reflect the messy reality of actually carrying out a study: Updates are continually made based on earlier results. I have seen this taken to the extreme of taking the complete data of a finished experiment "upstairs" to the resident SPSS expert, who is expected to deliver publishable "significances." Much more subtle (and common) is the distorting effect of reconstructive memory on slowly changing predictions: The prediction update is "written over" the original memory and the researcher is prepared to swear this is *the* "preplanned" analysis. Published analyses often bear little relationship to the original plan (cf. Millar, 1980; Tart, 1980).

Many a researcher feels his scientific integrity attacked if publicly suspected of cherry picking, and Cardeña (2015) characterizes Parker's suggestion as innuendo. But it is increasingly recognized that researchers are not the infallible authority figures of the late 19th Century; rather they are subject to the same human frailties as the people they study. The DM/EC study was carried out over a considerable period (two years for participant selection alone) and was interwoven with other work. Long drawn out and segmented activity seems particularly conducive to the subversive effect of reconstructive memory. No special precautions are mentioned to ensure that any analyses planned did, in fact, remain constant over time. The bottom line is that only preregistered analyses should be taken at full face value. Happily, EC has largely preempted suggestions of where to go from here by planning a followup study, for which the critical analyses to be performed have already been preregistered.

In summary, the DM/EC report follows the style of an exploratory study: Confirmatory analyses are not clearly distinguished from exploratory ones. "Significant" results claimed are tenuous if they are considered against a background of considerable multiplicity of analyses. In particular, the flagship correlation of psi *z* scores with state of consciousness measures bears telltale traces of "selection" rather than the signs of a real effect of the participants' state of consciousness.

References

Bretz, F., Hothorn, T., & Westfall, P. (2011). *Multiple comparisons using R*. New York, NY: CRC Press.
Cardeña, E. (2013a). A Swedish skirmish. *Mindfield, 5*, 4.
Cardeña, E. (2013b). From the editor's desk. *Mindfield, 5*, 40–41.
Cardeña, E. (2015). [Letter to the Editor]. *Journal of Parapsychology, 79,* 252–255.
Fahler, J., & Osis, K. (1966). Checking for awareness of hits in a precognition experiment with hypnotized subjects. *Journal of the American Society for Psychical Research*, *60,* 340–346.
Halle, B., Hesslow, G., Karlström, G., Lidin, S., Lindgren, G., Löfstedt, C., Nilsson, D., . . . Svensson, B. E. Y. (2012, October 31). Pseudovetenskap sprids okritiskt [Pseudoscience spreads uncritically]. *Svenska Dagbladet.* Retrieved from http://www.svd.se/
Honorton, C., Davidson, R., & Bindler, P. (1971). Feedback-augmented EEG alpha, shifts in subjective state, and ESP card-guessing performance. *Journal of the American Society for Psychical Research, 65,* 308–323.
Kaplan, R. M., & Irvin, V. L. (2015). Likelihood of null effects of large NHLBI clinical trials has increased over time. *PLoS ONE* , *10*(8), 1–12.
Loftus, E. F. (1996). *Eyewitness testimony*. Chicago IL: Harvard University Press.
Marcusson-Clavertz, D., & Cardeña, E. (2011). Hypnotizability, alterations in consciousness, and other variables as predictors of performance in a ganzfeld psi task. *Journal of Parapsychology, 75*, 235–259.
Millar, B. (1980). Rejoinder to Dr. Tart. *European Journal of Parapsychology, 3,* 114–116.
Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*, aac4716.
Oredsson, U. (2012). Han sänder sina tankar [He sends his thoughts]. *LUM, 7*, 16–17.
Palmer, J. (2013). JP publication policy: Statistical issues [editorial]. *Journal of Parapsychology, 77*, 5–8.
Parker, A. (2015). The jungle of hypnotic psi: Part 2. Research on relationships between psi and hypnosis. *Journal of Parapsychology, 79*, 37–52.
Popper, K. (1992). *The logic of scientific discovery.* New York, NY: Routledge (Original work published 1959).
Tart, C. T. (1980). [Letter to the Editor]. *European Journal of Parapsychology, 3,* 111–112.
Wasserstein, R. L., & Lazer, N. A. (in press). ASA statement on statistical significance and *p*-values. *American Statistician*.

Brian Millar

Denpassar, Bali
bmillar@ymail.com

To the Editor:

I am flattered that one of our papers (Marcusson-Clavertz & Cardeña, 2011) has provoked not only one (Parker, 2015; Parker, this issue) but two reactions, currently the letter by Dr. Parker's sometime collaborator Dr. Millar (Parker & Millar, 2014). I will first clarify some points and then describe our points of agreement and disagreement.

Regarding the quotation in the employee magazine *LUM* by Ms. Oredsson (2012), what I mentioned to her in our interview was that we had found a "large correlation" (a technical term for correlations greater than .50) between psi *z* scores and an altered-state scale. She thought that the general audience would not understand the technical term and chose the phrase of "a very strong indication of telepathy," a phrase which cannot be found in our original study. As anyone who has interacted with the media knows, one may ask a reporter to use more precise language but they may choose a different route. In a response to the Halle et al. (2012) attacks, I mentioned to newspapers that the evidence for psi in the ganzfeld rested in ganzfeld meta-analyses rather than in the findings of a single study such as ours (e.g., Cardeña, 2012).

As for the criticism by Halle et al (2012) about the overall close to chance hit rate (27%) in our study, they had reason to know of its lack of validity because before they published their letter I had sent them a copy of our paper, in which we specifically stated on page 244, "Evaluating overall psi was not a target of this process-oriented research." How could it be when we chose to include two groups (low hypnotizables and low dissociators) whom we thought might be random scorers or even psi-missers?

Now, with regard to Dr. Millar's own misinterpretation of our work, he stated that "In the DM/EC paper there is no explicit mention of the distinction between exploratory and confirmatory studies." Darned if on page 247 we did not write "We estimated the correlations between psi *z* scores and all of the PCI major dimensions. Except for the altered state dimension already discussed, these analyses can be considered *exploratory* (emphasis added)." Thus, the confirmatory analyses referring to the dependent variable (psi *z* scores) are reported only in Tables 1 and 2, not in 1 and 3 as he implies, and encompass 20 analyses for the dependent variable blocked according to the factors in our hypotheses. Millar writes rhetorically "Were there, in fact, only five pre-planned analyses...?" To which the answer is yes. We designed the experiment to have one dependent variable (*z* scores) and a few independent variables, mentioned in our hypotheses: hypnotizability, dissociation (as mediator or moderator of hypnotizability), belief in success in the experiment, previous ostensible psi experiences, and an overall experience of being in an altered state of consciousness. Our hypotheses were not post-hoc elucubrations but were based on previous research findings by others mentioned in our paper, including: (a) a meta-analysis showing a relation between success in a psi task and belief in one's own experimental success (Lawrence, 1993); (b) a report of a positive relation between reported ostensible psi experiences and experimental success (Honorton, 1997); and (c) a meta-analysis showing participants doing better in a hypnosis than in a nonhypnosis context (Stanford & Stein, 1994); and we had reason to predict that the former would be more likely to impact high hypnotizables. To our knowledge there had not been studies of dissociation and psi performance (as compared with reports of ostensible psi in everyday life), but there is previous research showing that hypnotizability and dissociation interact in important ways (e.g., Terhune & Cardeña, 2010), and thus it made sense to hypothesize that dissociation would mediate or moderate the effect of hypnotizability. Finally, the hypothesis that has seemed to most discombobulate the critics of our study, that experiencing an altered state of consciousness would correlate with psi scoring, was not picked out of thin air (or post-hoc findings), but was based on a number of previous studies, including one by Palmer, Khamashta, and Israelson (1979) with meditators (who are likely to perform similarly to high hypnotizables given their ability to focus) in which a scale representing experiencing an altered state of consciousness correlated positively with success in the ganzfeld psi task.

As far as Millar's remark about our not discussing the impact of multiple analyses (he curiously omits to mention the multiple analyses of his sometime-collaborator, e.g., Parker, Grams, & Pettersson, 1998), we were aware of it, which is why we reported only the hypothesis-driven analysis as confirmatory and did one Bonferroni correction to analyze a question that was similarly phrased to another. The usefulness of the Bonferroni and similar procedures is for unplanned, post-hoc comparisons, including those when the overall *F* is not significant but the researchers want to still carry out contrast analyses (Keppel, 1982; Rosenthal & Rosnow, 1985), which was not the case for our confirmatory analyses. And, of course, there is always the issue of balance between avoiding Type I (i.e., falsely reporting a significant difference) and Type II errors (i.e., falsely not reporting a significant difference). An exclusive regard with the former can lead to strangulating the generation and testing of alternative hypotheses (Fiedler, Kutzner, & Krueger, 2012). It is ironic that in his "corrected" Holm-method analysis, Dr. Millar continued to find support for the finding he most inveighs against, namely experiencing an altered state and psi scoring.

In any case, I believe that a number of statisticians would agree with me that statistical analyses, including possible corrections, should not be applied automatically and mindlessly, but Dr. Millar seems to have fallen prey to the worship of the ("corrected") *p* value, while ignoring calls for the "new statistics" that also or instead consider other indexes, including effect sizes such as the correlation values in our study (cf. Cumming, 2013). Dr. Millar states that "statisticians' statistics" would demand a correction of *p* values for more than one "family-wise" analysis per study, but there is not even consensus about what the ideal correction may be, or even if it is always a mandatory step (cf. Fiedler et al., 2012; Keppel, 1982; Rosenthal & Rosnow, 1985). And as mentioned elsewhere in this Letter, we did conduct a Bonferroni correction when we thought it was justified; the result he objects to remains significant even after different corrections, and we specified which were hypothesis-testing and which were exploratory analyses. Statisticians and psychologists may prioritize different aspects, with the latter being more involved in the theoretical context and previous empirical support relevant to a particular area of research, as well as with the potential generation of valuable hypotheses.

Dr. Millar implies that our result concerning the large correlation ($r = .74$, $p < .01$), which remains significant after his correction and explains more than half of the variance between *z* scores and the altered-state scale, is artifactual, perhaps the product of "the subversive effect of reconstructive memory" about what analyses we had planned, or worse. Actually, as mentioned earlier, our hypotheses and corresponding analyses were based on previous research and theory; had we "cherry-picked" our significant analyses, we would have had a perfect set of matching hypotheses and significant results, which clearly was not the case.

Having said that, as we also mentioned in our paper, we had a small *N*, which demands a larger effect difference to reach significance but makes the latter less reliable than results obtained with a larger *N* (Button et al., 2013). I do agree with Dr. Millar that preregistering a study, particularly in such a controversial area as parapsychology, is a good idea and may reduce innuendoes and "analytical retrocognitions" and assist us in establishing more reliable relations. In fact, we preregistered our follow-up study (both the KPU and the Open Science Framework registries were started later than our study) once the KPU was created. After all, it is not only the original researchers but also their critics who are vulnerable to the "human frailties" that Dr. Millar mentions. Having said that, I disagree with his conclusion that a failure to replicate many psychology (and parapsychology) experiments necessarily shows that they are spurious (cf. Open Science Collaboration, 2015). Failures to replicate may be due to the impact of methodological, social, and other contextual issues in research (Barrett, 2015). For instance, when replications matched more closely the initial study according to the original investigator in the Open Science Collaboration, the replications were considerably more successful (Gilbert, King, Pettigrew, & Wilson, 2016); for the case of ganzfeld psi research, see Bem, Palmer, and Broughton (2001).

Finally, despite Dr. Millar's statement that the positive correlation between psi *z* scores and the altered-state scale "bears characteristic traces of 'selection'," in the replication study that we preregistered, which we will soon submit for publication, we found again a significant relation between them.

References

Barrett, L. F. (2015, September 1). Psychology is not in crisis. Retrieved from http://www.nytimes.com/2015/09/01/opinion/psychology-is-not-in-crisis.html?_r=0

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.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafo, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*, 1–12.

Cardeña, E. (2012, October 31). Starkt statistiskt stöd för parapsykologi. (Replik) [Strong statistical support for parapsychology (Reply)]. *Svenska Dagbladet.* Retrieved from http://www.svd.se/starkt-statistiskt-stod-for-parapsykologi).

Cumming, G. (2013). The new statistics: Why and how. *Psychological Science, 27,* 7–29. Retrieved from www.tiny.cc/tnswhyhow

Fiedler, K., Kutzner, F., & Krueger, J. I. (2012). The long way from α-error control to validity proper: Problems with a short-sighted false-positive debate. *Perspectives on Psychological Science, 7*, 661–669.

Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on "Estimating the reproducibility of psychological science." *Science, 351,* 1037–1038. doi:10.1126/science.aad7243

Halle, B., Hesslow, G., Karlström, G., Lidin, S., Lindgren, G., Löfstedt, C., Nilsson, D., . . . Svensson, B. E. Y. (2012, October 31). Pseudovetenskap sprids okritiskt [Pseudoscience spreads uncritically]. *Svenska Dagbladet.* Retrieved

from http://www.svd.se/
Honorton, C: (1997). The ganzfeld novice: Four predictors of initial ESP performance. *Journal of Parapsychology, 61*, 143–158.
Keppel, G. (1982). *Design & analysis. A researcher's handbook* (2nd ed.). New York, NY: Prentice Hall.
Lawrence, T. R. (1993). Gathering in the sheep and goats: A meta-analysis of forced choice sheep-goat ESP studies, 1947–1993. *Proceedings of Presented Papers: The Parapsychological Association 36th Annual Convention*, 75–86.
Marcusson-Clavertz, D., & Cardeña, E. (2011). Hypnotizability, alterations in consciousness, and other variables as predictors of performance in a ganzfeld psi task. *Journal of Parapsychology, 75*, 235–259.
Open Science Collaboration, Estimating the reproducibility of psychological science (2015). *Science, 349*, aac4716. doi:10.1126/science.aac4716 pmid:26315443
Oredsson, U. (2012). Han sänder sina tankar [He sends his thoughts]. *LUM, 7*, 16–17.
Palmer, J., Khamashta, K., & Israelson, K. (1979). An ESP ganzfeld experiment with transcendental meditators. *Journal of the American Society for Psychical Research, 73,* 333–348.
Parker, A. (2015). The jungle of hypnotic psi: Part 2. Research on relationships between psi and hypnosis. *Journal of Parapsychology, 79*, 37–52.
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 seem to succeed by using their own psi. *Journal of Parapsychology, 78,* 39–85.
Rosenthal, R., & Rosnow, R. L. (1985). *Contrast analysis: Focused comparisons in the analysis of variance.* Cambridge, UK: Cambridge University Press.
Stanford, R. G., & Stein, A. G. (1994). A meta-analysis of ESP studies contrasting hypnosis and a comparison condition. *Journal of Parapsychology, 58*, 235–269.
Terhune, D. B., & Cardeña, E. (2010). Differential patterns of spontaneous experiential response to a hypnotic induction: A latent profile analysis. *Consciousness and Cognition, 19*, 1140–1150.

Etzel Cardeña

Department of Psychology
Lund University
Paradisgatten 5P
Lund 223 50, Sweden
etzel.cardena@psy.lu.se

To the Editor:

I enjoyed reading Loyd Auerbach's review of William Hall's book, *The World's Most Haunted House: The True Story of the Bridgeport Poltergeist on Lindley Street* in the Fall 2015 *JP*. As one of the investigators involved in that study, I want to record my own thoughts about it and about the book. One of our cultural icons is a good ghost story. This has not gone unnoticed by the media, the entertainment industry, book publishers, and a growing list of "psychic investigators" and "ghostbusters" with varying credentials and intentions. And then there's the scientific field of parapsychology, as represented by the Parapsychological Association (PA), a composite of credentialed scientists with publications in peer-reviewed science journals who attempt to apply scientific methodologies to explain, model, or predict the phenomena at the heart of a good ghost story. These are two different worlds; they are superficially similar but in fact vastly different.

One major difference between these two worlds is how they go about their business—and, indeed, what their business is! Viewing the two approaches side-by-side in the Bridgeport case highlights the contrasts. More importantly, it opens a discussion to which various investigators of possible RSPK cases might contribute.

The book focuses upon a highly publicized RSPK case in Bridgeport, Connecticut in 1974, when William J. Hall was a child of 10 growing up in that city. He has a vague recollection of the case and recalls that it was quickly "solved" by the local police as being a hoax. I was a research associate at Duke University's Psychical Research Foundation (PRF) at that time. Boyce Batey, a friend of the PRF who lived near Bridgeport, called us and convinced Bill Roll, Keith Harary, and myself that the case was not a "hoax," that it was worth investigating, and that the police chief (who'd publically closed the case) would welcome and even help facilitate our investigation, as would the focal family. Keith and I went there as soon as we could to make initial contact with Boyce, the police, and the family. I stayed on for another few weeks, interviewing the numerous people who'd been involved. Thus, the factual skeleton of data upon which William J. Hall's book hangs was largely constructed by me.

Years later the adult Hall took an interest in "things that go bump in the night" and embedded with the "psychic investigator," "ghostbuster" community. He created a syndicated news column, Magic and the Unknown, and searched for "proof" that ghosts are "real"; but these adventures generally ended in disappointment. When he was reminded of the Bridgeport case by a Facebook post, Hall took a closer look and found there was more than a "hoax" involved. He soon discovered a virtual treasure trove of readily accessible, high quality media accounts published between November 25 and 29, 1974, including detailed eye-witness reports of large scale events from police, a neighboring fireman, and journalists. He decided to reopen the investigation even though it was almost 40 years later. Hall soon discovered Boyce Batey, who had preserved the original investigative data from the case, including a lengthy reel-to-reel audio tape on which I had recorded many of my official interviews. Out of Hall's investigative efforts came this book.

To set the tone for reading this book, keep in mind that the blockbuster movie, *The Exorcist*, had opened in December, 1973! A year later, in November of 1974, crowds were still lining up at theatres in Bridgeport (and elsewhere) to see it. So it was a no-brainer that on Sunday morning, November 24, when the largest local Bridgeport radio station (WNAB) announced on air that there were strange, unexplained, exorcist-like activities going on at a house at xxx Lindley Street, there would be a literal traffic jam of rubber-neckers lining Lindley Street before long. In addition, numerous police, firemen, neighbors, and friends of the family, as well as radio, TV, and newspaper journalists, descended upon the now public address where the little family of three opened the door to all of them. There were reportedly dozens of people swarming around inside the house for much of that day, even while TVs, large recliner chairs, kitchen shelves, tables, wall-hangings, lamps, mirrors, and other large furniture flipped, flopped, and flew mysteriously about in view of all.

This zoo-like atmosphere continued into the next day until the next night, when two officers caught the 11-year-old girl who lived there perpetrating one rather insignificant false "event." The case was immediately declared solved and was closed. The Tuesday morning newspaper carried the headline, "FAMILY 'HAUNTED' NO LONGER; COPS SAY GIRL TELLS OF HOAX." But, as I discovered in the following weeks by interviewing many of the numerous eyewitnesses (the tape-recordings that Hall obtained from Boyce Batey), this case was more than a simple hoax.

In retrospect, the book's title—*The World's Most Haunted House*—may in fact be well-deserved. The number of eye-witnesses alone sets it apart from other RSPK cases. Although spontaneous observations are typically the most questionable form of data, a number of the events were quite compelling, due to multiple witnesses observing from different perspectives while all potential perpetrators were absent or otherwise controlled. Some events were witnessed by people with no expectancy set, such as friends or neighbors dropping by without having heard the radio reports.

Although the scrutiny of events reported to occur in RSPK cases usually dominates the spotlight, it's only a part of the story. The other part, the most important part from the family's perspective, is the extraordinary disruption and emotional turmoil that are the poltergeist's constant companions. Boyce Batey and I knew this part first-hand because we came to know the Goodin family. Hall had only second-hand knowledge of this disruptive factor. So I applaud him for his insistence on doing the best he could to capture this aspect of the situation in his book.

What Hall has done well is to take a simple chronological approach to the basic data in the first half of the book. This makes the sequence of events easy to follow. Whenever Hall drifts away from this chronology we move into relative chaos. Also, Hall is to be commended for being intent upon getting the facts of the case straight. I can myself attest to that, both from the book and from my personal interactions with Hall.

The second half of the book is a disaster, in my view. This may reflect a key difference between ghostbusters and scientific parapsychologists. Hall appears to buy into the ghostbusters' fallacy—that they are (he is) doing scientific research. He himself does not see that the worlds of the ghostbuster and scientist are quite separate, and it shows here. Starting with Chapter 13, on The Scientific Investigation, the writing disintegrates into a hodge-podge of random ideas, quotes, and ramblings that are disconnected, irrelevant, and often just a distraction from the central theme. And what is the central theme of this book and investigation? Hall, the ghostbuster, has only *one single ill-defined research question*—"Is it real?" He states his answer early on:

> …what happened in November 1974 at that now-iconic little house was "real" beyond any doubt. I found myself struggling to prove it a hoax. I reached a tipping point where the hoax story—what most people would call the most logical of stories, was the hardest—impossible, in fact, to prove at all. (p. 16)

This is where Hall departs (in my humble opinion) from science. This brings me back to my primary goal in writing this Letter. I will outline four specific issues this book highlights for me.

First of all, no single case with uncontrolled observations can be proven as "real"! There are always alternative explanations. (Think of a magic show where you were unable to explain a levitation or disappearance—did that "prove" levitation was "real"?). This is, at best, a single observation, an "*n*" of 1, in a nonexperimental study composed entirely of *post hoc* data.

Second, although this case has the earmarks of a valid RSPK case, that's the beginning of science, not the end. Hall has done a very creditable job of carefully presenting the facts of the case in the book and dealing with the hoax hypothesis. This is a prerequisite for the case to be entered into the database with other similar cases with credible witness reports. It is from such a database of hundreds of cases, spanning centuries in time, that we now recognize common features, such as rapping, object movements and disappearances, the concept of a "focal" person or "agent," the commonality of a tense family environment, and features suggestive of human involvement from which the RSPK rubric has emerged (Maher, 2016; Roll, 1976).

At this stage of the game, RSPK is of necessity "bottom-up" science in which we painstakingly add the details of carefully screened cases to the database in hopes of inching closer to an understanding of these fantastical phenomena. In contrast, many ghostbusters employ a "top-down" approach based upon an unfalsifiable hypothesis—that ghosts did it.

Third, in this scenario science is best served by the systematic collection of as much detail as possible concerning each event, not just the hoax negation (which can never be 100% eliminated, post hoc, for any individual event) but the circumstances, such as time of day, duration, number, and identity of all witnesses, location of the suspected agent and other family members, and even such things as temperature, weather, and activity in the neighborhood. Since RSPK cases often involve some faux events, each event needs to be considered independently. It would be ideal if each event could be rated by the investigating team for its degree of causal uncertainty.

Fourth, attempting to bring the RSPK events to a halt and following up afterwards may also contribute to science. The Bridgeport RSPK activity ended, in my opinion, when the family tension was relieved, which supports—but certainly does not "prove" —the RSPK model. At present, we have little data concerning the life trajectories of poltergeist agents after the phenomena have ceased, but I suspect this could be very useful for science to track.

So, for the scientist, eliminating the hoax theory is just and only that, and you are now left with nearly as much of a mystery as you started with. Thus, in science, concluding "it's not a hoax" is not the same as "it's real," *whatever that means*!

When I returned to Durham, PRF director W. G. Roll, Keith Harary, Boyce Batey, and I compiled a short paper on the Bridgeport case, which we presented at a scientific conference (Solfvin, Harary, & Batey, 1976). That was the only report we ever published of the case, because we did not think there was anything in it which advanced our scientific understanding of this fascinating unexplained phenomena. That is, it was another "good" case—but just another—to add to our accumulating database.

Overall, I am delighted that Hall took the trouble to resurrect the Bridgeport poltergeist case. Interacting with Hall on the telephone prior to publication, and reading the book after publication, has reinvigorated my own interest in these fascinating family phenomena. And looking backward in time, despite the limitations of post hoc analysis, I find there are a few things that I would have done a bit differently had I known then what I know today. Perhaps someday I'll write more on that topic!

References

Auerbach, L.. (2015). [Review of the book *The world's most haunted house: The true story of the Bridgeport poltergeist on Lindley Street*, by W. J. Hall]. *Journal of Parapsychology, 79*, 240–242.

Maher, M. (2016). Ghosts and poltergeists: An eternal enigma. In E. Cardeña, J. Palmer, & D. Marcussen-Clavertz (Eds), *Parapsychology: A handbook for the 21st century* (pp. 327–340). Jefferson, NC: McFarland.

Roll, W. G. (1976). *The poltergeist*. Metuchen, NJ: Scarecrow Press.

Roll, W. G. (1977). Poltergeists. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp. 382–413). New York, NY: Van Nostrand Reinhold.

Solfvin, G., Harary, B., & Batey, B. (1976). A highly publicized case of RSPK [Abstract]. *Journal of Parapsychology*, *40*, 48–49.

Jerry Solfvin

Center for Indic Studies
University of Massachusetts Dartmouth
285 Old Westport Road
North Dartmouth, MA 02747, USA
www.umassd.edu/indic/
jsolfvin@umassd.edu