

Journal of the Society for Psychical Research

Volume 81, Number 2

April 2017

- Exploring Precall Using Arousing Images and Utilising an
On-Line Memory Recall Practice Task.....*David J Vernon* 65
- Gathering in More Sheep and Goats: A Meta-Analysis of
Forced-Choice Sheep-Goat ESP Studies, 1994–2015
Lance Storm & Patrizio E. Tressoldi 79
- A Controlled Study of Psychometry Using Psychic and Non-
Psychic Claimants with Actual and False Readings Using a
Mixed-Methods Approach
Ian S. Baker, Jane Montague and Abigail Booth 108

BOOK REVIEWS

- Ghosts: A Haunted History* by Lisa Morton
Christian Jensen Romer 123
- Parapsychology: A Beginner's Guide* by Caroline Watt
Bryan J. Williams 125
- Psychology and the Near Death Experience: A Search for God*
by Roy L. Hill..... *Erika A. Pratte* 130
- The Self Does Not Die: Verified Paranormal Phenomena*
From Near-Death Experiences by Titus Rivas,
Anny Dirven and Rudolf H. Smit *Guy Lyon Playfair* 132
- Near Death Experiences: Understanding Visions Of The*
Afterlife by John Martin Fischer & Benjamin Mitchell-Yellin
Robert A Charman 133
- Ghostly Encounters: The Hauntings of Everyday Life*
by Dennis Waskul with Michele Waskul..... *Fiona Bowie* 139

CORRESPONDENCE

- Letters from Clay Thistleton and Carlos S. Alvarado142
- Notice regarding *Journal* editorship144

VOLUME 81, NUMBER 2

JOURNAL OF THE SOCIETY FOR PSYCHICAL RESEARCH

APRIL 2017

NOTES FOR AUTHORS

The Journal of the Society for Psychical Research

The Society for Psychical Research was founded in 1882 “to examine without prejudice or prepossession and in a scientific spirit those faculties of man, real or supposed, which appear to be inexplicable on any generally recognised hypothesis”. In keeping with most scientific bodies, the Society holds no corporate views, and any opinions expressed in this Journal or the Society’s other publications are those of the authors alone. However, for over a century the Society has published an impressive body of evidence for the existence of such faculties and the occurrence of so-called paranormal phenomena.

Editorial Board

Prof. John Poynton, MSc, PhD (Chair)

Carlos S. Alvarado, BS, MS, MA, PhD

Prof. Stephen E. Braude, PhD

Richard S. Broughton, BA, PhD

Prof. Bernard Carr, MA, PhD

Prof. Deborah Delanoy, BA, PhD

Alan Gauld, MA, PhD, DLitt

James Houran, MA

Prof. Adrian Parker, MA, PhD

Adrian Ryan

Stefan Schmidt, PhD

Paul Stevens, BSc, PhD

Prof. Donald West, MD, LittD, FRC Psych

Carl Williams, BA, PhD

Prof. Robin Wooffitt, BSc, PhD

Editor

Prof. Chris A. Roe, BSc, MSc, PhD, AFBPsS

Editorial Assistant

Charmaine Sonnex, BSc, MSc

Review Editor

Tom Ruffles, BSc, BA, MA, PhD, ARPS

The annual subscription to the *Journal*, which is published quarterly, is £40/\$88. This includes any Proceedings published. More information is available on the Society’s website at: www.spr.ac.uk

Issue Number 927

ISSN 0037–9751

© Society for Psychical Research

The Society for Psychical Research, 1 Vernon Mews, London W14 0RL

Telephone 020 7937 8984

The Editor of the *Journal of the Society for Psychical Research* welcomes original submissions in English on any aspect of psychical research so long as it is followed with sufficient rigour and awareness of the wider academic literature. The Editorial Board will consider any empirical approach, including (but not restricted to) experimental work, field investigations, case collections, questionnaires and interviews, as well as historical or theoretical work that can demonstrate an original contribution. Submissions are considered on the understanding that they have not been published before and are not concurrently under review by any other journal. All submissions are subject to anonymous review by subject experts as identified by the Editorial Board and will be accepted or rejected based on the recommendations of those reviewers. The email address for submissions is journal@spr.ac.uk and we encourage electronic submissions as Word, rich text, or PDF files.

Organization of Manuscript Papers should not normally exceed 7,000 words, which should include a short abstract of 50–100 words that describes the key aspects of the paper. Authors are responsible for the preparation of artwork for figures and these and any photographs submitted should be of quality suitable for reproduction.

In the preparation of manuscripts, authors are expected to adhere to APA style (see <http://www.apastyle.org/> or refer to previous issues of the *JSPR* — example copies can be provided on request). Note in particular that APA style requires that citations are given in the body text by author surname and date (e.g. “Smith & Jones, 2016”) and that a references list is included that includes all cited work listed in alphabetical order by author’s name. All periodical titles should be given in full.

Footnotes These should be indicated in the text by superscript numerals at the appropriate places, numbering consecu-

tively. The notes should be placed at the foot of the text page concerned, separated from the main text by a line. Although footnotes may include citations, they should not be used as a way of referencing others’ work.

We invite Letters to the Editor that comment on or otherwise develop ideas introduced by material published in the Journal. These should not exceed 1,000 words of text and 10 references.

Publication If the paper is accepted the author or first-named of joint authors will be sent copies of the typeset pages for checking and correction prior to printing (but proofs are not usually sent for letters). On publication a PDF copy of the work will be sent to the lead author.

COPYRIGHT

Authors retain copyright to their writings and take responsibility for any opinions expressed. However, if an article has been accepted for publication in the *Journal of the Society for Psychical Research* then the Society reserves the right to reproduce the material as part of a published journal, and to make it available in electronic as well as print formats. The material must not appear anywhere else (including on an Internet website) until it has been published by the Journal (or rejected for publication). After publication in the Journal, authors may use the material as they wish but should make appropriate reference to its prior publication in the Journal; for example: “This paper (material) first appeared in the *Journal of the Society for Psychical Research*, vol..., pp..., and is reproduced here with their permission”.

Permissions Authors are responsible for securing written permission to reproduce any published or unpublished material for which copyright is currently held. It is not normally necessary to secure permission to publish short quotations from published works.

© Society for Psychical Research, 2017
0037–9751 / 17



Established 1882

EXPLORING PRECALL USING AROUSING IMAGES AND UTILISING AN ON-LINE MEMORY RECALL PRACTICE TASK

BY DAVID J VERNON

ABSTRACT

The idea that future practice can somehow influence current behaviour has been examined within the paradigm of precognition. Previous work attempting to examine possible precognitive effects using a modified repetition priming task showed that participants were more accurate to respond to material they would see again in the future. Such a finding was taken to indicate that a task relying primarily on accuracy of performance, such as a memory recall task, could be a more sensitive measure of precognition, or 'precall'. Furthermore, utilising arousing images as opposed to everyday words may elicit a stronger precall effect, and by conducting such a study on-line it may be possible to reduce or even eliminate any potential experimenter effects. The prediction when completing such a task was that post-recall practice would lead to greater precall of those items practised compared with items not practised. Such an on-line precall study utilising emotive images was completed by 94 participants. However, comparison of the accuracy between images that were subsequently repeated and those that were not showed no evidence of a precall effect. Post recall practice did show an initial improvement in accuracy, but this plateaued after the second trial. The failure to find any evidence of precall could simply be indicative of the impossibility of such a notion. However, given that others have reported precall effects, the failure to find one in this study is discussed in terms of possible methodological factors inhibiting psi performance.

INTRODUCTION

Precognition refers to the ability to perceive and/or behave in a way that is influenced by a future event that would not be anticipated through any known inferential process (see, Mossbridge, Tressoldi, Utts, Ives, & Radin, 2014; Taylor, 2014). Such behaviour may be conscious or unconscious, cognitive or affective, and has been examined utilising a variety of methods. Over time, this has included Zener cards, dice, light configurations, as well as possible changes in physiology (see, Cardeña, Marcusson-Clavertz, & Palmer, 2015; Irwin & Watt, 2007). More recently, Bem (2011) utilised a range of standard psychology paradigms that incorporated a time reversed element to test for potential precognitive effects. These included a mere exposure paradigm, repetition priming, habituation and recall paradigms. In all cases the idea was that an individual's performance would be influenced by exposure to and/or processing of a stimulus after an original response has

been made and recorded. For instance, Bem's time reversed recall task would examine the positive effect on memory recall performance for items that are practised after the recall responses have been made. Given the commonplace finding that practice or rehearsal of items improves subsequent recall performance (see, Bahrnick, 2000) the question posed here is whether practice after the recall performance would have a reversed time influence such that recall performance is improved for those items practised compared with those that are not. Such an effect is generally referred to as a precall effect. Thus, precall represents the positive effect on memory recall performance for items that are practised after the recall test.

Bem (2011) reported on nine experiments that tested precognition across a range of paradigms and found that eight out of the nine studies showed evidence of a significant precognitive effect. Interestingly, the non-significant study did show a trend in the predicted direction. Furthermore, the studies utilising the precall type paradigm showed the most robust effect sizes (mean ES = 0.31) compared with the habituation (0.15), or priming (0.25) studies. However, attempts by a number of independent research teams have yielded conflicting results (see, e.g., Galak, LeBouf, Nelson, & Simmons, 2012; Ritchie, Wiseman, & French, 2012), though modifications in the precise methodology may to some extent account for such differences (see, Bem, Tressoldi, Rabeyron, & Duggan, 2015). Given the variability in findings, some have suggested utilising meta-analysis as a possible way of attempting to resolve such debates and identify possible generalised effects across samples, stimuli and protocols. For instance, early work by Honorton and Ferrari (1989) examining possible precognitive effects from 1935 to 1987 reported a small but significant effect of precognition. More recently, Storm, Tressoldi, and Di Risio (2012) examined studies from 1987 to 2010 and, though they found a slightly lower effect size, reported results that were still significantly above chance expectations. Finally, Bem et al. (2015) undertook a more comprehensive review of research conducted from 2000 to 2013 that included unpublished reports as well as conference proceedings, and once again found clear and decisive evidence in support of precognitive effects. Nevertheless, some still argue that the findings overstate the evidence and that when such data are re-analysed the evidence is 'not sufficient to sway an appropriately skeptical reader' (Rouder & Morey, 2011, p. 688). Unsurprisingly, this has resulted in some suggesting that the use of meta-analysis rarely succeeds in resolving ideological debates (Ferguson, 2014). This in turn has meant that such contentious and often ambivalent findings have led to calls for more research to be conducted in an effort to bring both sides of the argument together (Franklin, Baumgart & Schooler, 2014).

With this in mind, an initial study was conducted utilising a modified repetition priming paradigm to examine possible precognitive priming effects (Vernon, 2015). This modified priming task showed no evidence of precognition when looking at the response times, but did find that participants were more accurate to respond to words they would see again in the future (Vernon, 2015). It was noted that such an anomalous finding could represent either a Type I error, or suggest that the component of memory that relies on accuracy — as opposed to speed — may be more susceptible to precognitive influences.

While such a possibility is speculative, given the unknown nature of precognition, there is some support that tasks relying only on speed of response show less evidence of precognitive effects (e.g., Traxler, Foss, Podali & Zirnstein, 2012), compared with those relying on accuracy of performance (see Bem, 2011; Subbotsky, 2013). This would suggest that a memory task relying primarily on accuracy of performance, such as a memory recall task, could be a more sensitive measure of precognition. Hence, one aim of the current study was to utilise a memory recall task that was based on the accuracy of recall as opposed to the speed.

Alongside changes to the nature of the task from one that focuses on speed to one that relies more on accuracy it may also be that changing the stimuli and the nature of the post recall practice sessions could also influence potential precall effects. For instance, the repetition priming paradigm reported by Vernon (2015) used everyday words relating to living and non-living objects. However, others have suggested that possible precognitive effects may be proportional to the level of physiological arousal elicited by the stimuli used (see Lobach, 2009; Maier et al., 2014). The point here is that stimuli that elicit stronger feelings of pleasure and/or discomfort may be better suited to producing precognitive effects. Hence, using positive and negative arousing images rather than everyday words may lead to a more robust precall effect.

In addition, previous research has utilised a variety of post-recall practice tasks that include categorisation of the presented words and visualisation of the related image (Bem, 2011; Galak et al., 2012; Ritchie et al., 2012). The rationale for selecting these specific practice tasks is unclear and may simply relate to the fact that they were used in the initial studies by Bem (2011). However, there are two points that need to be considered here. First, researchers fail to report performance on these post-test practice tasks. That is, how well participants were able to classify and/or image the referents. Second, that such practice tasks can be completed without the need to 'remember' the stimuli. That is, neither the categorisation nor visualisation tasks require the participant to 'recall' the words/images and as such there is no requirement to strengthen the memory representation for those items. It has long been known that rehearsal and practice enhances standard memory recall of items (see, e.g., Bjork, 1988). However, it is not known yet if the retroactive memory effects evidenced by precall are subject to the same influences and potential limitations as standard memory effects; if so it would make sense to ensure that the memory representation for the practiced items was strengthened as much as possible. Hence, having a practice task that requires participants to recall the items multiple times would lead to a stronger memory representation, which in turn could lead to more robust precall effects. Furthermore, consistent with the transfer appropriate processing view of memory, and Taylor's (2014) block view of the universe, it may be that similarity in processing between the recall test and the post recall practice task facilitates precognitive performance (see Brown & Craik, 2000).

Thus, this study tested potential precall effects by using both positive and negative arousing images whilst incorporating a post-test recall practice

task. In addition, the task was run on-line to reduce any possible bias and/or experimenter effects (see, Schlitz, Wiseman, Watt, & Radin, 2006). The prediction was that in the test phase participants will recall more of the items that appear in the later post-test practice phase compared with those that do not.

METHOD

Pre-Registration with KPU

This study was pre-registered with the Koestler Parapsychology Unit and a copy of the raw data uploaded to the site (ref#1019, <https://koestlerunit.wordpress.com/study-registry/registered-studies/>).

Participants

The study was halted once 121 participants had begun the study with 94 completing it (completion rate of 77.68%). This was based on a-priori power analysis showing that an N of 90 would be required (see KPU ref#1019). These 94 participants were opportunity sampled via an advertised web-link to the on-line study and completed the study voluntarily. Of the 94 participants, 26 were male and 68 female, with ages ranging from 22 to 62 (mean 42.9, SD 11.61).

Materials

The experiment utilised Qualtrics software (www.qualtrics.com) to build and present the study on line. This incorporated a Revised Paranormal Belief Scale (RPBS: Tobacyk, 2004) to assess participants' belief in anomalous events and a selection of 28 images from the International Affective Picture System's (IAPS) database (Lang, Bradley, & Cuthbert, 1997). These 28 images consisted of 14 positively- and 14 negatively-valenced items (see Appendix A). The images differed significantly in terms of valence (Positive = 7.19, Negative = 3.52; $t(26) = 14.47$, $p = 0.001$, 95% CI (3.14, 4.18), $d = 5.5$) but were matched for mean arousal level (Positive = 5.86, Negative = 5.78; $t(26) = 0.343$, $p = 0.734$, 95% CI (-0.395, 0.553), $d = 0.13$). The 28 images were then used to construct 8 sub-lists each containing 14 images (7 positive and 7 negative) with each sub-list matched for mean valence and arousal levels (see Appendix B).

Design

The six stages of the experiment are illustrated in Figure 1. The first stage provided a brief written outline of the study, stating that it aimed to examine extra sensory perception (ESP) and that this would involve viewing images that contained both positively and negatively valenced items which may elicit an emotional response and that if exposure to such images would be thought to have any negative impact then there was a recommendation not to continue. This was followed by a multiple choice question obtaining informed consent, a captcha question to exclude possible bots participating and then a question regarding demographics and finally the revised paranormal belief scale. The second stage followed on immediately and presented the 28 images, cropped to a width of 700px and a height of 525px,

with each image appearing in the centre of the screen along with a written identifier (e.g., name) written in Arial font size 36pt. Each image, along with its name, was presented on screen for 3 seconds. The third stage presented an open text box along with instructions to recall and enter in any order the names of as many of the previously seen 28 images as possible within 3 minutes. A clock also appeared in the top left of the screen counting down the time. Stage four presented participants with a sub-set of 14 images in the same format and for the same time duration as the initial images. These sub-sets of images were completely counterbalanced to ensure that each image occurred equally often in the repeated and non-repeated conditions. Stage five presented an open text box along with instructions to type in the name of the images just seen in any order within 2 minutes. Again, a clock appeared in the top left of the screen counting down the time. Stages 4 and 5 were then repeated three more times. This meant that across the main recall and post-recall practice stages each participant saw 84 images. Stage 6 required participants to answer whether they had left their PC or switched to another window/application during the study and then provided debrief information along with the experimenters contact details.

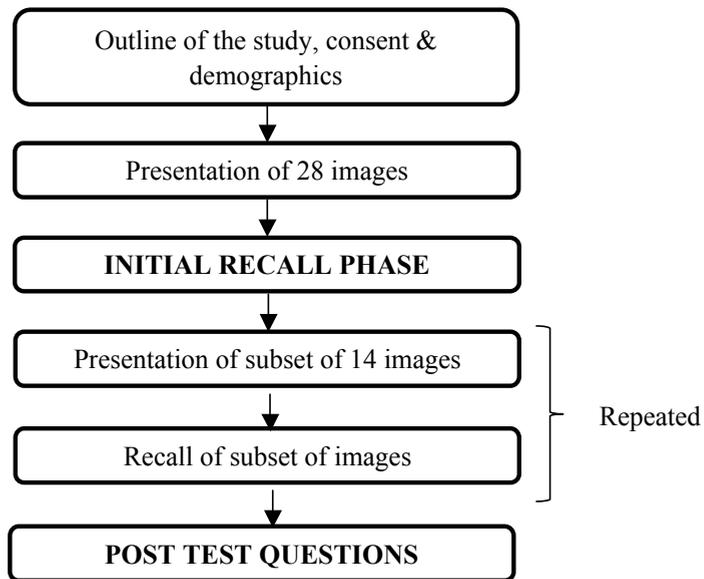


Figure 1. The six stages of the experiment

Procedure

The study was conducted on-line using Qualtrics software to deliver all information and stimuli and record all participant input via their keyboard. The study began by presenting an information screen informing participants

they are about to participate in a study exploring ESP, although the precise nature of this was not made clear until the final debrief at the end of the study. Once informed consent was obtained participants progressed to an information capture screen and entered demographic details and completed the revised paranormal belief scale (Tobacyk, 2004). The precise instructions given to participants then stated they would “be presented with a selection of both positive and negative images. Each image will remain on screen for 3 seconds. Please attend to the images and do not write anything down”. Following this the software presented all 28 images, with the appropriate image name appearing above the image, in a random sequence with each image/name appearing on screen for 3 seconds. The software used an inbuilt Mersenne Twister pseudorandom number generator (PRNG) to randomly select the order of stimuli presentation. Once all images had been shown participants then completed a surprise recall test. Precise instructions were “your task is now to recall as many of the images you have just seen and write their names in the box below. You have 3 minutes to do this. You can write them in any order and spelling isn’t important”. A timer on screen counted down from 3 minutes to provide an indication of how much time remained. Following this the software pseudo-randomly presented participants with one of four sub-lists of 14 images (see Appendix B), with the proviso that the PRNG evenly select the four sub-lists. This sub-list of 14 images (7 positive and 7 negative) was then presented in exactly the same way as the original list of 28 images. Once this had been completed another recall screen appeared with an open text box and a 2-minute timer. Participants then had 2 minutes to recall, in any order, as many of the just seen 14 images as they could in the allotted time. This presentation of the same sub-list of 14 images followed by a recall stage was repeated a further three times. Following this, a check screen asked participants if at any time during the study they shifted screens to check emails, looked away from their PC, or wrote down the words to help their recall. Finally, participants were provided with an information/debrief screen containing contact details of the Principal Investigator should they wish to obtain more information.

Ethics

Full University Faculty ethics approval was obtained for this study (Ref: 15/SAS/213C).

RESULTS

The RPBS was coded according to Tobacyk (2004) to create the 7 sub-scales of: traditional religious belief; Psi; Witchcraft; Superstition; Spiritualism; Extraordinary Life Forms, and Precognition. Precall was measured using level of recall accuracy for images recalled in stage 3 that were later repeated in stage 4 compared with those that were not-repeated. Given the requirement for participants to type in the name of the image, it is possible that a name could be misspelt or that a name may only be partially typed due to the time restriction. To deal with this all incorrectly spelled items were viewed by two external judges, blind to the aims of the study, to

ascertain whether they sufficiently identified the appropriate image. For partially typed responses a key criterion used was the requirement that there be a greater than 50% level of mapping between the letters and placements of the partially typed input and the name of the image.

RPBS Data

Data on the seven sub-scales of the RPBS are presented in Table 1. The means of the sub-scales were generally lower than those reported by Tobacyk (2004), in particular the subscale of ‘Traditional Religious Belief’ showed a much lower mean than the 6.3 reported by Tobacyk (2004).

Table 1

Showing mean belief levels for each of the seven sub-components of the RPBS

	Mean	SD
Traditional Religious Belief	3.56	2.00
Psi	3.11	1.40
Witchcraft	3.18	1.88
Superstition	1.61	1.03
Spiritualism	3.09	1.82
Extraordinary Life Form	3.20	1.14
Precognition	2.83	1.56

Precall Data

Ninety-four participants were each exposed to 28 images, creating a total of 2,632 trials. Of these, there were 51 (1.9%) that required additional consideration by two coders blind to the aims of the study due to spelling and/or grammar issues. The two coders who examined these items agreed 100% on the outcome. This included 7 instances of accepting ‘motorbike’ for ‘motorcycle’, 8 instances of accepting ‘cockroaches’ for ‘cockroach’, 18 instances of accepting ‘lightening’ for ‘lightning’, 1 instance of accepting ‘lighting’ for ‘lightning’, 1 instance of accepting ‘jaguer’ for ‘jaguar’, 10 instances of accepting ‘windsurfer’ for ‘windsurfers’ and 6 instances of accepting ‘skydiver’ for ‘skydivers’. There were also 14 (0.5%) intrusions which did not refer to any of the images seen but were invariably semantically related (e.g., ‘cheetah’ and ‘leopard’ in place of jaguar) and were excluded from the analysis.

A repeated measures t-test was conducted on recall scores comparing level of recall of images that were repeated with those that were non-repeated. A 2-tailed test was used to allow for the possibility that post-recall repetition of the images could impair precall performance (see, Ritchie et al., 2012). This showed that the level of mean recall for repeated images did not differ from images not-repeated (respective means: 7.28 vs. 7.38), $t(93) = 0.374$, $p = 0.710$, 95% CI (-0.604, 0.413), $d = 0.05$. The effect sizes of the precall scores for positive images was $d = -0.11$, for negative images was $d = 0.05$.

The precall score for positive and negatively valenced images along with their respective baseline can be seen in Table 2.

Of the 94 participants that took part, 35 (37.2%) reported in the post-study questionnaire that they were either distracted or switched to another application (e.g., to check emails, Facebook) during the study. When the main analysis was re-run, restricting the sample to those that did not report any such distractions or switching there was still no difference in the level of mean recall for repeated images compared to images not-repeated (respective means: 7.0 vs. 7.18), $t(58) = 0.574$, $p = 0.568$, 95% CI (-0.836, 0.463), $d = 0.08$.

To examine possible links between participant belief in paranormal events correlations were conducted between participants' precall scores and their scores on the RPBS (see Table 3). None of these correlations was significant (all $ps > 0.3$).

Table 2

Showing mean number of images (with SD) recalled (out of a total of 28) in the repeated and the non-repeated conditions for positively and negatively valenced images.

	Positive		Negative	
	Repeated	Non repeated	Repeated	Non repeated
Mean	3.27	3.43	4.01	3.94
Standard deviation	1.56	1.34	1.37	1.37

Table 3

Showing correlation coefficients (with significance values) between precall score and the seven sub-scales of the RPBS

	correlation	significance
Traditional Religious Belief	-.009	.93
Psi	.032	.75
Witchcraft	.106	.30
Superstition	-.053	.61
Spiritualism	-.068	.51
Extraordinary Life Form	-.016	.88
Precognition	-.016	.88

Post Recall Data

The pattern of post recall performance was examined using a repeated measures analysis of variance (ANOVA) with single factor of Time containing 4 levels (time1, time2, time3 and time4). The assumption of sphericity was not met, Mauchly's $W(5) = 0.816$, $p < 0.01$, hence the Greenhouse-Geiser correction was used when interpreting the ANOVA. This showed a main

Table 4

Showing mean number of images (with SD) recalled (out of a total of 28) in the repeated and the non-repeated conditions for positively and negatively valenced images.

	Post recall			
	Time 1	Time 2	Time 3	Time 4
Mean	10.56	11.54	11.80	11.81
Standard deviation	2.02	1.80	1.99	1.95

effect of Time $F(2.67, 248.9) = 16.201, p < 0.001, Mse = 2.36, \eta^2 = 0.148$. Pairwise comparisons using a Bonferroni correction to control for inflated Type I errors showed a significant increase in mean recall from time1 to time2 (respective means: 10.56 and 11.54) $p < 0.001, 95\% CI(-1.473, -0.484), d = 0.5$. There was no further change in mean recall from time 2 to time 4 (all $ps > 0.5$), see Table 4.

DISCUSSION

The results show no evidence for any precall effect when using both positive and negative arousing images. The post recall practice phases do show an initial increase in recall accuracy, but this plateaus after the second practice session, suggesting that two post-recall practice phases may be sufficient to establish a robust memory with continued practice having little or no effect.

There are in essence two alternative views that can be taken when attempting to interpret the pattern of data here. On the one hand, the data may show no precall effects simply because there were none to show, since such effects represent a scientific impossibility. This view would be consistent with others who report null effects when examining the possibility of precognition (e.g., Galak et al., 2012; Ritchie et al., 2012), and might bolster the argument that any such reported effects could be accounted for more simply in terms of improper statistical analysis (e.g., Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011), or even in terms of potential fraud (see e.g., Stokes, 2015). Of course, the argument that precognition is impossible, or any cognitive process for that matter, may say more about our current understanding of science, or the lack of it, than it does about the nature of such unusual phenomena. On the other hand, it may be that precognition represents a real effect and this study simply failed to find it. This approach would be more consistent with those reporting significant precognitive effects (Bem, 2011; Maier et al., 2014; Subbotsky, 2013) as well as the positive overviews obtained from the various meta-analyses (Bem et al., 2015; Honorton & Ferrari, 1989; Steinkamp, Milton, & Morris, 1998). At this moment in time, it is not possible to clearly identify which of these interpretive possibilities is correct. Hence, in an attempt to remain critical, yet open minded, which should be the hallmark of any scientist, the author offers the following discussion points as possible reasons why no potential precall effect

was found in the current study. This includes the level of participants' belief in such phenomena, whether they are relaxed and/or distracted during the task, the potential arousing nature of the stimuli themselves, and the slow/explicit nature of the task.

It is possible that individual belief may have influenced the behaviour of participants in this study. For example, research has shown that belief can be an important predictor of success in psi experiments (see e.g., Palmer, 1971). Where individuals can be classified into the unflattering categories of either sheep (i.e., believers) or goats (i.e., non-believers) research shows that those with a higher belief in the paranormal tend to score above chance, whereas those with a lower level of belief score either at chance or below it (see e.g., Parker, 2000). Given that the initial information page explicitly informed participants they would be taking part in a study looking at possible ESP effects, it is likely that their belief could have influenced the outcome. Furthermore, the mean overall score on the RPBS for the current group was 77.6, which is slightly lower than the reported norm of 89.1 by Tobacyk (2004). Hence, it is possible that the opportunity sample utilised in this study may have contained individuals with a more sceptical frame of mind, which in turn may have reduced any possible precall effects. Furthermore, recruiting potentially sceptical participants may have led to a reduction in variability, which could account for the lack of any correlation between scores on the belief scale and precall scores. Future research could address this by either attempting to recruit participants from specific target groups thought to have higher levels of belief in psi phenomena, and/or pre-screen all participants to ensure a target level of belief is expressed prior to having them take part in the experiment.

Alongside belief in psi phenomena it is possible that participants in the current study were not sufficiently relaxed during the memory task for any possible precall effects to emerge. For instance, Bem (2011) in his original suite of experiments began by presenting participants with an image whilst simultaneously playing new-age type music in an effort to help them relax. Such a view would be consistent with the early findings of Braud (1974) suggesting that relaxed participants perform better in psi related tasks (see also, Honorton, 1977). However, no images/music or relaxation induction stage were included in the current study. As such, future research could address this by including a specific 'relaxation induction' phase in an effort to help the participants relax and potentially facilitate any possible precall effects.

Given that the current study was also run on-line it is possible that participants may have experienced other 'distractions' during the task that took them away from the focus of the study. For instance, of the 94 participants who took part, 35 (37.2%) reported in the post-recall question phase that they had either switched applications and/or been distracted by something in their environment during the completion of the task. Such a possibility would be consistent with the suggestions of Braud (2002) who suggested that psi 'type' effects are invariably weak signals that can be overwhelmed and/or obscured by noise or distractions. Of course, it is not always possible to control all aspects of the environment, particularly when using an on-line

delivery method as in the current study. Nevertheless, greater emphasis could be made at the beginning of such a study of the need to complete the task in a quiet place away from any external distractions.

The current study used images from the IAPS database that were classified as either positive or negatively arousing; however, it is possible that the level of valence and/or arousal of these images may not have been sufficiently emotive. For instance, Maier et al. (2014) in their study on precognition used images from the IAPS with a positive valence of 7.57 and negative valence of 1.73, whereas the images used in the current study which had a positive valence of 7.19 and a negative valence of 3.52. Interestingly, Maier et al (2014) found a precognitive effect for the negative images but not for the positive images. Given this, it may be that using more emotive images would elicit a precall effect. Such a possibility would be consistent with suggestions that the more emotive the stimulus the more likely a psi effect will emerge (see e.g., Lobach, 2009; Radin, 2004).

Finally, the precall task used here involved presenting stimuli over time and allowing up to 3 minutes for participants to recall them. Such time would allow participants to adopt a more deliberate, conscious and attentive strategy for responding, all of which would naturally reflect explicit 'slow' processes. It may be that a task that is more reliant on slow explicit processes overshadows and/or reduces the possible influence of any psi-based behaviour, whereas a task that is more reliant on faster implicit processes may be more amenable to eliciting such anomalous effects. This idea is consistent with the view put forward by Bargh and Ferguson (2000) that psi behaviour may be better understood and explored using more indirect and/or implicit measures. To some extent it is also supported by the recent meta-analysis from Bem et al. (2015) that showed more robust precognitive effects for the fast-thinking type tasks, with no clear evidence of precognition when the slower thinking protocols were used. However, this picture is neither consistent nor clear. For instance, in the studies reported by Bem (2011) it is the slower explicit recall task that shows the greatest precognitive effect. In contrast, a repetition priming task requiring fast responses reported by Vernon (2015) failed to show any evidence of precognition. Furthermore, it may also be worth noting that researchers could be confounding the distinction between explicit/implicit processes with slow/fast thinking tasks. Hence, this distinction between explicit/slow and implicit/fast may be useful but needs to be more clearly examined to allow the contribution of each aspect to be teased apart. As such, it may represent a potentially fruitful avenue for future researchers.

In conclusion, no evidence of precall was found for an on-line task using emotive images. However, a number of methodological points are offered for consideration, including the level of belief in psi phenomena of the participants, the use of a relaxation induction, possible distractions and the emotive nature of the stimuli, which may have contributed to this null result.

ACKNOWLEDGEMENTS

This research was supported in part by a small grant awarded to the author by The Society for Psychical Research. Thanks also go to Dr Tammy Dempster and Dr Lynn Nichols for help with data processing and coding.

School of Psychology, Politics & Sociology
Canterbury Christ Church University
Canterbury, Kent.
CT1 1QU
 david.vernon@canterbury.ac.uk

REFERENCES

- Bahrick, H.P. (2000). Long-term maintenance of knowledge. In E. Tulving, & F.I.M. Craik (Eds.), *The Oxford handbook of memory* (pp. 347–362.). Oxford: University Press.
- Bargh, J.A., & Ferguson, M.J. (2000). Beyond behaviourism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*, 925–945.
- Bem, D.J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*, 407–425.
- Bem, D.J., Tressoldi, P., Rabeyron, T., & Duggan, M. (2015). Feeling the future: A meta-analysis of 90 experiments on the anomalous anticipation of random future events. *F1000 Research*, *4*, 1–33. doi: 10.12688/f1000research.7177.2
- Bjork, R.A. (1988). Retrieval practice and the maintenance of knowledge. In M. Gruneberg, P. Morris, & R. Sykes (Eds.), *Practical aspects of memory: Current research and issues*. (Vol. 2, pp. 396–401.). Chichester: John Wiley.
- Braud, W.G. (1974). Relaxation as a psi-conducive state. *Bulletin of the Psychonomic Society*, *3*(2), 115–118.
- Braud, W.G. (2002). Psi-favourable conditions. In V.G. Rammohan (Ed.), *New frontiers of human science: A festschrift for K. Ramakrishna Rao* (pp. 95–118). Jefferson, NC: McFarland.
- Brown, S.C., & Craik, F.I.M. (2000). Encoding and retrieval of information. In E. Tulving, & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 93–107.). Oxford: University Press.
- Cardeña, E., Marcusson-Clavertz, D., & Palmer, J. (2015). Reintroducing parapsychology. In E. Cardeña, J. Palmer & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 1–11). Jefferson, NC: McFarland & Company, Inc.
- Ferguson, C.J. (2014). Comment: Why meta-analyses rarely resolve ideological debates. *Emotion Review*, *6*, 1–2. doi: 10.1177/1754073914523046
- Franklin, M.S., Baumgart, S.L., & Schooler, J.W. (2014). Future directions in precognition research: More research can bridge the gap between skeptics and proponents. *Frontiers in Psychology*, *5*(907), 1–4. doi: 10.3389/fpsyg.2014.00907
- Galak, J., LeBouf, R.A., Nelson, L.D., & Simmons, J.P. (2012). Correcting the past: Failures to replicate psi. *Journal of Personality and Social Psychology*, *103*(6), 933–948. doi: 10.1037/a0029709
- Honorton, C. (1977). Psi and internal attention states. In B.B. Wolman (Ed.), *Handbook of parapsychology* (pp. 435–472.). New York: Van Nostrand Reinhold.
- Honorton, C., & Ferrari, D.C. (1989). Future telling: A meta-analysis of forced-choice precognition experiments, 1935–1987. *Journal of Parapsychology*, *53*, 281–308.
- Irwin, H.J., & Watt, C. (2007). *An introduction to parapsychology (5th ed.)*. Jefferson, NC: McFarland & Co.

- Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (1997). *International Affective Picture System (IAPS): Technical manual and affective ratings*. Gainesville, FL. The Center for Research in Psychophysiology, University of Florida.
- Lobach, E. (2009). Presentiment research: Past, present and future. In C.A. Roe, L. Coly & W. Kramer (Eds.), *Utrecht II: Charting the future of parapsychology* (pp. 22–45). New York: Parapsychology Foundation.
- Maier, M.A., Buchner, V.L., Kuhbandner, C., Pflitsch, M., Fernandez-Capo, M., & Gamiz-Sanfeliu, M. (2014). Feeling the future again: Retroactive avoidance of negative stimuli. *Journal of Consciousness Studies*, 21(9–10), 121–152.
- Mossbridge, J.A., Tressoldi, P., Utts, J., Ives, J.A., & Radin, D. (2014). Predicting the unpredictable: Critical analysis and practical implications of predictive anticipatory activity. *Frontiers in Human Neuroscience*, 8(146), 1–10. doi: 10.3389/fnhum.2014.00146
- Palmer, J. (1971). Scoring in ESP tests as a function of belief in ESP. *Journal of the American Society for Psychical Research*, 66, 1–26.
- Parker, A. (2000). A review of the Ganzfeld work at Gothenburg University. *Journal of the Society for Psychical Research*, 64, 1–15.
- Radin, D. (2004). Electrodermal presentiments of future emotion. *Journal of Scientific Exploration*, 18(2), 253–273.
- Ritchie, S.J., Wiseman, R., & French, C.C. (2012). Failing the future: Three unsuccessful attempts to replicate Bem's retroactive facilitation of recall effect. *PLOS One*, 7(3), e33423.
- Rouder, J.N., & Morey, R.D. (2011). A Bayes factor meta-analysis of Bem's ESP claim. *Psychonomic Bulletin & Review*, 18(4), 682–689. doi: 10.3758/s13423-011-0088-7
- Schlitz, M., Wiseman, R., Watt, C., & Radin, D. (2006). Of two minds: Sceptic-proponent collaboration within parapsychology. *British Journal of Psychology*, 161, 1–11. doi: 10.1348/000712605X80704
- Steinkamp, F., Milton, J., & Morris, R.L. (1998). A meta-analysis of forced-choice experiments comparing clairvoyance and precognition. *Journal of Parapsychology*, 62, 193–218.
- Stokes, D.M. (2015). The case against psi. In E. Cardena, J. Palmer & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 42–48.). Jefferson, NC: McFarland & Company.
- Storm, L., Tressoldi, P., & Di Risio, L. (2012). Meta-analysis of ESP studies, 1987–2010: Assessing the success of the forced-choice design in parapsychology. *Journal of Parapsychology*, 76(2), 243–273.
- Subbotsky, E. (2013). Sensing the future: Reversed causality or a non-standard observer effect? *The Open Psychology Journal*, 6, 81–93.
- Taylor, J. (2014). The nature of precognition. *Journal of Parapsychology*, 78(1), 19–38.
- Tobacyk, J.J. (2004). A revised paranormal belief scale. *The International Journal of Transpersonal Studies*, 23, 94–98.
- Traxler, M.J., Foss, D.J., Podali, R., & Zirnstein, M. (2012). Feeling the past: The absence of experimental evidence for anomalous retroactive influences on text processing. *Memory and Cognition*, 40, 1366–1372. doi: 10.3758/s13421-012-0232-2
- Vernon, D. (2015). Exploring precognition using a repetition priming paradigm. *Journal of the Society for Psychical Research*, 79, 65–79.
- Wagenmakers, E., Wetzels, R., Borsboom, D., & van der Maas, H.L. (2011). Why psychologists must change the way they analyze their data: the case of psi. Comment on Bem (2011). *Journal of Personality and Social Psychology*, 100(3), 426–432.

APPENDIX A: Listing the 28 images (14 positive and 14 negative) used in the study, with valence and arousal ratings.

	Positive Image			Negative Image			
	IAP#	Valence	Arousal	IAP#	Valence	Arousal	
Jaguar	1650	6.65	6.23	Snake	1110	3.84	5.96
Waterfall	5260	7.34	5.71	Spider	1201	3.55	6.36
Skydivers	5621	7.57	6.99	Dog	1302	4.21	6.00
Mountains	5700	7.61	5.68	Shark	1930	3.79	6.42
Windsurfers	5623	7.19	5.67	Bomb	2692	3.36	5.35
Baby	2660	7.75	4.44	Cockroach	1274	3.17	5.39
Fireworks	5910	7.8	5.59	Gun	6610	3.6	5.06
Lightning	5950	5.99	6.79	Tornado	5971	3.49	6.65
Cakes	7220	6.91	5.3	Tank	6940	3.53	5.35
Pizza	7350	7.1	4.97	Boxer	8060	5.36	5.31
Gymnast	8470	7.74	6.14	Toilet	9301	2.26	5.28
Motorcycle	8251	6.16	6.05	Solider	9160	3.23	5.87
Pilot	8300	7.02	6.14	Skull	9480	3.51	5.57
Money	8501	7.91	6.44	Ship	9600	2.48	6.46
Mean		5.31	5.88	Mean		5.41	5.77

APPENDIX B: The 8 sub-lists (consisting of 4 practice lists and 4 non-practice baseline lists) created from the original list of 28 images with valence and arousal ratings.

	Items	Mean Valence	Mean Arousal
Practice 1	Snake, Spider, Dog, Shark, Bomb, Cockroach, Gun, Lightning, Cakes, Pizza, Gymnast, Motorcycle, Pilot, Money	5.31	5.88
No practice baseline 1	Jaguar, Waterfall, Skydivers, Mountains, Windsurfers, Baby, Fireworks, Tornado, Tank, Boxer, Toilet, Soldier, Skull, Ship	5.41	5.77
Practice 2	Jaguar, Waterfall, Skydivers, Mountains, Windsurfers, Baby, Fireworks, Tornado, Tank, Boxer, Toilet, Soldier, Skull, Ship	5.41	5.77
No practice baseline 2	Snake, Spider, Dog, Shark, Bomb, Cockroach, Gun, Lightning, Cakes, Pizza, Gymnast, Motorcycle, Pilot, Money	5.31	5.88
Practice 3	Jaguar, Waterfall, Skydivers, Shark, Bomb, Cockroach, Gun, Snake, Tank, Toilet, Gymnast, Motorcycle, Pilot, Money	5.28	5.89
No practice baseline 3	Tornado, Spider, Dog, Mountains, Windsurfers, Baby, Fireworks, Lightning, Cakes, Pizza, Boxer, Soldier, Skull, Ship	5.44	5.76
Practice 4	Snake, Spider, Dog, Shark, Windsurfers, Baby, Fireworks, Lightning, Cakes, Pizza, Gymnast, Soldier, Skull, Ship	5.36	5.82
No practice baseline 4	Jaguar, Waterfall, Skydivers, Mountains, Bomb, Cockroach, Gun, Tornado, Tank, Boxer, Toilet, Motorcycle, Pilot, Money	5.36	5.83



Established 1882

GATHERING IN MORE SHEEP AND GOATS: A META-ANALYSIS OF FORCED-CHOICE SHEEP-GOAT ESP STUDIES, 1994–2015

BY LANCE STORM & PATRIZIO E. TRESSOLDI

ABSTRACT

The terms ‘sheep’ and ‘goat’ were introduced by Gertrude Schmeidler — sheep are those who accept the possibility of ESP occurring under given experimental conditions, while goats are those participants who reject the possibility. In statistical tests of psychic ability, Schmeidler found that sheep tended to score above chance, while goats (rather than scoring at chance) tended to score below chance. This scoring differential is known as the sheep-goat effect (SGE). This study is a meta-analysis of the SGE in forced-choice literature, being a continuation from where Lawrence (1993) left off. The period of analysis was 1994 to 2015. The authors retrieved 49 studies reported by 43 investigators. The mean *ES* for ESP = .045, mean $z = 0.75$, Stouffer $Z = 5.23$ ($p = 8.47 \times 10^{-6}$), and the mean trial-based SGE = 0.034, mean $z = 0.24$, Stouffer $Z = 1.67$ ($p = .047$). Thus, our SGE is on par with Lawrence’s reported “ $r = 0.029$ ”. There was no relationship between study quality and ESP effect or SGE, but there was a significant incline in the SGE over a period of 22 years. The SGE did not vary significantly with belief measure used. Bayesian analysis of the same dataset yielded results supporting the ‘frequentist’ finding that the null hypothesis should be rejected. These and other findings are generally comparable to Lawrence’s, altogether indicating a “belief-moderated communications anomaly” in the forced-choice ESP domain that has been effectively uninterrupted and consistent for almost 70 years.

INTRODUCTION

Schmeidler (1943, 1945) categorised participants in paranormal experiments as either those who think that ESP is possible under a given experimental condition (‘sheep’), or those who rejected this possibility (‘goats’). The terms ‘sheep’ and ‘goat’ were adopted by Schmeidler from a New Testament simile that describes how a shepherd “separates the sheep from the goats” (Matthew, 25: 31–33). It is claimed that sheep tend to score significantly above mean chance expectation (MCE) on psi tasks, whereas goats tend to score at or below MCE — this effect was termed the sheep-goat effect (SGE) and it was usually measured on a sheep-goat variable (SGV). Literature reviews looking at the SGE found some support for a belief-based psi-scoring differential between sheep and goats (see Palmer 1971, 1972, 1977; Schmeidler & McConnell, 1973), and Lawrence’s (1993) meta-analysis of forced-choice studies dating back to 1947 demonstrated an accumulative SGE that was small in size but highly significant — sheep consistently scored

better than goats. Lawrence described the SGE as being “robust over a wide range of potential moderating variables and experimental manipulations” (p. 75). These variables included experimental setting, sample size, knowledge of results, and study quality.

This study is a meta-analytic review of the sheep-goat forced-choice ESP literature from 1994 to 2015 aimed at determining whether a comprehensive up-to-date database shows a significant differential effect, and thus confirms Lawrence’s (1993) original meta-analysis. Before we describe the present study, which is a follow-up of Lawrence’s meta-analysis, a brief review of the sheep-goat literature is now presented.

The Sheep-Goat Effect

Probably the best-known process-oriented variable that has been examined in relation to psi performance is the sheep-goat variable (SGV) which is usually measured with paranormal belief scales or questions. Schmeidler (1943, 1945) started with simple questions (e.g., “Do you accept the possibility of ESP under the conditions of the experiment?”) that were meant to gauge how a participant thought they might perform in an ESP experiment. In other words, Schmeidler sought to measure belief in psi performance. For Schmeidler, participants’ response scores would be the means by which the experimenter could separate the sheep from the goats as defined above, but these questions did not indicate actual level of psychic ability, and Schmeidler’s early studies did not draw distinctions between the sheep-goat effect *per se* and extraneous factors such as relaxation and experimenter effects (see Schmeidler & McConnell, 1973).

Since Schmeidler’s time, the scope of the SGV has been considerably widened, to the point where it often simply refers to the effects on performance of paranormal belief or disbelief in the abstract sense. In fact four distinct meanings of the term sheep-goat variable (referred to as “criteria”) have been identified (Palmer, 1971, pp. 391–394): (i) the participant believes ESP is possible during an ESP experiment; (ii) the participant believes in ESP in the abstract or theoretical sense; (iii) the participant personally believes they have psychic ability; and (iv) the participant believes he/she can or has scored above-chance in an ESP test. Currently, the widespread use of scales such as the Australian Sheep-Goat Scale (Lange & Thalbourne, 2002; Thalbourne, 1995; Thalbourne & Delin, 1993) more than cover these criteria and items on alleged experience of psi are also included under the rubric of the SGV.

A sizeable number of tests of the sheep-goat effect have been conducted, and there are a number of previous literature reviews on the effect (Palmer, 1971, 1972, 1977; Irwin, 1993; Schmeidler & McConnell 1973; Thalbourne, 2010), including one major meta-analysis (Lawrence, 1993). In particular, Palmer (1971) presented an analysis of the sheep-goat effect in a number of studies published during the period 1947 to 1970, and found that 6 of 17 experiments (35%) produced “significant confirmations of the sheep-goat hypothesis” (p. 402). Of these 17, 13 (76%) were in the predicted direction. Second, Palmer (1977) reported 7 new experiments since his 1971 study, where 5 (71%) produced significant effects in the direction hypothesized, and

6 out of 7 (86%) were in the predicted direction though not necessarily significant.

Lawrence's (1993) above-mentioned review only looked at the forced-choice literature, but it is nevertheless informative and thorough. For our purposes, a typical forced-choice study requires the identification of concealed targets such as symbols, letters, shapes, numbers, and so on. That is, the target-guess is "one of a limited range of possibilities which are known to [the participant] in advance" (Thalbourne, 2003, p. 44). Forced-choice differs from free-response in that the latter "describes any test of ESP in which the range of possible targets is relatively unlimited and is unknown to the percipient" (Thalbourne, 2003, p. 44).

Lawrence (1993) accumulated 73 forced-choice studies (4,500 participants, 685,000 guesses), and he calculated an SGE of 0.029, with a highly significant Stouffer Z of 8.17, $p = 1.33 \times 10^{-16}$. Eighteen studies (24%) showed a significant SGE at $p = .05$. The mean z was 0.96, and mean SGE per investigator was 0.026. He also found that study quality and SGE had not changed in 46 years. The file-drawer estimate was 1,726 (23 unreported, nonsignificant studies for every one successful study). The so-called 'file-drawer problem' refers to the claim that there exist many unpublished studies with chance effects, and if these studies were taken into account, the significant effects found in meta-analyses would reduce to non-significance (Hyman, 1985; Rosenthal, 1984). Such a high file-drawer estimate suggests that too many unpublished studies are required for it to be a viable explanation of the observed effect.

In addition, Lawrence (1993) determined that of the 65 studies that reported the order of administration of the belief scale and ESP task there were 15 studies where the sheep-goat scale was given after the ESP task ($ES = 0.044$), and 50 studies where the scale was given before the ESP task ($ES = 0.028$). He thus found suggestive evidence of an artefact (an order effect in which knowledge of one's performance at the ESP task might affect one's level of declared belief in ESP), though the difference between effect sizes was not significant. Although the number of studies stating the kind of feedback (i.e., 'no feedback' and 'trial-by-trial feedback') given was "insufficient" (p. 81), Lawrence nevertheless advised that measures of belief be given before the ESP task. Lawrence concluded that there was a "belief-moderated communications anomaly" (p. 75) — i.e., a sheep-goat effect.

The Forced-Choice Domain

The present study is a meta-analysis of forced-choice studies that feature tests on the sheep-goat effect (SGE). However, we will also assess the overall ESP performance (herein termed ES) of the studies independent of SGEs (see H1 in the Results section). Thus, results of the analyses below are not to be confused or equated with those found in the meta-analysis by Storm, Tressoldi, and Di Risio (2012). First, the Storm et al. meta-analysis covered forced-choice ESP studies from a different period (i.e., dating from 1987 to 2010). Second, the Storm et al. meta-analysis included any forced-choice study whether or not sheep-goat effects were investigated. To make it clear, H1 is tested to confirm an ESP effect independent of the SGE.

It should be noted that Storm et al. (2012) formed a homogeneous dataset of 72 forced-choice studies that yielded a very weak but significant mean ES of 0.01 (Stouffer $Z = 4.86$, $p = 5.90 \times 10^{-7}$). There was no evidence that these results were due to low-quality design or selective reporting. The clairvoyance studies did not produce a significantly higher mean ES than the precognition studies, and target type did not make a difference to effect size.

Storm et al. (2012) noted that effects did not vary between investigators, but they did find suggestive evidence that the number of choices (k) per trial was related to the z score. They also found evidence of a linear incline in ES values indicating that effect sizes have increased over the period 1987 to 2010. Such inclines are not usual in parapsychology (Bierman, 2001; Hövelmann, 2015), but there is mounting evidence that, regardless of experimental domain, neither should declines in effects over the years be expected, or taken as a norm (for example, see Baptista, Derakhshani, & Tressoldi, 2015).

Study Design and Hypotheses

In this meta-analysis a sheep-goat effect and an ESP effect are sought in the database of forced-choice studies dating from 1994 to 2015. Chronologically, the meta-analysis continues from the end of Lawrence's (1993) review period.

The study will also investigate whether methodological factors account for the SGE — that is, whether the SGE can be explained normally as artefacts of study quality, a limited pool of researchers/laboratories, and/or publication bias. Also to be assessed are:

- year of study which is related to decline/incline effects, to see if the SGE has changed over the years (Baptista et al., 2015; Bierman, 2001; Hövelmann, 2015);
- the various sheep-goat (belief) measures used in the dataset, to discern whether the SGE depends on how one measures paranormal belief, which is “a potential source of confusion” (Lawrence, 1993, p. 80);
- the differences between telepathy, precognition, and clairvoyance: Storm et al. (2012) found a significant difference only between precognition and telepathy in their dataset of forced-choice studies;
- differences between target types, since some targets may be more appealing to participants than others, and this preference could influence psi performance and the SGE; however, Storm et al. (2012) found no significant difference between target types;
- the ‘ k -choices’ hypothesis, where it is claimed that the number of target choices, k , in a given trial is related to effect size (Timm, 2000). Timm argued that effect size measures have limited use if they do not adequately account for k , and Storm et al. (2012) showed that “ z scores tended to increase as k increased” (p. 256).

Finally, a Bayesian analysis will be conducted on our dataset. Bayesian analysis in parapsychology is a recent approach aimed at providing an alternative to classical ‘frequentist’ or null hypothesis significance testing (NHST), although NHST is the main thrust of this paper (see Bem, Utts, &

Johnson, 2011; Rouder & Morey, 2011; Tressoldi, 2011). The benefit of Bayesian meta-analysis lies in the fact that it provides a measure of the probability of a phenomenon (H1) and its non-existence (H0) called posterior odds, which are the product of summary statistics (so-called Bayes factors, or probability ratios) and prior odds. More specifically, we will perform a hierarchical random effect Bayesian Parameter Estimation, which emphasises what is referred to as the explicit posterior probability distribution of the parameter values (e.g., the mean, *SD*, *ES*); that is, the conditional probability that is assigned after the data obtained from experiments are taken into account and the prior probability is updated. In this Bayesian approach to assessing null values, we set up a range of parameter values, including the null value, and use Bayesian inference to compute the relative credibility of each of these values. These parameter values are referred to as the prior distribution, or ‘prior’, and must be specified in the definition of the model to be tested.

Thus, the following hypotheses are proposed.

- H1. Forced-choice sheep-goat studies produce statistical evidence of a communications anomaly known as ESP.
- H2. There is a sheep-goat scoring differential in scoring at ESP tasks (i.e., a sheep-goat effect; SGE).
- H3. *ES* and SGE vary in strength according to target types.
- H4. Number of choices (*k*) per trial is positively related to *z* score.
- H5. Bayesian analysis of the forced-choice sheep-goat studies yields statistical evidence of a sheep-goat effect (SGE).

Other factors relating to the SGE not expressly hypothesised will be tested post hoc.

METHOD

Study Retrieval

Suitable studies were found from a comprehensive search of the following peer-reviewed journals: *Australian Journal of Parapsychology*, *European Journal of Parapsychology*, *Explore: The Journal of Science and Healing*, *Journal of Parapsychology*, *Journal of Scientific Exploration*, *Journal of the American Society for Psychological Research*, *Journal of the Society for Psychological Research*, and *Proceedings of the Annual Convention of the Parapsychological Association*. The search period for the meta-analysis was 1994 to 2015, which continues from Lawrence (1993) — his period of assessment was 1947 to 1993. However, while the period of analysis for the present meta-analysis is 1994 to 2015, we make one exception with the necessary inclusion of a multiple-study article by Haraldsson (1993) which was not available for analysis by Lawrence at the time of his literature search (T. Lawrence, personal communication, August 24, 2016).

Insofar as there are many other non-specialist journals that feature parapsychological studies, exhaustive internet searches of databases were also conducted, including searches of PsycINFO, PsycARTICLES, Informat and Web of Science. The following keywords and subject headings were entered in the search: ‘extrasensory perception’, ‘ESP’, ‘forced-

choice', 'telepathy', 'clairvoyance', 'precognition', 'anomalous cognition', 'parapsychology', 'paranormal', 'paranormal belief', 'psi', 'psi belief', 'sheep-goat', and 'sheep-goat effect'.

Selection Criteria

The following seven selection criteria were adopted, and studies were excluded if any of these criteria were not met — studies had to:

- use a forced-choice design;
- include sheep-goat measures (see next list);
- test ESP (i.e., telepathy, clairvoyance, and precognition, therefore excluding studies that only tested psychokinesis);
- use human participants only (not animals);
- test more than two participants to avoid the inherent problems that are typical in case studies;
- use randomization procedures for selection of targets that could not be manipulated by the experimenter or participant;
- provide sufficient information (e.g., number of trials and hit counts) so that the hit rates, z scores, effect sizes (where $ES = z/\sqrt{n}$), etc. could be calculated.

The sheep-goat measures were:

- Anomalous Experiences Inventory (Gallagher, Kumar, & Pekala, 1994);
- Australian Sheep-Goat Scale (Thalbourne, 1995);
- Icelandic Sheep-Goat Scale (Haraldsson & Houtkooper, 1992);
- Paranormal Belief Scale (Tobacyk, 2004);
- Sheep-goat psi belief (Palmer, 1972);
- Single-item sheep-goat questions (SGQs), or two-, three-, or four-item SGQs (e.g., 'belief in telepathy?'; 'existence of ESP probable?'; 'psi-able?'; 'believe in psychic phenomena?'; etc.).

Procedure

For each study, the following factors were checked: (1) the criteria adopted for selecting participants; (2) number of participants; (3) number of trials; (4) type of ESP task (telepathy, clairvoyance, or precognition); (5) number of choices/alternatives (k) in the tasks; (6) total number of hits; and (7) correlations of psi outcomes with the sheep-goat measure. The direct hits measure was preferred as it provides a more "conservative" result (see Honorton, 1985, p. 54), and is easier to grasp intuitively.

Databases (Excel and SPSS) were prepared listing all the forced-choice studies that utilised sheep-goat measures in the period 1994 to 2015 (including Haraldsson, 1993). The databases included target types, k choices, correlation values, quality ratings, and sheep-goat measure. With these data, proportion-of-hits values could be compared with proportions expected by chance (i.e., mean chance expectation; MCE). The databases also included the details listed in Appendix A1, which lists author and publication date, trial counts, hit counts, z scores, ES values, and type of ESP task (telepathy, clairvoyance, or precognition). From the trial and hit counts, z scores (if not

already given in the study) were calculated for sheep and goats combined, and for sheep and goats separately, using an online binomial exact-probabilities calculator (<http://faculty.vassar.edu/lowry/binomialX.html>) — see Appendix B (Formulae 1 and 2).

Of primary importance was the comparison of sheep ESP performance with goats' ESP performance. To conduct this comparison, the z score of the difference in two proportions using Formula #10 from Edge, Morris, Palmer, and Rush (1986, p. 154) was used, since this formula was used by Lawrence (1993, p. 76). This formula gives a critical ratio (CR) as a z score (Appendix B Formula 3). For studies that did not give actual sheep and goat data, Kendall's tau, Spearman's rho, and Person's r values were converted to z scores using formulae given in Appendix B (Formulae 4 to 7).

Forced-choice studies were rated for quality based on criteria used by Honorton and Ferrari (1989), and Steinkamp, Milton, and Morris (1998):

- Number of trials pre-planned;
- Appropriate randomisation (using electronic apparatuses or random tables);
- Random target positioning;
- Automated scoring;
- Safeguard(s) against data manipulation by participants;
- Experimenter(s) blind to target identity.

RESULTS

Descriptive Statistics

Appendix A1 lists a total of 49 forced-choice experiments drawn from 33 articles by 43 investigators that satisfied the above criteria for inclusion in the meta-analysis (see References for articles by name — articles marked with an asterisk indicate papers included in the meta-analyses). There were a total of 113,431 trials and 32,864 hits. Samples ranged from 60 trials to 18,150 trials. The total number of participants was 4,164, ranging from as few as six, to as many as 600 in a single study. Of the 49 studies, 20 (41%) were precognition, 17 (35%) were clairvoyance, 3 (6%) were telepathy, and 9 (18%) could not be classified (but could have been either precognition or clairvoyance).

Four-choice and five-choice designs were the most frequent (39 studies, or 79%): 27 (55%) used a four-choice design, and 12 (25%) used a five-choice design, 5 (10%) used a 2-choice design, and 5 (10%) used a six-choice design.

In only 16 studies (33% of studies), were specific data for sheep and goats found (see Appendix A2), from which a CR value could be calculated. Consequently, specific hit rates for sheep and goats in 67% of studies were either not provided, or could not be calculated. These 67% gave alternative measures (i.e., Kendall's tau, Spearman's r_s values, or Pearson's r values) to indicate relationships between paranormal belief and the given psi outcome (i.e., hit score). Nevertheless, from these values it was possible to calculate z scores equivalent to CR(z) scores following Lawrence's (1993, p. 76) procedure, and corresponding ES values, where $ES = z/\sqrt{n}$ (see Appendix B, Formulae 4

to 7). Thus, for all 49 studies, it was possible to calculate a mean z and mean ES value. To avoid confusion, we refer to the ESP effect as ES, and the sheep-goat effect as SGE, which is CR(ES) derived from CR(z).

Quality Ratings

Two judges were used to rate the studies for their quality (judges were not aware of the sources of these papers so that results could not affect ratings). One judge holds a PhD and is a full-time researcher with 12 years of experimental experience in parapsychological investigations, and the other also holds a PhD and has seven years' experience in parapsychological investigations. They answered "Yes" or "No" to each of the criteria. The study score is the ratio of points awarded with respect to the items applicable (minimum score is $1/6 = 0.17$; maximum score is $6/6 = 1.00$). Thirty-five studies out of 49 (71%) received a perfect score from at least one judge. On average, a majority of criteria (i.e., 4 or more out of 6) were met in 35 of the 49 studies (i.e., 71%). These results would suggest that study quality is at an acceptable standard, albeit with some scope for improvement.

Cronbach's alpha for the two judges' ratings was 0.76 indicating satisfactory inter-rater reliability. The correlation between mean Quality Scores and ES values was weak and not significant, $rs(47) = -0.11$ ($p = .442$, two-tailed), suggesting that effect size was not an artefact of poor experimental design. The correlation between mean Quality Scores and sheep-goat CR(ES) values (i.e., SGE) was also weak and not significant, $rs(47) = -0.13$ ($p = .381$, two-tailed), suggesting once again that effect size was not an artefact of poor experimental design.

We note the relationship between mean Quality Score and year of publication is negative but not significant, $rs(47) = -0.19$ ($p = .193$, two-tailed). In other words, by our criteria, study quality has not worsened across the span of over 20 years.

Planned Analyses

The method of mean result (Rosenthal, 1984) was used to calculate single values that are representative of the sample data. That is, measures of outcome (z 's and ES's) are totalled and averaged to give mean values. The major problem with the database is that investigators used various instruments to measure the SGE, but that problem had no direct bearing on our capacity to calculate standardised outcomes, as can be seen from the last paragraph in Descriptive Statistics section above.

A cumulative forest plot was prepared to illustrate the sheep-goat effect (SGE) over time. This plot is a cumulative representation of studies illustrating possible time trends in effect sizes as new studies are added. We added studies successively by their publication year, and our Figure 1 shows a tendency for SGE values to increase, thus giving evidence that the accumulation over time is not attributable to older SGE values, nor study quality as just indicated above.

We also include a funnel plot to indicate alleged publication bias (see Figure 2). The recommended y-axis measure of 'precision' is the inverse of SE (i.e., $1/SE$) of the SGE values (i.e., CR[ES]), and the x-axis is SGE values. The

Begg rank correlation test was significant, $z = 2.84$, $p = .004$ (two-tailed), but this test is not recommended in meta-analysis (Jin, Zhou, & He, 2015). However, Egger's linear regression test, which is considered "high power" (Jin et al., 2015, p. 348) was not significant, $t(47) = 1.23$, $p = .225$ (two-tailed).

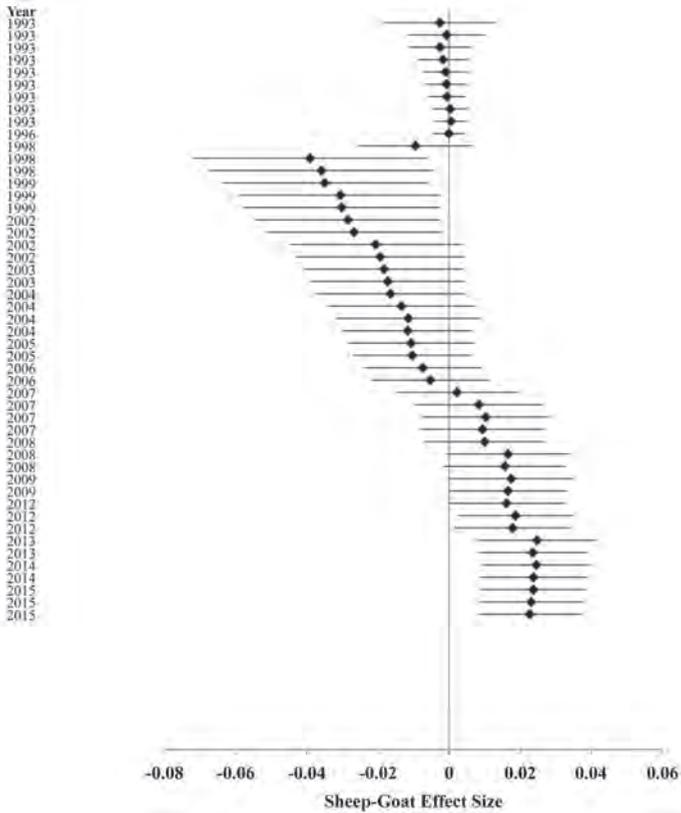


Figure 1. Cumulative forest plot showing the sheep-goat effect has increased on average over time 1993–2015 ($N = 49$)

Publication bias is not necessarily indicated just because a funnel plot shows outliers (Lau et al., 2006; Sterne & Harbord, 2004). A given funnel plot might merely indicate that the smaller studies had less methodological rigour than the larger studies, or the smaller studies deliberately targeted participants that were likely to respond better to a given treatment compared to larger samples. In the case of ESP studies, small- N studies tend to be comprised of participants 'selected' for their psi ability, or are over-represented by enthusiastic 'first responders' who are invariably sheep. Either of these two factors can inflate effect sizes, whereas it is very difficult to populate large samples in the same way(s) right up to the end of testing. While the information available from the studies was not sufficient to test for these

biases, we did find that SGE correlated significantly and negatively with trial counts, suggesting that the smaller studies had the stronger effects, $r_s(47) = -0.37, p = .005$ (two-tailed).

Following conventional meta-analytic procedure, step-by-step attempts were made to trim the database of outliers, but when the exclusion rate reached nearly 50% of studies (and counting), the procedure was abandoned. Such techniques can demonstrate the folly of attempting to construct homogeneous databases. Consequently, it was decided that all planned and post hoc tests would be conducted on the full heterogeneous dataset ($N = 49$), and nonparametric tests were conducted where appropriate.

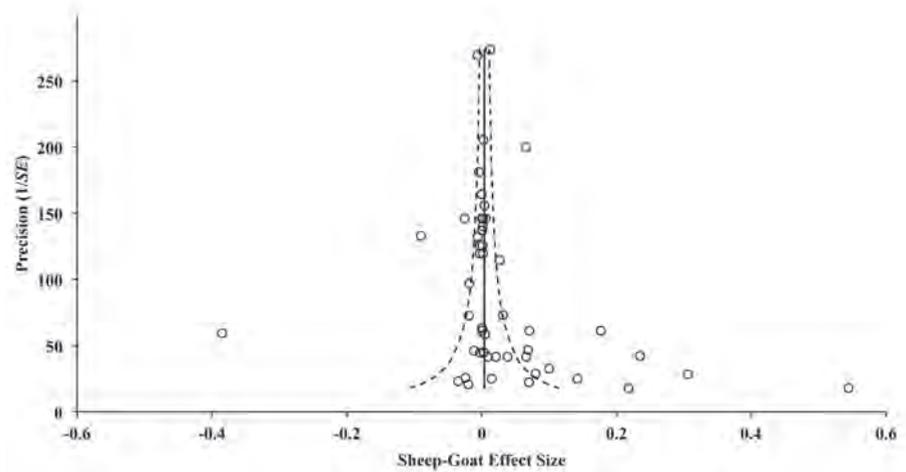


Figure 2. Funnel plot indicating outliers mainly to the right of the mean SGE line (central solid line; the dotted lines indicate regression curves)

H1: *Forced-choice sheep-goat studies produce statistical evidence of a communications anomaly known as ESP.* The full data set of studies ($N = 49$) yielded mean $z = 0.75$ ($SD = 1.57$; min. = -1.67 ; max. = $+5.76$), mean $ES = 0.045$ ($SD = 0.12$; min. = -0.04 ; max. = $+0.71$), and Stouffer $Z = 5.23$ ($p = 8.47 \times 10^{-8}$). This statistic is usually interpreted as a fixed effect, so we also employ a random effects test of significance using a one-sample t-test which gave a significant result, $t(48) = 3.33, p = .002$ (two-tailed). This statistic has greater value when generalizing to other studies (Pashler, 2002), and we find that 30 (61%) of the 49 studies have positive z scores, indicating a disproportionately greater number with positive effects, which we can say may be generally true in the population. Ten (20%) of the studies are independently significant ($\alpha < .05$), with nine indicating psi-hitting, but only one indicating psi-missing. Ninety-five percent confidence intervals (CIs), indicating a range in which the population parameter may lie, are as follows: z scores, $[0.30, 1.20]$; ES values, $[0.009, 0.080]$. Note that neither of these 95% CIs includes MCE (i.e., zero), which means that we can be 95% confident that zero (or no effect) is not a reasonable possibility for the true effect.

Using Darlington and Hayes's (2000, p. 503, Table 2) tabled data, we made a 'file-drawer' estimate of the number of unpublished and nonsignificant articles that would need to exist to reduce our significant finding to a chance result. If nine individual results are significant with $\alpha = .05$, then pooled p is less than or equal to .05 only if the total number of studies is up to 95. In other words, we find a "fail-safe N " of 95 unpublished studies must exist in total in the file-drawer. Note that such a number of unpublished studies is unlikely to exist. This method does not require the assumption that the mean effect size of the missing studies is zero, as "all the missing effect sizes may be highly negative" (Darlington & Hayes, 2000, p. 500).

H2: *There is a sheep-goat scoring differential.* Appendix A2 lists all 16 sheep-goat studies that gave actual trial and hit counts for sheep and goats. For comparative purposes, statistics are given for this smaller dataset, followed by statistics for the full dataset ($N = 49$). The 16 studies yielded a mean sheep-goat $CR(z) = 0.40$ ($SD = 1.14$; min. = -2.12 ; max. = $+1.84$), and mean $CR(ES) = 0.077$ ($SD = 0.20$; min. = -0.39 ; max. = $+0.54$). Ninety-five percent confidence intervals (CIs) are: z scores, $[-0.21, 1.01]$; ES values, $[-0.030, 0.183]$. Note that these 95% CIs include values of MCE (i.e., zero). Stouffer $Z = 1.58$ ($p = .057$), indicating a marginal sheep-goat effect. A one-sample t -test result was not significant, $t(15) = 1.39$, $p = .186$ (two-tailed). If we follow a random-effects model, generalization of the effect would not be permissible in this instance, even though 12 (75%) of the 16 studies had positive z scores — it is possible that the nonsignificant result is due to low N . Four (25%) are independently significant ($\alpha < .05$), with two indicating a sheep-goat effect, and two indicating a reversed sheep-goat effect.

In contrast, the full dataset ($N = 49$) yielded a mean sheep-goat $CR(z) = 0.24$ ($SD = 1.01$; min. = -2.17 ; max. = $+2.71$), mean $CR(ES) = 0.034$ ($SD = 0.12$; min. = -0.39 ; max. = $+0.54$), Ninety-five percent CIs are: z scores, $[-0.05, 0.53]$; ES values, $[-0.0003, 0.069]$. Note that both CIs include MCE. Stouffer $Z = 1.67$ ($p = .047$), indicating a significant sheep-goat effect. A single-sample t -test result was not significant, $t(48) = 1.65$, $p = .106$ (two-tailed). Of the 49 studies, 32 (65%) had positive z scores. Six (12%) are independently significant ($\alpha < .05$), with three indicating a sheep-goat effect, and three indicating a reversed sheep-goat effect.

Overall, both sets of results for the different datasets are similar, and both suggest a sheep-goat effect based on our primary outcome measure (Stouffer Z), though we are cautious about making inferences to the population given the CIs.

H3: *ES and SGE vary in strength according to target types.* We created five target types: Zener cards, pictures/photos, I Ching (hexagram) symbols, shapes (e.g., rectangles, doors), and racing game (greyhounds). We first tested the difference between target types for the full sample independent of the sheep-goat effect. A One-Way ANOVA test showed that ES varies significantly among target types, $F(4, 44) = 4.87$, $p = .002$ (two-tailed). Table 1 shows the mean ES values for each target type. The differences were between I Ching symbols and shapes ($p = .027$), I Ching symbols and Zener cards ($p = .033$), I Ching symbols and pictures/ photos ($p = .001$), and a marginally significant

difference between I Ching symbols and racing game (greyhounds; $p = .066$). There were no other significant ES differences.

Table 2 shows that I Ching symbols elicited the strongest mean SGE, which was significantly higher than the mean SGEs for pictures/photos and shapes. A One-Way ANOVA test showed that the SGE varies significantly among target types, $F(4, 45) = 3.34, p = .018$ (two-tailed). The differences were between I Ching symbols and pictures/photos ($p = .015$), and I Ching symbols and shapes ($p = .017$). There were no other significant ES differences.

Table 1

Forced-Choice Target Types: Mean ES Values and SDs

Target Type	<i>N</i>	Mean <i>ES</i>	<i>SD</i>
<i>I Ching</i> symbols	8	.195	.255
Shapes	9	.035	.032
Zener Cards	6	.021	.050
Pictures/photos	23	.008	.033
Racing Game	3	-.002	.007

Table 2

Forced-Choice Sheep-Goat Target Types: Mean CR(ES) Values and SDs

Target Type	<i>N</i>	Mean CR(ES)	<i>SD</i>
<i>I Ching</i> symbols	8	.159	.191
Zener Cards	6	.040	.075
Racing Game	3	.023	.015
Pictures/photos	23	.009	.057
Shapes	9	-.015	.147

H4: *Number of choices (k) per trial is positively related to z score.* There are six values of k : 2, 4, 4.07,¹ 5, and 6. The number of trials and hits are presented, with hit-rates and mean z scores, in Table 3. In accordance with Timm’s (2000) conjecture, increasing the number of target alternatives k leads to increased levels of significance in ESP experiments, which can be discerned from z scores. In fact, Table 3 shows that the highest hit-rates (and the most significant) are yielded by studies that use a ‘ $k = 5$ ’ design, and there is a steady trend across levels of k , which is true for overall ESP effects, and the sheep-goat effect (SGE). However, for overall ESP, and for the SGE, the trends are undermined by the results for the ‘ $k = 6$ ’ studies.

A Kendall’s tau-b test on k values and z scores was significant, $rT(47) = 0.19, p = .041$ (one-tailed). Thus, z scores tended to increase as k increased. However, CR(z) scores did not correlate significantly with k , $rT(47) = 0.05$,

¹ Derived from the harmonic mean of two k choices for I Ching second hexagrams = $[16/63 + 15/63]/2$.

$p = .331$ (one-tailed). While there may be some advantage in using high k -choice designs in forced-choice designs, there is no evidence that the SGE is enhanced by increases in k choices. Nevertheless, the ' $k = 5$ ' design seems nominally superior for both kinds of effects.

Table 3

Hit-Rates for Five Forced-Choice Sheep-Goat Studies by Number of Choices (K)

K-Choices	N	Trials	Hits	Hit-Rate (%)	Mean Z	PMCE	Mean CR(z)
2-choice	5	24,004	11,961	49.83	-0.52	.302	-0.23
4-choice	24	48,293	12,239	25.34	1.74	.041*	0.27
4.07-choice	3	354	101	28.53	1.67	.048*	0.24
5-choice	12	39,095	8,269	21.15	+5.68	<.001**	0.55
6-choice	5	1,685	294	17.45	+0.83	.203	-0.17

H5: *Bayesian analysis yields statistical evidence of the SGE.* A Bayesian estimate of the mean SGE and its 95% credible intervals was made using 'RStan' which is the R interface to 'Stan',² an open-source package for obtaining Bayesian inference using a variant of the Hamiltonian Monte Carlo method known as the 'No-U-Turn sampler'. In Bayesian analyses, θ is the parameter we are attempting to estimate based on our database: θ is conceptualised as a random variable and, thus, is considered to fluctuate over time (Field, 2009). We used a random hierarchical model. Our prior subjective belief regarding θ was 0.1 with a standard deviation of 100. The prior distribution of θ was specified as normal, and was estimated with 5,000 iterations, each starting with different and dispersed initial values for the model, and four chains applied to the random hierarchical model presented in Appendix C.

Figure 3 depicts the posterior probability distribution of θ . That is, the conditional probability that is assigned after the SGE of each of the 49 experiments is taken into account and the prior probability of θ is updated. The across-experiments μ (i.e., mean SGE; see Figure 3) was above chance. The 95% Highest Density Interval (HDI)³ of posterior probability related to the SGE ranged from 0.001 to 0.06 (mean SGE = 0.03). As shown in Figure 3, θ is included in the 95% HDI. Thus, the null hypothesis is rejected confirming the results obtained with the frequentist random model.

Post Hoc Analyses

Does the sheep-goat effect vary across psi modalities (i.e., clairvoyance, precognition, and telepathy)? The following test results are presented to give some idea as to the possible quantitative differences that may exist between these three psi modalities. We ran a Univariate ANOVA, entering the sheep-goat measure CR(ES) as Dependent Variable (DV), and Modality as the

² For details, see <<http://mc-stan.org/interfaces/rstan>>

³ The HDI indicates the most plausible 95% of the values in the posterior distribution.

Independent Variable (IV). Unfortunately, there were only three telepathy studies, and nine studies could not be classified as either clairvoyance or precognition (these were merged into a fourth ambiguous modality). The strongest SGE came from precognition, followed by telepathy and clairvoyance. However, there was no significant SGE difference between the modalities, $F(3, 28) = 1.78, p = .175$ (two-tailed).

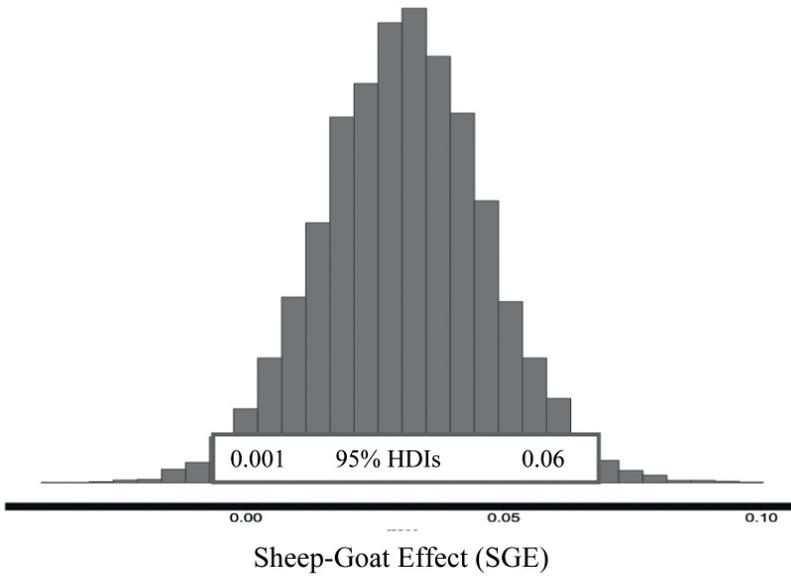


Figure 3. Highest Density Interval (HDI) related to the SGE (Mean SGE = 0.03). All values inside the HDI, 0.001 to 0.06, marked by vertical strokes | have higher credibility than values outside the interval, where the interval includes 95% of the respective distribution

Table 4

Psi Modality and SGE: Mean CR(ES) Values and SDs

Modality	N	Mean CR(ES)	SD
Precognition	20	.068	.141
Telepathy	3	.022	.116
Clairvoyance	17	.014	.124
Clairvoyance/Precognition	9	.0003	.003

Is the sheep-goat effect measurable using various paranormal belief measures? To ascertain whether the SGE depends on how one measures belief in the paranormal, the same Univariate ANOVA from the previous analysis included the range of psi measures listed above in Selection Criteria.

We formed one belief measure with the 1-, 2-, 3-, and 4-item paranormal belief questions combined into a single category. The other four measures were: ‘ASGS’, ‘ISGS’, ‘PBS’, and ‘AEI-PB’.

The SGE was more readily detected with the PBS and the AEI-PB, followed by the ASGS, whereas the ISGS and short scales (1 to 4 questions) were very poor at measuring the SGE. However, the SGE was not significantly different among measures, $F(4, 28) = 1.87, p = .144$ (two-tailed).

Table 5

Paranormal Belief Measure and SGE: Mean CR(ES) Values and SDs

Belief Measure	<i>N</i>	Mean CR(ES)	<i>SD</i>
PBS	5	.184	.221
AEI-PB	3	.089	.127
ASGS	17	.039	.078
ISGS	9	-.003	.009
1, 2, 3, or 4 item PBQs	15	-.010	.120

Does the sheep-goat effect vary across investigators? To ascertain whether the SGE was the result of a high success rate amongst a limited pool of experimenters and/or laboratories, the same Univariate ANOVA included experimenter/laboratory groups. We could not test Experimenter/Laboratory interaction, as we found that a number of experimenters had worked in more than one specific laboratory. We formed five mutually exclusive groups with at least two studies in each: ‘Haraldsson’, ‘Dalkvist / Westerlund’, ‘Roe’, ‘Storm / Thalbourne’, and ‘Miscellaneous’ (this last group was comprised of mainly single studies from North and South American researchers, as this was the best way to include all studies in the analysis while still maintaining optimum degrees of freedom).

The ‘Miscellaneous’ group produced the strongest sheep-goat effects, followed by ‘Storm / Thalbourne’ and ‘Roe’. However, the SGE was not significantly different among groups, $F(4, 28) = 0.09, p = .985$ (two-tailed).

Table 6

Investigators and SGE: Mean CR(ES) Values and SDs

Investigator Group	<i>N</i>	Mean CR(ES)	<i>SD</i>
‘Miscellaneous’	10	.109	.233
‘Storm/Thalbourne’	10	.053	.099
‘Roe’	14	.012	.030
‘Haraldsson’	13	-.002	.008
‘Dalkvist’	2	-.048	.059

Are there declines in the sheep-goat effect? The relationship between the sheep-goat CR(ES) values and year of publication is positive and significant,

$r(47) = 0.33, p = .019$ (two-tailed), although the same is not true for (non-sheep-goat) ESP ES values and year of publication, $r(47) = 0.21, p = .139$ (two-tailed) — see Figure 4. In other words, the SGE has increased across the span of about 20 years, though ESP effects generally have not increased significantly, but in both cases (as shown in the second paragraph of the section Quality Ratings), increases in effects have been independent of study quality.

The linear trend line formula for ESP (in the form $y = m.x + b$, where m is the slope and b is the y-intercept) is: $ES = (.0037 \times \text{YEAR}) - 7.4$. The linear trend line formula for SGE is: $CR(ES) = (0.0056 \times \text{YEAR}) - 11.26$. The SGE slope is slightly steeper (R^2 linear = .112) compared to ESP (R^2 linear = .046). The incline for the SGE is significant, $F(1, 47) = 5.92, p = .019$ (two-tailed), but the incline for ESP is not significant, $F(1, 47) = 2.26, p = .139$ (two-tailed).

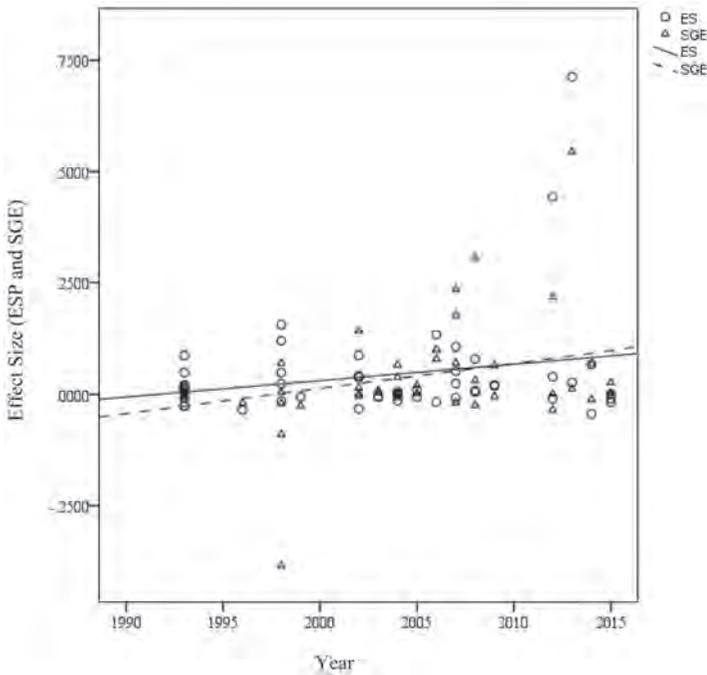


Figure 4. Scatterplot of ESP effect and sheep-goat effect (SGE) inclining over nearly two decades (1994–2015). Only the SGE slope (dotted line) is significant

Is there predictive consistency in the SGE database? We determined counts of four types of study in the large dataset ($N = 49$): (i) significant SGE, (ii) non-significant SGE, (iii) non-significant reversed SGE, and (iv) significant reversed SGE. We compared these counts to the statistics presented by Lawrence (1993, p. 82). Figure 5 is a pie chart showing study outcome percentages. These percentages vary somewhat from those of both Lawrence (1993), and Palmer (1971):

- Significant SGE (i.e., psi-hitting) is much lower at 6% compared with Lawrence's 27.3%, and Palmer's 16%;
- Non-significant SGE is higher at 62% compared with Lawrence's 47%, but comparable with Palmer's 69%;
- Non-significant reversed SGE at 26% is comparable with Lawrence's 24.2%, but higher than Palmer's 15%;
- Significant reversed SGE is higher at 6% compared with Lawrence's 1.5%, and Palmer's < 1%.

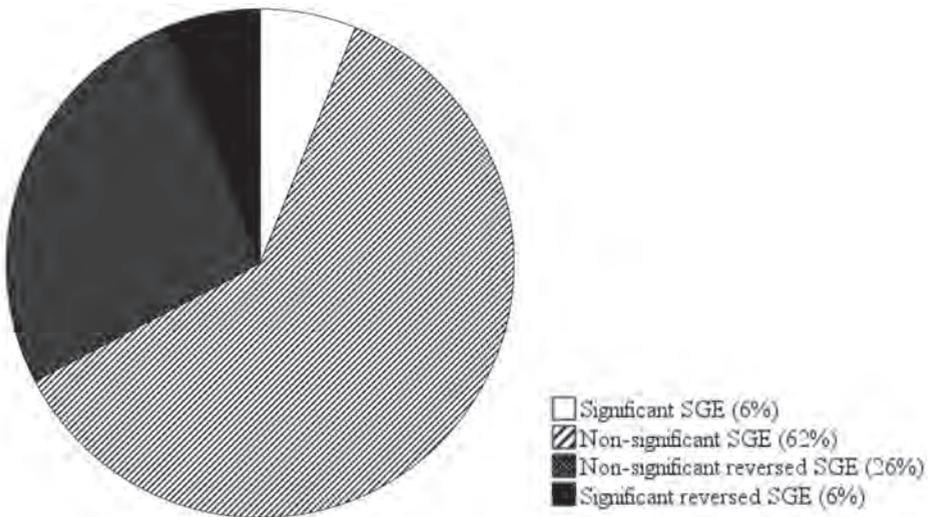


Figure 5. Study outcome percentages (N = 49)

DISCUSSION

The ESP effect yielded a significant Stouffer Z , and we find that results for the two different datasets ($n = 16$, and $N = 49$) were similar (including the two significant Stouffer Z values), both of which suggest an SGE. Interestingly, the SGE effect is as convincing in the much smaller dataset ($n = 16$), that had actual sheep and goat trial counts and hit counts, as it is in the full dataset ($N = 49$), 33 studies of which were a hotchpotch of correlational values needing conversion. Furthermore, demonstrating an SGE with a database of only 16 studies is particularly impressive. Consequently, a point we must raise is that actual sheep and goat data may be crucial in these sorts of analyses, and it behoves investigators to bear this in mind in future studies.

Lawrence (1993) presented a mean z score of 0.96 compared to our lower values of 0.40 ($n = 16$) and 0.24 ($N = 49$), yet Lawrence's mean SGE of 0.029 was lower than our two SGE values, 0.077 ($n = 16$) and 0.034 ($N = 49$). As far as effect size is concerned, we have not only replicated Lawrence's result, but

have shown that the SGE has increased in mean strength since his study — an effect which is not due to improvements in design quality.

In testing H3, we found that target types matter, and so may the number of choices in a given trial. However, while ecologically-based targets may be a better alternative, there was no statistical evidence that investigators should shift to ' $k = 5$ ' designs, though that is our recommendation. In finding that using I Ching symbols as target type made a considerable (significant) difference to the SGE, it may be that these resonate more with sheep than goats because sheep put more stock in the outcomes (i.e., meaningful readings are provided with each outcome hexagram that the participants generate). In other words, the 'reward' *per se* is valued more highly by sheep, whereas goats (being sceptical of divinatory systems) simply do not care. Perhaps this is because other symbols (even 'greyhounds' where there should be vested financial interest) tend not to stir the imagination and/or have little ecological validity in the laboratory if meaning in the task is paramount. In pseudo-gambling studies, perhaps much higher stakes would rouse the necessary motivation. We suggest future forced-choice studies must re-consider the importance of these factors.

Finally, we used a Bayesian approach to estimate the probability of the mean SGE. The 95% HDI of posterior probabilities concerning the SGE did not include zero. Thus, the null hypothesis was rejected. Importantly, these Bayesian findings support the significant result we obtained using a classical frequentist approach. Nevertheless, we are mindful that "Bayesian methods utilize [a] 'degree of belief' interpretation of probability to model all uncertainty" (Utts, Norris, Suess, & Johnson, 2010, p. 2). Indeed, the statistician Gelman (2008) contended that, "as scientists we should be concerned with objective knowledge rather than subjective belief" (p. 2).

Our post hoc analyses revealed some interesting and pertinent results. Strongest sheep-goat effects came from precognition studies, but they were not significantly different from any other psi modality. And, if our results are anything to go by, while we would advise investigators to use 'tried-and-true' (i.e., established) paranormal belief measures wherever possible, choice of measure may come down to giving greater credence to the hit-rates (see Table 5) than to a significant p value, so that a tried-and-true belief measure might mean, for example, the R-PBS, or the AEI-PB, or even the ASGS. Lawrence (1993) also found no evidence that the SGE depends on how paranormal belief is measured. He did find that the ASGS was matched in performance by Schmeidler's (1943) Criterion 1 question ("Do you believe it is possible that ESP can be shown under the conditions of this experiment?") with an SGE score of 0.032 for both measures. In our meta-analysis, studies that used the ASGS produced a mean SGE of 0.039, on par with Lawrence's statistic. Lawrence (1993, p. 80) also commented on the conspicuous absence in parapsychology of the "Paranormal Belief Scale (Tobacyk and Milford, 1983)". Times have changed — the R-PBS (or the PBS) was used in five studies, and it proved to be the best performer, though not significantly so. It may be argued that behaviour is more predictable when attitudes are measured most narrowly whereas the R-PBS, taken to be a diffuse measure

of paranormal belief, is not a precise measure of attitude, and therefore not a good predictor of psi behaviour. While we agree that measuring attitude with precision may tend to be a better way of predicting behaviour, perhaps the psi behaviours of sheep and goats are an exception to the rule, so that broader measures of paranormal belief more accurately identify sheep and goats. As Lawrence says, “a sheep on one scale may be a goat on another” (p. 80), and we do not put this down to how the data is split, so due caution is warranted.

We found no significant differences between investigator groups. Though there seemed to be some performance differences, we must accept that the variations are within MCE, which must surely be good news for researchers because it means that everyone is working to the same standards and proven expectations.

There was no evidence of chronological declines in the mean ES, or the SGE — quite the reverse was found for the SGE. But since quality is not improving in tandem with these increases, any optimistic response should be tempered by our failure to explain why it is happening, unless it is in the experimental designs. Figure 1 shows that the most impressive period of growth in the SGE was from 1998 to 2007, but from 2007 to the present, SGE growth has been slower. A study of these periods may reveal which designs are superior. Lawrence (1993) found no evidence that study quality was related to the SGE, or that quality had improved between 1947 and 1993. Oddly, Lawrence suggested that the SGE had declined over the years, but he presented an extremely weak correlation to support that claim, “ $r(72) = -.083, p. = 80$ ”, backed up with a non-significant t test result, “ $t(72) = .705, p = .48$ ” (p. 80). Given our result of no statistical evidence of declines in either SGE or quality over our test period (1994 to 2015), coupled with Lawrence’s results, we extend the period of forced-choice ESP testing to nearly five decades (i.e., 1947 to 2015), during which time, the SGE and study quality have remained consistent.

We hasten to mention the fact that quality has been optimal in a majority of cases, and even though there is room for improvement, it does not seem to make a difference to the SGE given current forced-choice designs. Investigators may argue that they do everything in their power to produce unassailable effects for the sake of scientific truth and to keep the critics and sceptics at bay, but if the SGE can be strengthened (or for that matter, the ESP effect in general), the key may lie more in design intricacies and other issues, such as target type. For example, if we glance again at Tables 1 and 2, we see that ecological considerations in relation to target types might be key in our attempts to answer the question, ‘What is ESP for?’ This is hardly a new question for parapsychologists (Broughton, 1988), and we have raised these issues elsewhere in the context of the ecological use of psi (see Storm et al., 2012, p. 260).

Our inspection of predictive consistency of the SGE produced different outcomes from those of Lawrence (1993) and Palmer (1971), as far as significant hit-rates for both sheep and goats are concerned. While non-significant rates for both SGE and reversed-SGE were on a par with at least one past estimate, there was a drop in sheep performances, and a rise from

goats. The decrease from sheep is disconcerting, but the increase from goats is consistent with our expectations that goats, if they are to demonstrate any kind of psi in ESP tests, should demonstrate psi-missing.

Believers and sceptics may debate all the above findings ad nauseum for there may be cause for dispute given some degree of ambiguity in the test results. Hopefully, debate will hinge on questions of theoretical interest such as why a performance difference between sheep and goats should come about in the first place (i.e., how does a psychological mental state trigger a paranormal effect)? Nevertheless, some significant effects are difficult to dispute. For that reason, Palmer's (1971, p. 405) conclusion, which is repeated in Palmer (1972), is effectively the same nearly half a century later: "the data presently available support the hypothesis of a genuine SGE" (p. 1).

Conclusion

The forced-choice datasets have never been larger, with two meta-analyses — Lawrence's (1993), and the present study — suggesting a genuine SGE for the domain, but the effect is weak in both cases. One caveat worth repeating is that the SGE should not be taken as conclusive or necessarily reliable, and it should go without saying that the SGE cannot be taken as indicative of other domains.

For the future, it is also worth repeating that researchers need to give more attention to investigating the SGE via the administration of sheep-goat measures, and record hits for sheep and for goats — failing to do so is a costly oversight given that the arguably better database with actual sheep-goat data ($n = 16$) is relatively small compared with those tested in other meta-analyses. This situation may have eventuated because researchers have shifted their focus away from the SGE, perhaps because it is nowadays a taken-for-granted assumption, yet far from categorically proven.

Is there a perfect forced-choice design? We would answer in the affirmative and recommend that investigators (i) gather sheep and goat data (trial and hit counts); (ii) use either the R-PBS, the AEI-PB, or the ASGS to measure paranormal belief; (iii) implement a five-choice design; (iv) adopt ecologically relevant and valid set-ups in the laboratory (although Zener card symbols do not fare too badly); and (v) test precognition which is actually easier to implement, and is a design advantage in that it eliminates sensory leakage problems because target selection comes after participants' responses rather than before as in telepathy and clairvoyance designs.

In closing, we note that Lawrence's (1993) advice, "one area for future SGE research is to look at it in free response settings" (p. 83), has not been heeded. More than two decades later, the SGE in the free-response domain is still largely unexplored. In that regard, Lawrence also mentions that it would be worthwhile analysing goat's mentation "to get a grasp of what psi missing in more ecologically valid circumstances actually looks like" (p. 83). We have already covered this point, and must wait to see what investigators can dream up in the coming years.

ACKNOWLEDGEMENTS

The first author thanks the Society for Psychical Research for financial support of this project. Our thanks go to Dr Lorenzo Di Risio for help in the quality ratings of the studies.

*Brain and Cognition Centre
School of Psychology
University of Adelaide
Adelaide, South Australia 5005
lance.storm@adelaide.edu.au*

*Dipartimento di Psicologia Generale
Università di Padova
Padua, Italy
patrizio.tressoldi@unipd.it*

REFERENCES

- Baptista, J., Derakhshani, M., & Tressoldi, P. (2015). Explicit anomalous cognition: A review of the best evidence in ganzfeld, forced-choice, remote viewing and dream studies. In E. Cardeña, J. Palmer & D. Marcusson-Clavertz (Eds.), *Parapsychology. A handbook for the 21st century* (pp. 192–214). Jefferson, NC: McFarland.
- Bem, D.J., Utts, J., & Johnson, W.O. (2011). Must psychologists change the way they analyze their data? *Journal of Personality and Social Psychology*, *101*, 716–719. doi:10.1037/a0024777
- 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 Loocke (Ed.), *The physical nature of consciousness* (pp. 269–292). New York, NY: Benjamins.
- *Billows, H., & Storm, L. (2015). Believe it or not: A confirmatory study on predictors of paranormal belief, and a psi test. *Australian Journal of Parapsychology*, *15*, 7–35.
- Broughton, R.S. (1988). If you want to know how it works, first find out what it's for. In D.H. Weiner & R.L. Morris (Eds.), *Research in parapsychology 1987* (pp. 187–202). Metuchen, NJ: Scarecrow Press.
- *Dalkvist, J., & Westerlund, J. (1998). Five experiments on telepathic communication of emotions. *Journal of Parapsychology*, *62*, 219–253.
- Darlington, R.B., & Hayes, A.F. (2000). Combining independent p values: Extensions of the Stouffer and binomial methods. *Psychological Methods*, *5*(4), 496–515.
- Edge, H.L., Morris, R.L., Palmer, J., & Rush, J.H. (1986). *Foundations of parapsychology*. London: Routledge & Keegan Paul.
- Field, A. (2009). *Discovering statistics using SPSS (3rd ed.)*. Los Angeles: Sage.
- Gelman, A. (2008). *Objections to Bayesian statistics*. Retrieved, March 13, 2011, from: ba.stat.cmu.edu/journal/2008/vol03/issue03/gelman.pdf.
- Gallagher, C., Kumar, V.K., & Pekala, R.J. (1994). The anomalous experiences inventory: Reliability and validity. *Journal of Parapsychology*, *58*, 402–428.
- *Haraldsson, E. (1993). Are religiosity and belief in an afterlife better predictors of ESP performance than belief in psychic phenomena? *Journal of Parapsychology*, *57*, 259–273.
- Haraldsson, E., & Houtkooper, J.M. (1992). Effects of perceptual defensiveness, personality, and belief on extrasensory perception tasks. *Personality and Individual Differences*, *13*(10), 1085–1096.
- *Haraldsson, E., Houtkooper, J.M., Schneider, R., & Bäckström, M. (2002). Perceptual defensiveness and ESP performance: Reconstructed DMT ratings and psychological

- correlates in the first German DMT-ESP experiment. *Journal of Parapsychology*, 66, 249–270.
- *Hitchman, G.A.M., Roe, C.A., & Sherwood, S.J. (2012). A re-examination of nonintentional precognition with openness to experience, creativity, psi beliefs, and luck beliefs as predictors of success. *Journal of Parapsychology*, 76(1), 109–145.
- *Hitchman, G.A.M., Roe, C.A., & Sherwood, S.J. (2015). The relationship between lability and performance at intentional and non intentional versions of an implicit PMIR-type psi task. *Journal of Parapsychology*, 79(1), 65–86.
- Hövelmann, G.H. (2015). On the usefulness of parapsychology for science at large. In E. Cardena, J. Palmer, & D. Marcussen-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 389–398). Jefferson, NC: McFarland.
- Honorton, C. (1985). Meta-analysis of psi ganzfeld research: A response to Hyman. *Journal of Parapsychology*, 49, 51–91.
- Honorton, C., & Ferrari, D.C. (1989). “Future telling”: A meta-analysis of forced-choice precognition experiments, 1935–1987. *Journal of Parapsychology*, 53, 281–308.
- *Houran, J., & Lange, R. (2012). I Ching outcomes from experimental manipulations of transliminality and paranormal belief. *Australian Journal of Parapsychology*, 12, 39–58.
- *Houtkooper, J.M. (2003). An ESP experiment with natural and simulated sferics: Displacement scores and psychological variables. *European Journal of Parapsychology*, 18, 49–64.
- *Houtkooper, J.M., Schienle, A., Stark, R. & Vaitl, D. (1999). Atmospheric electromagnetism: A possible disturbing influence of natural sferics on ESP. *Perceptual and Motor Skills*, 89, 1179–1192.
- Hyman, R. (1985). The ganzfeld psi experiment: A critical appraisal. *Journal of Parapsychology*, 49, 3–49.
- Irwin, H.J. (1993). Belief in the paranormal: A review of the empirical literature. *Journal of the American Society for Psychical Research*, 87, 1–39.
- Jin, Z.-C., Zhou, X.-H., & He, J. (2015). Statistical methods for dealing with publication bias in meta-analysis. *Statistics in Medicine*, 34, 343–360. DOI: 10.1002/sim.6342
- *Lange, R., & Houran, J. (2013). Towards a replicable formula for significant I Ching outcomes. *Australian Journal of Parapsychology*, 13, 9–25.
- Lau, J. Ioannidis, J.P.A., Terrin, N., Schmid, C. H., & Olkin, I. (2006). The case of the misleading funnel plot. *The British Medical Journal*, 333, 597–600.
- 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* (pp. 75–86). Durham, NC: Parapsychological Association.
- *Luke, D.P., Delaney, D., & Sherwood, S.J. (2008). Psi may look like luck: Perceived luckiness and beliefs about luck in relation to precognition. *Journal of Society for Psychical Research*, 72, 193–207.
- *Luke, D. & Morin, S. (2014). Exploration of the validity and utility of a reward contingency in a non-intentional forced-choice precognition task. *Journal of the Society for Psychical Research*, 78, 207–218.
- *Luke, D. & Zychowicz, K. (2014). Comparison of outcomes with non-intentional and intentional precognition tasks. *Journal of Parapsychology*, 78(2), 223–234.
- *Luke, D., Zychowicz, K., Richterova, O., Tjurina, I. & Polonnikova, J. (2012). A sideways look at the neurobiology of psi: Precognition and circadian rhythms. *NeuroQuantology*, 10(3), 580–590.
- Palmer, I. (1971). Scoring in ESP tests as a function of belief in ESP: Part I. The sheep-goat effect. *Journal of the American Society for Psychical Research*, 65, 373–408.
- Palmer, J. (1972). Scoring in ESP tests as a function of belief in ESP: Part II. Beyond the sheep-goat effect. *Journal of the American Society for Psychical Research*, 66, 1–25.

- Palmer, J. (1977). Attitudes and personality traits in experimental ESP research. In B.B. Wolman (Ed.), *Handbook of parapsychology* (pp. 175–201). New York: Van Nostrand Reinhold.
- *Palmer, J. (2009). Decision augmentation in a computer guessing task. *Journal of Parapsychology*, *73*, 119–136.
- *Parra, A., & Argibay, J.C. (2007). Comparing psychics and non-psychics through a ‘token-object’ forced-choice ESP test. *Journal of the Society for Psychical Research*, *71*, 80–90.
- Pashler, H. (Ed.) (2002). *Steven’s handbook of experimental psychology: Methodology in experimental psychology, Volume 4 (3rd ed.)*. New York: Wiley.
- *Pitman, J.A., & Owens, N.E. (2004). The effect of manipulating expectations both before and during a test of ESP. *Journal of Parapsychology*, *68*, 45–63.
- *Roe, C.A., Davey, R., & Stevens, P. (2003). Are ESP and PK aspects of a unitary phenomenon? A preliminary test of the relationship between ESP and PK. *Journal of Parapsychology*, *67*, 343–366.
- *Roe, C.A., Davey, R., & Stevens, P. (2004). Arousal and performance at ESP and PK tasks using a common protocol. *European Journal of Parapsychology*, *19*, 29–43.
- *Roe, C.A., Davey, R., & Stevens, P. (2005). Are ESP and PK Aspects of a unitary phenomenon? The effects of deception when testing the relationship between ESP and PK. *Journal of the Society for Psychical Research*, *69*, 18–32.
- Rosenthal, R. (1984). *Meta-analytic procedures for social research*. Beverley Hills, CA: Sage.
- Rouder, J.N., & Morey, R.D. (2011). A Bayes factor meta-analysis of Bem’s ESP claim. *Psychonomic Bulletin and Review*, *18*, 682–689. doi: 10.3758/s13423-011-0088-7
- *Rudski, J.M. (2002). Hindsight and confirmation biases in an exercise in telepathy. *Psychological Reports*, *91*, 899–906.
- Schmeidler, G.R. (1943). Predicting good and bad scores in a clairvoyance experiment: A preliminary report. *Journal of the American Society for Psychical Research*, *37*, 103–110.
- Schmeidler, G.R. (1945). Separating the sheep from the goats. *Journal of the American Society for Psychical Research*, *39*, 47–49.
- Schmeidler, G.R., & McConnell, R.A. (1973). *ESP and personality patterns*. Westport, CT Greenwood.
- *Simmonds-Moore, C.A., & Moore, S.L. (2009). Exploring how gender role and boundary thinness relate to paranormal experiences, beliefs and performance on a forced-choice clairvoyance task. *Journal of the Society for Psychical Research*, *73*, 129–143.
- *Steinkamp, F., (1998). Psychology, psi, and the Web. *Journal of the American Society for Psychical Research*, *92*(3), 256–278.
- Steinkamp, F., Milton, J., & Morris, R.L. (1998). A meta-analysis of forced-choice experiments comparing clairvoyance and precognition. *Journal of Parapsychology*, *62*, 193–218.
- Sterne, J.A.C., & Harbord, R.M. (2004). Funnel plots in meta-analysis. *The Stata Journal*, *4*(2), 127–141.
- *Storm, L. (2006). A parapsychological investigation of the I Ching: The relationships between psi, intuition, and time perspective. *Journal of Parapsychology*, *70*, 121–141.
- *Storm, L. (2008). Investigations of the I Ching: I. Relationships between psi, time perspective, paranormal belief and meaningfulness. *Australian Journal of Parapsychology*, *8*, 103–127.
- *Storm, L., Ertel, S., & Rock, A.J. (2013). Paranormal effects and behavioural characteristics of participants in a forced-choice psi task: Ertel’s ball selection test under scrutiny. *Australian Journal of Parapsychology*, *13*, 111–131.
- *Storm, L., & Thalbourne, M.A. (1998–1999). The transliminal connection between paranormal effects and personality in an experiment with the I Ching. *European Journal of Parapsychology*, *14*, 100–124.
- *Storm, L., & Thalbourne, M.A. (2005). The effect of a change in pro attitude on paranormal performance: A pilot study using naïve and sophisticated skeptics. *Journal of Scientific Exploration*, *19*, 11–29.

- Storm, L., Tressoldi, P.E., & Di Risio, L. (2012). Meta-analyses of ESP studies 1987–2008: Assessing the success of the forced-choice design in parapsychology. *Journal of Parapsychology*, 76, 243–273.
- Thalbourne, M.A. (1995). Further studies of the measurement and correlates of belief in the paranormal. *Journal of the American Society for Psychical Research*, 89, 233–247.
- *Thalbourne, M. (1996). An attempt to predict precognition scores using transliminality-relevant variables. *Journal of the Society of Psychical Research*, 62, 129–140.
- Thalbourne, M.A. (2003). *A glossary of terms used in parapsychology*. Charlottesville, VA: Puente.
- Thalbourne, M.A. (2010). The Australian Sheep-Goat Scale: Development and empirical findings. *Australian Journal of Parapsychology*, 10, 5–39.
- Timm, U. (2000). The problem of finding an optimal measure of effect size for psi experiments. *Journal of Parapsychology*, 64, 253–254.
- Tobacyk, J.J. (2004). A revised Paranormal Belief Scale. *International Journal of Transpersonal Studies*, 23, 94–98.
- Tressoldi, P.E. (2011). Extraordinary claims require extraordinary evidence: The case of non-local perception, a classical and Bayesian review of evidences. *Frontiers in Psychology*, 2, 117. doi:10.3389/fpsyg.2011.00117
- Utts, J., Norris, M., Suess, E., & Johnson, W. (2010). The strength of evidence versus the power of belief: Are we all Bayesians? In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society*. Voorburg, the Netherlands: International Statistical Institute.
- Walker, D.A. (2013). Converting Kendall's Tau For correlational or meta-analytic analyses. *Journal of Modern Applied Statistical Methods*, 2(2), 525–530.
- *Walsh, K., & Moddel, G. (2007). Effect of belief on psi performance on a card guessing task. *Journal of Scientific Exploration*, 21, 501–510.
- *Westerlund, J., & Dalkvist, J. (2004). A test of predictions from five studies on telepathic group communication of emotions. *Proceedings of presented papers: The Parapsychological Association 47th Annual Convention* (pp. 268–277). Durham, NC: Parapsychological Association.
- * Wilson, S., & Hamlin, I. (2007). Implicit learning in a card prediction task. *European Journal of Parapsychology*, 22, 3–29.

APPENDIX A1: Forced-Choice Studies by Author/Year, Task, Trials, Hits, Z , and ES Values (1993 to 2015)

Study	Author(s)	Year	Task	Trials	Hits	Z	ES
1	Billows & Storm	2015	CL	2100	423	0.00	0.000
2	Dalkvist & Westerlund	1998	TE	4410	2172	-0.98	-0.015
3a	Haraldsson I	1993	CL/PR	2960	758	0.74	0.014
3b	Haraldsson II	1993	CL/PR	2960	768	1.17	0.022
3c	Haraldsson III	1993	CL/PR	3280	782	-1.51	-0.026
3d	Haraldsson IV	1993	CL/PR	4320	1016	5.76	0.088
3e	Haraldsson V	1993	CL/PR	3680	907	-0.48	-0.008
3f	Haraldsson VI	1993	CL/PR	3520	895	0.56	0.009
3g	Haraldsson VIII	1993	CL/PR	4000	956	-1.59	-0.025
3h	Haraldsson IX	1993	CL/PR	4000	1031	1.11	0.018
3i	Haraldsson X	1993	CL/PR	4000	1086	3.12	0.049
4a	Haraldsson, Houtkooper, Schneider, & Bäckström	2002	CL	2680	719	2.16	0.042
4b	Haraldsson, Houtkooper, Schneider, & Bäckström	2002	PR	2680	632	-1.67	-0.032
5	Hitchman, Roe & Sherwood	2012	PR	750	201	1.10	0.040
6	Hitchman, Roe & Sherwood	2015	PR	500	245	-0.22	-0.010
7	Hitchman, Roe & Sherwood	2015	PR	500	247	-0.40	-0.018
8	Houran & Lange	2012	PR	60	27	3.43	0.443
9	Houtkooper	2003	CL	4560	1130	-0.32	-0.005
10	Houtkooper, Schienle, Stark, & Vaitl	1999	CL	4000	991	-0.31	-0.005
11	Lange & Houran	2013	PR	60	34	5.52	0.713
12	Luke, Delanoy & Sherwood	2008	PR	1000	285	2.52	0.080
13	Luke & Morin	2014	PR	410	115	1.37	0.068
14	Luke & Zychowicz	2014	PR	400	92	-0.87	-0.044
15	Luke, Zychowicz, Richterova, et al.	2012	PR	100	48	-0.10	-0.010
16	Palmer	2009	PR	6400	1330	1.55	0.019
17	Parra & Argibay	2007	CL	444	228	0.52	0.025
18a	Pitman & Owens, Exp 1	2004	CL	275	53	-0.23	-0.014
18b	Pitman & Owens, Exp 2	2004	CL	600	121	0.05	0.002

19	Roe, Davey & Stevens	2003	CL	240	39	-0.09	-0.006
20	Roe, Davey & Stevens	2004	CL	240	41	0.09	0.006
21	Roe, Davey & Stevens	2005	CL	240	39	-0.09	-0.006
22a	Rudski	2002	TE	100	24	0.88	0.088
22b	Rudski	2002	TE	100	22	0.38	0.038
23	Simmonds-Moore & Moore	2009	CL	5250	1095	1.54	0.021
24a	Steinkamp	1998	CL	489	91	1.09	0.049
24b	Steinkamp	1998	PR	476	84	0.52	0.024
25a	Storm	2006	PR	200	48	-0.24	-0.017
25b	Storm	2006	PR	153	47	1.65	0.133
26a	Storm	2008	PR	150	38	0.07	0.006
26b	Storm	2008	PR	122	30	0.09	0.008
27	Storm, Ertel & Rock	2013	CL	12000	2529	2.93	0.027
28a	Storm & Thalbourne	1998-9	PR	93	30	1.50	0.156
28b	Storm & Thalbourne	1998-9	PR	79	24	1.06	0.119
29	Storm & Thalbourne	2005	CL	6750	1377	0.81	0.010
30	Thalbourne	1996	PR	990	232	-1.10	-0.035
32a	Walsh & Moddel	2007	CL	600	146	2.60	0.106
32b	Walsh & Moddel	2007	CL	600	133	1.28	0.052
33	Westerlund & Dalkvist	2004	CL	18150	9069	-0.08	-0.001
34	Wilson & Hamlin	2007	PR	1760	434	-0.30	-0.007

APPENDIX A2: Sheep-Goat Forced-Choice Studies by Author/Year, Task, Trials, Hits, Z, and ES Values (1994 to 2015)

Study	Author(s)	Year	Task	Total Trials	Total Hits	Sheep Trial	Sheep Hits	Goat Trial	Goat Hits	Z	ES	Sheep-Z	Sheep ES	Goat Z	Goat ES
1	Billows & Storm	2015	CL	2100	423	1100	226	1000	197	0.000	0.000	0.410	0.012	-0.200	-0.006
2	Dalkvist & Westerlund	1998	TE	4410	2172	2190	1043	2220	1129	-0.980	-0.015	-2.200	-0.047	0.790	0.017
8	Houran & Lange	2012	PR	60	27	30	15	30	12	3.430	0.443	2.780	0.508	1.640	0.299
11	Lange & Houran	2013	PR	60	34	30	20	30	14	5.520	0.713	4.630	0.845	2.400	0.438
16	Palmer	2009	PR	6400	1330	3200	695	3200	635	1.550	0.019	2.410	0.043	-0.200	-0.004
17	Parra & Argibay	2007	CL	444	228	264	145	180	83	0.520	0.025	1.540	0.095	-0.970	-0.072
18b	Pitman & Owens (1)	2004	CL	275	53	75	16	200	37	-0.230	-0.014	0.140	0.016	-0.440	-0.031
18a	Pitman & Owens (2)	2004	CL	600	121	300	61	300	60	0.050	0.002	0.070	0.004	0.000	0.000
24a	Steinkamp	1998	CL	489	91	446	78	43	13	1.090	0.049	0.400	0.019	2.180	0.332
24b	Steinkamp	1998	PR	476	84	456	81	20	3	0.520	0.024	0.570	0.027	-0.100	-0.022
26a	Storm	2008	PR	150	38	83	25	67	13	0.070	0.006	0.950	0.104	-0.920	-0.112
26b	Storm	2008	PR	122	30	68	16	54	14	0.090	0.008	-0.070	-0.008	0.060	0.008
27	Storm, Ertel & Rock	2013	CL	12000	2529	5940	1263	6060	1266	2.930	0.027	2.420	0.031	1.720	0.022
29	Storm & Thalbourne	2005	CL	6750	1377	4350	889	2400	488	0.810	0.010	0.700	0.011	0.380	0.008
32a	Walsh & Moddel	2007	CL	600	146	300	81	300	65	2.600	0.106	2.960	0.171	0.650	0.038
32b	Walsh & Moddel	2007	CL	600	133	300	70	300	63	1.280	0.052	1.370	0.079	0.360	0.021

APPENDIX B: Calculations of Statistics Used in the Analyses (Z Scores and Effect Sizes)

All z scores for studies where $n \leq 1000$ (where n = number of trials) were calculated from Exact Binomial P values:⁴

$$P_{(k \text{ out of } n)} = \frac{n!}{k!(n-k)!} * (p^k)(q^{n-k})$$

Formula 1

where n = number of trials, k = number of hits, p = probability of a hit, and q = probability of no hit (i.e., $1 - p$).

If $np \geq 5$ and $nq \geq 5$, binomial probabilities were estimated by way of the binomial approximation of the normal distribution, according to the formula:

$$z = \frac{(k - M) \pm .5}{\Sigma}$$

Formula 2

where $M = np$ (the mean of the binomial sampling distribution), and $\sigma = \sqrt{npq}$ (the SD of the binomial sampling distribution). NB: This formula includes a continuity correction ($\pm .5$) that yields negative z scores for chance scoring. One can see that repeated use of Formula 2, given the appropriate data, will yield conservative outcomes in meta-analyses because the mean z and ES values are consistently nudged in a negative direction.

The critical ratio of the SGE difference:

$$CRd(z) = \frac{(H_1/n_1) - (H_2/n_2) \pm 0.5(1/n_1 + 1/n_2)}{\sqrt{[(pq/n_1) + (pq/n_2)]}}$$

Formula 3

where H_1 and H_2 are the total number of hits in each group, n_1 and n_2 are the corresponding numbers of trials, p = chance probability of a hit (the reciprocal of the number of target alternatives), q = chance probability of a miss (i.e., $1 - p$), and ± 0.5 means adjust towards zero.

Haraldsson (1993) used Kendall's tau as an effect size measure of the relationship between paranormal belief and ESP task performance. Walker (2003) states: "With a presented value of tau, we can calculate a Pearson's r using Kendall's (1970, p. 126) formula" (p. 526):

$$r = \sin(.5 * \pi * \text{tau})$$

Formula 4

Effect size (ES or r) calculations were made using the formula:

$$r = z \sqrt{ON}$$

Formula 5

⁴ Source: <http://faculty.vassar.edu/lowry/binomialX.html>

This formula was also used to calculate z scores, given known r and N values. Use of this formula mainly follows the ‘simplicity’ precedent set by Honorton and Ferrari (1989, p. 283).

A number of studies used Spearman’s *rho* (r_s) as the effect size measure (Luke et al., 2008; Luke & Morin, 2014; Luke & Zychowicz, 2014; Roe et al., 2003, 2004, 2005; Simmonds-Moore & Moore, 2009; Thorisson et al., 1991) from which a z score could be calculated. First, the Fisher transformation is calculated, where:

$$F(r) = (0.5 * \ln[(1 + r)/(1 - r)])$$

Formula 6

If $F(r)$ is the Fisher transformation of r (the sample Spearman rank correlation coefficient), and n is the sample size, then

$$z = \text{sqrt } F(r) * (n - 3) / 1.06$$

Formula 7

The z score for r approximately follows a standard normal distribution under the null hypothesis of statistical independence ($\rho = 0$).

APPENDIX C: Random Hierarchical Effect Model

```
# Random hierarchical effect model

model.intercept <- '
data {
  int<lower=0> N;
  real g[N];
  real<lower=0> s[N];
}

parameters {
  real gamma[N];
  real mu;
  real<lower=0, upper=100> sigma;
}
model {
  mu ~ normal(0.1, 100);
  gamma ~ normal(mu, sigma);
  g ~ normal(gamma, s);
}
,
```



Established 1882

A CONTROLLED STUDY OF PSYCHOMETRY USING PSYCHIC AND NON-PSYCHIC CLAIMANTS WITH ACTUAL AND FALSE READINGS USING A MIXED-METHODS APPROACH

BY IAN S. BAKER, JANE MONTAGUE AND ABIGAIL BOOTH

ABSTRACT

This study examines whether it is possible using psychometry, the purported ability used by psychic claimants to glean personal information from objects, to “read the essence” of an individual from an object they have handled (Parra & Argibay, 2006, 2007, 2009a). This experiment was comprised of two key conditions: in the first, 45 participants handled two objects which were then passed on to a psychic claimant and a non-psychic respectively who handled the items and reported the information they gleaned from them; in the second, 45 participants also handled two items but were assigned readings that were generated for those in the first condition, thus receiving a false reading. Mixed ANOVA analysis demonstrated a significant main effect of accuracy between groups, with readings by non-psychics being rated as more accurate overall than readings by psychics (against the hypothesised direction), and no significant main effect for real versus fake readings, or a significant interaction between the main effects. There were no significant differences for behavioural influence, levels of paranormal belief between groups, or number of words in the different readings. Subsequent thematic analysis revealed three themes in the readings and considerable stylistic differences between the two groups, with non-psychics producing more examples of Barnum-style statements, which may have been rated as more accurate by participants.

INTRODUCTION

Psychological essentialism is the concept that people and objects are identified by category membership based on some inherent property or essence (Medin & Ortony, 1989). This cognitive bias is thought to develop during childhood (Gelman, 2004), and continue into adulthood influencing how individuals behave toward a particular type of person or object, specifically subjective perceptions of value and authenticity (Frazier, Gelman, Wilson & Hood, 2010; Gelman 2004; Hood, 2009). Although essentialism has connotations in social contexts for how individuals relate to each other (Haslam, Bastian, Bain & Kashima, 2006; Keller, 2005; Prentice & Miller, 2007), the focus of this study was the relationship between essentialism and objects. The notion of extreme value attributed to objects as a consequence of essentialism is apparent in the context of celebrity memorabilia; for example, at auction in 1996 a bidder paid \$772,500 for a set of John F. Kennedy’s golf clubs and \$48,875 for a tape measure from the Kennedy household (Newman,

Diesendruck & Bloom, 2011). One may argue that these items could have been purchased for their intended function, but the fact that collectors have also bought gum chewed by a celebrity for \$160 (in this case, Britney Spears: Newman, Diesendruck & Bloom, 2011) suggests there is some inherent meaning or value in such items. This is concurrent with the belief that the objects have become imbued with the essence of their owner, which transforms golf clubs and tape measure into carriers of this essence (Fernandez & Lastovicka, 2011; Newman, Diesendruck & Bloom, 2011).

Within the concept of authenticity, people feel associations with people and places, particularly after items have come into direct contact with them. This is linked to the idea of contagion; the act of ‘contaminating’ an object or place with the essence of an individual or event (Fernandez & Lastovicka, 2011; Frazier et al., 2009; Nemeroff & Rozin, 1994; Newman et al., 2011). There are positive and negative views of contagion. For example, participants in one study revealed that in choosing between an old, battered guitar played by their musical hero and a brand new one of the same make and model, every participant chose the one belonging to the musical hero (Fernandez & Lastovicka, 2011). This may be attributed to mere association and proximity (Argo, Dahl & Morales, 2006); however, some participants went on to demonstrate that they had contagion beliefs by further saying they believed that the guitar was imbued by ‘the gene prints’ of the musician, and their star power and ability would be ‘channelled’ through the guitar to whoever was playing it (Fernandez & Lastovicka, 2011). In contrast to this positive view in which individuals explicitly wanted to handle an object, Nemeroff and Rozin (1994) discovered that participants were reluctant to wear a cardigan that was claimed to have been worn by a serial killer in comparison to a new one.

Further to the belief that an object can hold an ‘essence’ of a person, there are individuals who claim to have the ability to psychically glean information about others by way of psychometry, defined as “a type of anomalous cognition (or ESP) which permits a psychic or ‘sensitive’ to receive impressions using a physical object as an inductor or instrument for information” (Parra & Argibay, 2009a, p. 57). This is the mechanism often claimed to be used by psychic detectives assisting with crimes when people, living or dead, are missing (Lyons & Truzzi, 1991), although it has been argued that there is little experimental support for such practices being successful (O’Keefe & Alison, 2000; Wiseman, West & Stemman, 1996). Nevertheless, it was anticipated in the current study that, if someone had such an ability, they would be able to glean accurate information about participants by reading their essence from an object only they have handled.

Parra and Argibay (2006, 2007, 2009a) noted that research into psychometry was limited to qualitative analysis due to problems that arose in evaluating free-response material. They also noted there has been little interest in exploring this phenomenon in ordinary, non-psychic individuals. By way of illustration, their 2006 study tested the accuracy of a psychometry-based strategy for assessing personal objects compared with a non-psychometry strategy using pictures sealed in envelopes. Four volunteers were used as targets, carrying identical objects around with them for fifteen

days, and 71 participants were recruited to obtain impressions from the objects and pictures. This already presents a problem of contamination: if so many people are handling these initial four objects it cannot be determined whether any impressions gleaned are from the original four targets. After being blind-scored by the original targets, it was concluded that participants performed more accurately in the picture condition than the psychometry condition, offering support for the claim that stimulation of visual imagery is more conducive for accuracy than the use of token objects, at least amongst a non-psychic population. There are further problems with this particular study, contamination aside; there were no psychic claimants as controls by which to measure accuracy performance and all of the participants believed in psychic ability so there was no control for scepticism. They also did not specify why it was 15 days that the target items were handled for, or mention whether time in contact with the object was a factor. Parra and Argibay's (2006, 2007, 2009a) research has acted as the inspiration for the current study, but we hoped to address several of the limitations outlined above.

The current study aim was to investigate whether essence is imbued into an object, such that information about the object's owner may be gleaned by individuals claiming to have psychic abilities and those who do not. Parra and Argibay's research has focused upon non-psychics or 'ordinary people' (2006, p. 299), or the categorisation of participants as 'psychics' or 'non-psychics' according to questionnaire responses (2007, 2009a). In each of these three studies, they had 70–80 participants handle the target objects from between two and six target persons. However, such an approach may introduce a 'contagion effect', whereby the handling of the targets by multiple people may result in the targets being associated with the 'essences' of multiple individuals. Therefore, in this study we intended to reverse the design in order to minimise this effect, with four 'readers' (two self-reported 'psychics' and two non-psychics) providing readings from the objects that the participants held, so that the only people to physically handle the objects would be the participant and the reader.

In terms of the behavioural impact of receiving a reading, many people visit alleged psychics, including celebrities, and pay money for these services which may be viewed by users as either a means of satisfying curiosity or a spiritual advisory service (Frank, 2011; Sancho, 2001). This leads to the assumption that individuals who seek out such services for the latter intend to use the information they have been given by the psychic claimant in order to enrich and improve their lives, resulting in changes to behaviour. It was therefore hypothesised that readings from psychics would be rated as more accurate than readings from non-psychics, and that readings from psychics would be rated as producing greater changes in behaviour than readings from non-psychics. As a control, we proposed to provide participants in a 'fake' reading condition with readings that were not intended for them but were originally for participants in the 'real' reading condition. It was anticipated that such readings would be rated as less accurate and would result in lower levels of influence on future behaviour for both the psychic and non-psychic groups.

METHOD

Design

The experiment employed a 2x2 mixed ANOVA design that manipulated claimant type (within subjects, psychic versus non-psychic) and reading type (between subjects, ‘real’ versus ‘fake’), as illustrated in Figure 1.

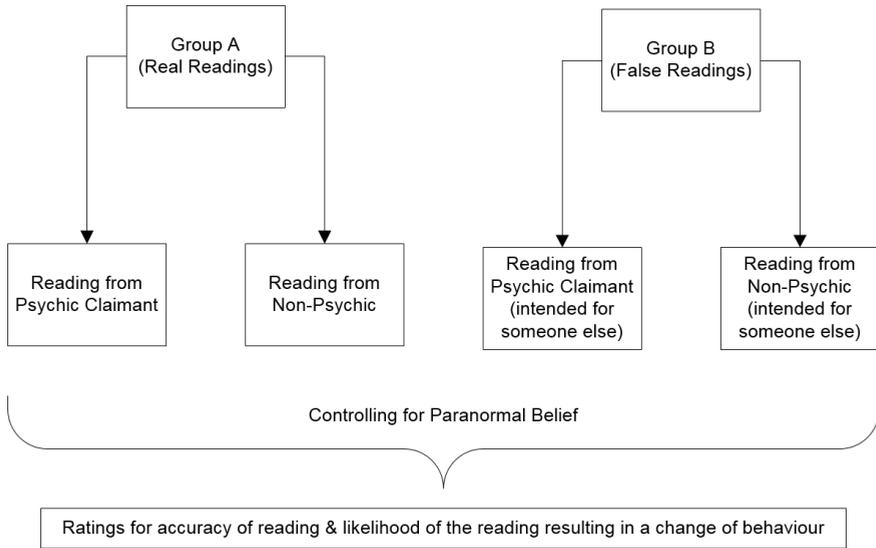


Figure 1: Summary of the 2 x 2 mixed design with accuracy and impact upon future behaviour as the dependent measures

The dependent variables consisted of two questions: The first question asked participants “how would you rate the accuracy of your reading?”, and responses were made on a 7-point Likert scale ranging from “completely inaccurate” to “completely accurate”. The second question asked participants “how likely will this reading influence your behaviour?”, and responses were made on a 5-point Likert scale ranging from “not at all” to “significantly so”.

The revised Paranormal Belief Scale (rPBS: Tobacyk, 2004) was used to check for differences in the levels of paranormal belief between the groups.

Participants

In total 90 participants were recruited to take part in the study; this was based upon a prospective power analysis ($d = .61$) from Parra and Argibay’s (2009a) work. The only exclusion criteria were that participants had to be over 18 years of age and able to give informed consent. Participants were recruited on an opportunity basis; invitations to take part were delivered via social networking sites such as Facebook, sampling undergraduates at the University of Derby, and placing posters in a New Age shop in Nottingham.

Materials

Two bags containing 100 marbles were used. These were chosen as they were easy to handle and are factory-produced items likely to have come into little or no human contact during the manufacturing process, thus minimising the possibility of contagion. The psychic claimants who volunteered to do the readings were two self-employed females who claim to regularly produce accurate psychic readings for paying customers using a mixed methodology of tarot cards and psychometry. No prior tests were conducted to ensure the validity of their claims. The non-psychic claimants were two females who were approached by the experimenter (AB) and agreed to take part when they were informed all they would have to do was handle each marble individually and note down any impressions they felt from it pertaining to the participants who had handled them. They did not claim to possess any psychic ability. No prior tests were conducted to ensure the validity of their claims.

Procedure

The procedure is summarised here and clarified further in Figure 2. Participants were provided with two marbles which were carefully dropped into their hands from the bag in to avoid contamination from the experimenter. They were asked to keep these marbles either in their hand or about their person for approximately fifteen minutes. During this time, all participants completed Tobacyk's (2004) revised paranormal belief questionnaire (rPBS). Once completed, participants were asked to place their marbles into separate envelopes and seal them.

For the 'real' reading condition, the envelopes were provided to the psychic and non-psychic claimants. The readers were instructed to open the envelopes one at a time and handle the marble inside and note down any impressions gleaned from handling the object. The experimenter was not present when the readings took place so as not to influence the reader and to remain blind to which participant the readings applied to at this stage. Once the experimenter had all envelopes and readings back from the readers, participants were sent an e-mail containing both readings (the source of which was not identified) along with a scale to rate the accuracy and behavioural influence for each.

Participants in the 'fake' reading condition were randomly assigned a reading from the psychic claimant and non-psychic that were actually for participants in the first condition, and asked to rate the accuracy and behavioural influence for each. All participants were debriefed and informed of the full details of the study once the data was received, and those participants in the 'fake' reading condition received a specific debrief explaining why they received the readings that they did.

RESULTS

An initial independent samples t-test found that there was no significant difference in the levels of paranormal belief between participants in the 'real' reading condition ($M = 3.25$, $SD = 1.24$) and those in the 'fake' reading condition ($M = 3.62$, $SD = 1.10$), $t(89) = 1.318$, $p = .192$, $d = .28$, indicating that this should not have impacted upon the main analysis.

Means and standard deviations for Accuracy and Behavioural Influence Ratings for each condition are shown in Table 1.

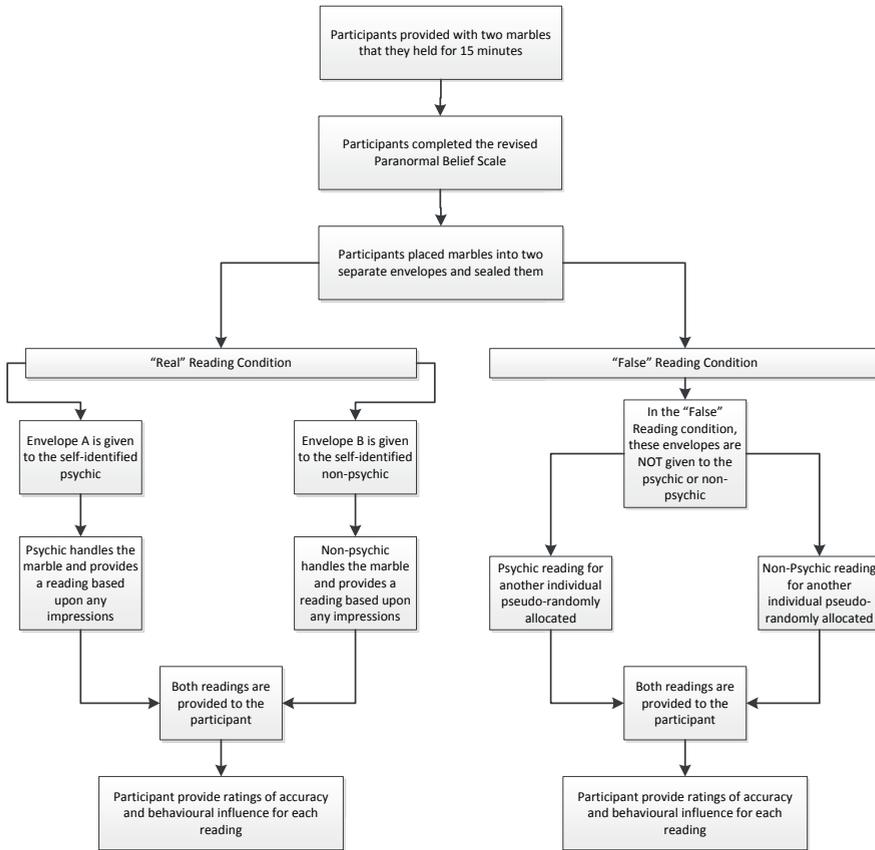


Figure 2: Flowchart summarising the experimental procedure

A 2x2 mixed ANOVA analysis revealed that there was a significant main effect of claimant type on perceived accuracy of the reading ($F(1,88) = 15.14$, $p < .001$, $\eta^2 = .07$), with non- psychic claimant readings being rated as being more accurate than psychic claimant readings. In addition, there was not a significant main effect for reading type ($F(1,88) = .15$, $p = .70$, $\eta^2 < .001$), and there was not a significant interaction effect ($F(1,88) = 2.28$, $p = .13$, $\eta^2 = .01$).

Furthermore, a 2x2 mixed ANOVA analysis revealed that there were no main effects of claimant type ($F(1,88) = 2.66$, $p = .11$, $\eta^2 = .01$) or reading type ($F(1,88) = .96$, $p = .33$, $\eta^2 = .01$) on the influence of the reading of future behaviour, nor was there a significant interaction effect ($F(1,88) = .05$, $p = .82$, $\eta^2 < .001$).

As the results indicated that the type of claimant had a significant impact upon perceived accuracy, but with non-psychic claimants being perceived as being more accurate, a post-hoc analysis examined the number of words in the readings for the two groups of claimants. This examined if non-psychic claimants may have provided more information, giving participants a higher opportunity to judge that an aspect of that information applied to them. However, a Welch's t-test revealed that there was not a significant difference between the overall number of words given in the readings for the psychic claimants ($M = 26.89$, $SD = 15.36$) and non-psychic claimants ($M = 29.67$, $SD = 11.18$), $t(80.432) = .981$, $p = .33$, $d = .22$.

Table 1.

Means (and standard deviations) for the accuracy ratings and behavioural influence ratings for all conditions

		claimant type			
		Psychic	Non-Psychic	Total	
reading type	Real Readings	Accuracy Rating	3.78 (1.78)	4.36 (1.64)	4.07 (1.73)
		Behavioural Influence	1.96 (1.21)	2.22 (1.30)	1.96 (1.20)
Fake Readings		Accuracy Rating	3.51 (1.87)	4.82 (1.40)	4.17 (1.77)
		Behavioural Influence	2.20 (1.23)	2.40 (1.20)	2.30 (1.20)
Total		Accuracy Rating	3.64 (1.82)	4.59 (1.54)	
		Behavioural Influence	2.08 (1.21)	2.31 (1.24)	

DISCUSSION OF EXPERIMENTAL FINDINGS

The experimental results suggested that overall the readings of non-psychics were rated as being significantly more accurate than the readings of psychics, regardless of whether or not the reading was intended for that particular individual. This did not support any of the hypotheses, which assumed that psychics would provide more accurate and behaviour-changing readings for intended recipients compared with the other groups, and therefore any interpretation of the experimental findings must be treated with caution. This is contrary to the findings of Parra and Argibay (2007) who found that individuals claiming ESP ability scored higher in a forced-choice psi test using a token object than those claiming only ESP experiences, and to Parra and Argibay's (2009a) work, which cautiously concluded that psychics provided more accurate psychometry readings than non-psychics. However, our findings are broadly consistent with those of O'Keeffe and Alison (2000), who reported that non-psychics performed better than psychic claimants when examining objects from crimes.

An initial interpretation of these results might conclude, then, that non-psychics can provide significantly more accurate readings than psychics. However, this interpretation fails to take account of the overall difference

across both types of reading ('real' and 'fake'); in other words, the nature of all readings provided by the psychics and the non-psychics appeared to significantly differ from one another. If the psychics were merely less accurate, the results should indicate that the reading from the psychics provided to someone for whom it was not intended (the 'fake') would be rated as no different from the readings produced by the non-psychics, i.e., there should be no differences across all conditions. This difference suggests the possibility of a fundamental difference in the approach that both groups of readers adopted when providing the readings.

One possible explanation for these differences might be that there was an inadvertent use of Barnum statements. First coined by Forer in 1949, this refers to generic statements that could be applicable to any individual who would be able to mentally mould trivial information to fit in with their lives (Furnham & Schofield, 1987; Roe, 1995; Roe & Roxburgh, 2013). It is conceivable that non-psychics, when asked to provide 'readings' of a person based upon holding a marble, may have defaulted to providing generic statements that could apply to many people and would be more likely to be rated in the middle of the accuracy scale. In contrast, the psychics may have attempted to provide much more 'targeted' information. In some cases, particularly relevant information by the psychics may have resonated and been rated as 'completely accurate'. Conversely, if such information was irrelevant, it would have been rated as being the opposite extreme of the scale. Although some of the readings may have been highly accurate, if more readings were viewed as irrelevant then this would have impacted upon the overall comparison with the more general Barnum statements produced by the non-psychics. This is supported to some degree by the larger standard deviation for accuracy ratings for the psychic versus the non-psychic readings. It may also help to explain the lack of significant differences in the impact of the readings upon future behaviour, as the readings lacked the psychological impact to provoke such change.

In order to explore these issues further, it was decided that a qualitative thematic analysis of the readings would help to elucidate the details of the narratives used.

QUALITATIVE THEMATIC ANALYSIS AND INTERPRETATION

The qualitative analysis was conducted by JM, who was blind to the conditions that each reading referred to (i.e., it was not revealed which readings were from the psychics and which were from the non-psychics). The readings were tabulated in two columns, A and B; within these, each reading (90 in total) was individually categorised in order to produce a series of overall themes relating to the content. On a first pass of the data the overall difference between the two groups did not appear to simply be that the non-psychics were providing significantly more information (as measured by number of words in a reading) than the psychics; this matched the post-hoc quantitative analysis presented above. More elements might have been provided that an individual could find relevant, but since that was not the case, another explanation was sought through a more in-depth analysis of the readings.

A noticeable aspect highlighted through this more in-depth analysis was the different styles in which the readings were written, particularly evidenced by the use (or not) of personal pronouns throughout the readings. On reviewing this particular feature, it became evident that the non-psychics wrote in the second person (39 out of 45 readings referred explicitly to the reader) and the psychics wrote much more objectively (42 out of 45 readings made reference to the reader in the third person; only one reading included both first and second person references). A small number of readings, predominately from the psychics (16 for the psychics and 3 for non-psychics), incorporated self-referential aspects such as “I see this person being...” [Case 4 - Psychic] and “I feel this is a very happy person...” [Case 16 - Psychic]. In addition, 18 out of the 45 readings from psychics were in the form of disjointed or unrelated sentences, generally relating to specific details, whereas 44 out of the 45 readings from non-psychics were generalizable, had socially desirable aspects, and were formulated as full sentences that were related to one another. For example, the following are readings from a psychic and non-psychic relating to the same individual:

Female. Don't feel I have ever met this person. (Felt slightly alien, more so than the others?). Mental about a computer game — NOT an app like bejewelled, like Zelda or WoW — Proper computer game. Guilty pleasure of TV show — something like dancing on ice or something similar — friends wouldn't think it! [Case 6 - Psychic]

You are an outgoing and fun person that likes to socialize, however you also have a sensitive side which you need to nurture. [Case 6 - Non-Psychic]

In addition to noting these visible, stylistic differences, three themes encompassing all 90 readings (i.e., from both the psychic and non-psychic groups) were identified. These were: “biographical details”, “lifestyle”, and “thoughts versus feelings”.

Biographical Details

The theme “biographical details” was further organised into two subthemes: internally-focused and externally-focused details. The internally-focused readings made reference to personal characteristics such as gender, appearance and individuals' likes and dislikes. For example:

Artistic flair — doodles or drawings especially dark ones. Poetry also? Deffo Creative. Nickname (significant). Really open and confident. Academic too. Sleeps badly, often. Anxious over what-ifs. Feel I can really relate to this energy. [Case 5 - Psychic]

A logical thinker and natural mathematician. Do not be afraid to embrace your imagination and spiritual side a bit more, travelling will open your eyes so do not pass up on the opportunity. You are different to those you were studying with and this is not a bad thing. [Case 29 - Non-Psychic]

The stylistic differences in these two examples fit with the overall patterns discussed above, particularly relating to the use of third/first person pronouns and disjointed statements/full sentences. The overall content of both, however, is similar, highlighting, for example, aspects relating to personal traits and their general approach to life. Other readings contrasted with these, referring to aspects such as place, relationships, interests and activities. For example:

Can't tell gender. Connections to Leeds/Yorkshire think maybe grew up or went to uni there. Feels like 3–5 years have no idea why — Pics of Leeds station (train) and Yorkshire cathedral. Not just visited lots. Blue-green gem on bracelet or ring is sentimental — Shown as oystershell blue. Like an opally/streaked/multi-tonal effect. [Case 2 - Psychic]

Feeling a connection with someone... can't pick up the gender. Age: Older than 30 [Only concentrated on older than/younger than]. Originally from or strong family connections to North East ENGLAND. Hull to Newcastle edge. Feel relief as if I've just saved a life. Like pulled someone out of a perilous danger. Saved an actual life. [Case 39 - Non-Psychic]

In contrast to those readings that referred in large part to the internal characteristics of an individual, these readings explicitly include reference to others and some also include aspects such as key places or geographical areas. The non-psychic example illustrates the self-referential nature of some of the readings, but appears to go beyond this and draw attention to a much more personal and emotional state of the reader. It also presents one of the few non-psychic readings that were not written with direct reference to the reader.

Lifestyle

The theme “lifestyle” focuses more upon things about the person such as their educational background, creativity, reference to colour, work-life balance, health, and their approach to life in general. For example:

Feminine. Horses/paddocks — can visualise show jumps in a field/paddock. Feels psychic too, dabbled in Wicca and can sense aged presences/stone tape (Only way I could think of to describe being able to read atmospheres). Randomly — Hates Gin. [Case 13 - Psychic]

Female. Very booky — Reference books/fiction lots of books. Goth — Lots of black/dark colours — Images of that girl from NCIS with short short fringe. Poss. pet rat or gerbil definitely small rodent pet. [Case 43 - Psychic]

It's not too late to kick start those health and lifestyle changes you've been thinking about recently, there will be positive improvements as soon as you start making them. [Case 32 - Non-Psychic]

You like to live life in the fast lane, and struggle to relax. At times though you need to take a moment to reflect and recharge your batteries. [Case 13 - Non-Psychic]

This group of readings overlaps in some part with some of the aspects seen in the above theme, i.e., referring to likes, dislikes and so on. In these readings, however, those aspects are very often explicitly linked to a particular approach to life or lifestyle choices and changes. They draw together both internal aspects, such as bookishness and health, with external aspects, such as demonstrating femininity or outward appearance, to comment more specifically on the individual's engagement with the world.

Thoughts versus Feelings

This differs from the internally-focused biographical details sub-theme in that this category of readings relates specifically to emotions and/or cognitions. For example:

A practical thinker, not one to be taken for a fool yet very compassionate. [Case 1 - Psychic]

A practical and reliable person, you are a good person to turn to regarding a problem or crisis and a calming influence on friends and peers. [Case 31 - Non-Psychic]

I feel this person is experiencing a lot of anxiety at the moment, it's OK to relax. [Case 27 - Psychic]

You are very compassionate towards others, and this would help towards a career in this area. Helping people seems to be your forte. [Case 17 - Non-Psychic]

Again, overlaps between all three themes can be seen here. The readings in this group draw a contrast between emotions such as anxiety and compassion, and rational thought such as being a 'practical' or a 'logical' thinker. There was no obvious difference in how these were used between the psychics and non-psychics and both types of reference were made by both groups, as can be seen above.

OVERALL DISCUSSION

The experimental findings indicated that the readings provided by non-psychics were rated as being significantly more accurate than the readings provided by psychics, regardless of whether or not the reading was provided for a particular individual. This suggested that non-psychics generally provide readings that individuals find to be more accurate to them than psychics do, although this finding should be treated cautiously as it is opposite to the hypothesised direction. However, as interesting as this finding was, the experimental findings on their own did not provide a complete picture, as they did not reveal in what way these readings differed.

The use of qualitative thematic analysis of both sets of readings by a researcher who was blind to the source of each reading produced a richer understanding of the nature of the different readings. The qualitative analysis provided several key findings. Firstly, there appeared to be differing styles between the readings; non-psychics tended to write in the second person, whereas psychics wrote predominantly in the third person. This may have made the non-psychic readings easier for different people to relate to, thus making the accuracy ratings higher. Secondly, self-referential features were more common amongst the psychics compared to the non-psychics. This can provide a more personalised experience for the individual receiving the reading. As such, if the reading is regarded as being accurate, it might feel more relevant to that individual and therefore be rated higher in accuracy, but this would be much less relevant if it was deemed inaccurate. This may help to account for the slightly broader spread of accuracy scores for the psychic readings compared to the non-psychic readings overall. In addition, the use of disjointed or unrelated sentences was much more common amongst the psychic readings compared with the non-psychic readings, and this can make the readings more difficult to read, which may have impacted their perceived accuracy. Finally, the most important stylistic difference between the readings was that the non-psychic readings were almost universally generalizable in nature and commonly contained socially-desirable aspects.

The qualitative comments were organised within three different themes: "biographical details", "lifestyle", and "thoughts versus feelings". Biographical details, including both internal and external aspects, were noticeable in 19 of the non-psychic readings and 23 of those from psychics. Lifestyle comments

occurred in 21 of the non-psychic readings and 13 of the psychic readings. Thoughts versus feelings were seen in readings from 6 of the non-psychics and 8 of the psychics.

Overall, the quantitative and qualitative findings suggest that the psychics approach a task such as this very differently from the non-psychics. The qualitative findings indicate that the psychics are less likely to generate Barnum-style statements; a criticism that has been levelled at psychics previously and criticised in turn (Roe, 1995; Roe & Roxburgh, 2013), and ironically in this study has resulted in the readings from the psychics as being rated as less accurate than the non-psychic readings.

This would also help to explain why none of the conditions would lead to a significantly different degree of change of future behaviours than any of the others. Although previous research has argued that individuals can make life decisions based upon what a psychic has told them (Roe, 1996), this would be unlikely to happen in this study due to the psychic readings being rated as less accurate than the non-psychic readings. This would also imply that, although the readings from the non-psychics were rated as being more accurate, they would not lead to any significantly higher levels of behavioural change due to them having minimal psychological impact; which is consistent with generic, Barnum-style statements.

The study does have a number of limitations. It only utilised two psychics and two non-psychics and the findings presented here might be idiosyncratic to them; only additional replications will uncover this further. In addition, unlike other studies (e.g., Beischel & Schwartz, 2007) there was no prior screening of any abilities (or lack thereof) of the four individuals and it is possible that the psychics may not have had any mediumistic abilities at all. However, although this was somewhat controlled for by the use of 'fake' readings, as if any of the four consistently produced highly accurate readings tailored to particular individuals, they would have also been consistently rated as less accurate for the fake readings, as the readings would not have applied to that individual. Finally, the use of self-identified psychics is arguably no different to previous research into psychometry that categorised 'psychics' as those who answered "yes" to extrasensory abilities on a self-report questionnaire (Parra & Argibay, 2009b). With regard to the method of psychometry itself, the marbles provided a well-controlled object for individuals to hold and, unlike Parra and Argibay's (2006, 2007, 2009a) research, where multiple participants may have handled the objects, only the participant and reader held them in order to minimise potential contamination. However, they are not a particularly noteworthy object to imbue an 'essence' to; many of the items referred to in the introduction were unique or had significance to the person, and marbles are unlikely to be meaningful either psychologically or psychically. In addition, participants only held the object for 15 minutes, which may or may not have been enough time to imbue the marble with their 'essence'. There are no clear indications as to how long an appropriate duration to imbue and essence might be; it is possible that 15 minutes is simply not long enough time. Future research, therefore, should aim to manipulate this duration to assess the effect upon potential readings, particularly as the manipulation of duration has not been previously

conducted. Finally, although the psychics were happy to participate, the well-controlled nature of the study involved psychometry using methods that they would not typically use, particularly regarding the controls and isolation from the participants, and this may have influenced their readings. A more ecologically valid approach would have been ideal, but is very challenging to achieve due to the experimental controls that are necessary.

Many of these limitations could be addressed in future research. Multiple psychics and non-psychics could be used, potentially with the introduction of screening methods. It is difficult to find objects that are both individualistic and mass-produced, in order for them to be both a significant object to the individual but also not to have characteristics that the reader could infer information about. However, these need not be mutually exclusive. One option might be a mass-produced object (e.g., ping-pong balls) that are then personalised by the individual and kept with them for a longer period of time so that some association with the object is engendered (e.g., a week). Appropriate controls would need to be introduced to the readers to prevent sensory contamination (e.g., blindfolds). The degree of attachment to the object could also be measured to see if this in itself correlates with accuracy or the likelihood of the change of behaviour. It would be ideal if the object had high emotional valence for the individual, but the challenge is to ensure that such an object does not have any unique identifying features that could lead to sensory contamination. Future research should also seek to incorporate a mixed-methods approach in a more purposeful manner; although the two approaches provided a more complete picture in this study, the use of qualitative approaches was only taken when the experimental study found results contrary to the hypothesised direction. Taking conversation analysis insights (e.g., Robinson, Tate & Heritage, 2015; Heritage & Robinson, 2006), one approach might be to take readings and manipulate them; for example, either as written in the first or the third person to explore any impact this might have on how individuals relate to them.

In conclusion, the mixed-methods approach that was ultimately adopted here has revealed interesting findings. The study acknowledges Parra and Argibay's (2006, 2007, 2009a) criticisms of previous research by using a quantitative approach and exploring psychometry by using non-psychics. However, the quantitative approach alone would have merely indicated that non-psychics are rated as providing more accurate readings than psychics, but without exploring how or why. Similarly, a qualitative approach alone would have suggested differences in the narratives of the readings, but without necessarily demonstrating the impact that this has. The joining of the methodological approaches provides a richer and more in-depth understanding of the findings of the study; that the non-psychic readings were consistently rated as being more accurate, but this is primarily due to stylistic differences with the non-psychic readings being presented more clearly, in the second person, and universally generalizable and socially-desirable in nature.

*Department of Life Sciences
College of Life & Natural Sciences
University of Derby, UK*

I.S.Baker@derby.ac.uk

REFERENCES

- Argo, J., Dahl, D.W., & Morales, A.D. (2006). Consumer contamination: How customers react to products touched by others. *Journal of Marketing*, 70(2), 81–94.
- Beischel, J. & Schwartz, G.E. (2007). Anomalous information reception by research mediums demonstrated using a novel triple-blind protocol. *Explore*, 3, 23–27.
- Fernandez, K.V., & Lastovicka, J.L. (2011). Making magic: Fetishes in contemporary consumption. *Journal of Consumer Research*, 38(2), 278–299.
- Frank, N. (2011). *Psychic addicts numerous and largely unacknowledged*. Psychology: Suite 101 (from the world wide web <http://natalie-frank.suite101.com/psychic-addicts-numerous-and-largely-unacknowledged-a393885>) accessed 12.03.2012.
- Frazier, B., Gelman, S.A., Wilson, A., & Hood, B. (2010). Picasso paintings, moon rocks and hand-written Beatles lyrics: Adults' evaluations of authentic objects. *Journal of Cognition and Culture*, 1(9), 1–14.
- Furnham, A., & Schofield, S. (1987). Accepting personality test feedback: A review of the Barnum Effect. *Current Psychological Research and Reviews*, 6(2), 162–178.
- Gelman, S.A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, 8(9), 404–409.
- Haslam, N., Bastian, B., Bain, P., & Kashima, Y. (2006). Psychological essentialism, implicit theories, and intergroup relations. *Group Processes Intergroup Relations*, 9(1), 63–76.
- Heritage, J., & Robinson, J.D. (2006). The structure of patients' presenting concerns: Physicians' opening questions. *Health Communication*, 19, 89–102.
- Hood, B. (2009). *Supersense: From superstition to religion. The brain science of belief*. London: Constable.
- Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology*, 88, 686–702.
- Lyons, A., & Truzzi, M. (1991). *The blue sense: Psychic detectives and crime*. New York: Mysterious Press.
- Medin, D.L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). New York: Cambridge University Press.
- Nemeroff, C., & Rozin, P. (1994). The contagion concept in adult thinking in the United States: Transmission of germs and interpersonal influence. *Ethos*, 22, 158–186.
- Newman, G.G., Diesendruck, G., & Bloom, P. (2011). Celebrity contagion and the value of objects. *Journal of Consumer Research*, 38(2), 215–228.
- O'Keeffe, C., & Alison, L. (2000). Rhetoric in 'psychic detection'. *Journal of the Society for Psychical Research*, 64(858), 26–38.
- Parra, A., & Argibay, J.C. (2006). Comparing a free-response test using object and without object conditions: Exploring the 'token object' effect on an un-gifted sample. *Proceedings of presented papers: The Parapsychological Association Convention*, 299–307.
- Parra, A., & Argibay, J.C. (2007). Comparing psychics and non-psychics through a 'token object' forced-choice ESP test. *Journal of the Society for Psychical Research*, 71(887), 80–90.
- Parra, A., & Argibay, J.C. (2009a). Psychics vs. non-psychics in 'face-to-face' and 'remote' token-object reading conditions. *Australian Journal of Parapsychology*, 9(1), 57–69.
- Parra, A., & Argibay, J.C. (2009b). An experimental study with ordinary people for testing 'sacred' objects through Psi detection. *Journal of the Society for Psychical Research*, 73(894), 41–49.
- Prentice, D.A., & Miller, D.T. (2007). Psychological essentialism of human categories. *Current Directions in Psychological Science*, 16, 202–206.
- Roe, C. (1995). Pseudopsychics and the Barnum Effect. *European Journal of Parapsychology*, 11, 76–91.

- Roe, C. (1996). Clients' influence in the selection of elements of a psychic reading, *Journal of Parapsychology*, 60, 43–70.
- Roe, C.A. & Roxburgh, E.C. (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: Evidence and beliefs*. California: Praeger. pp. 177–203.
- Robinson, J.D., Tate, A., & Heritage, J. (2015). Agenda-setting revisited: When and how do primary-care physicians solicit patients' additional concerns? *Patient Education and Counselling*, 99(5), 718–723.
- Sancho, J. (2001). *Beyond Entertainment? Research into the acceptability of alternative beliefs, psychic and occult phenomena on television*, Ofcom.org (from the world wide web <http://www2.ofcom.org.uk/static/archive/bsc/pdfs/research/occult.pdf>) accessed 23.06.16.
- Sherman, M. (2002). *Why people believe weird things: Pseudoscience, superstition and other confusions of our time*. New York: Henry Hoult.
- Tobacyk, J. (2004). A revised paranormal belief scale. *The International Journal of Transpersonal Studies*, 23(3), 94–98.
- Wiseman, R., West, D., & Stemman, R. (1996). An experimental test of psychic detection. *Journal of the Society for Psychical Research*, 61(842), 34–45.
- Wooffitt, R. (2006). *The language of mediums and psychics: The social organization of everyday miracles*. Hampshire: Ashgate Publishing Limited.

BOOK REVIEWS

GHOSTS: A HAUNTED HISTORY by Lisa Morton, Reaktion Books, London, 2015, 208pp, £16.00 (hardback) ISBN 978-1-780235172

What is the job of a reviewer? If it is to say if a book is readable, entertaining and of merit then my job is simple, for this book is all three. If I were writing a general review for a newspaper, that would suffice, and I could go on to praise the book for its many strengths. As a reviewer for this *Journal*, however, I feel I should comment on whether this book might be of value to those whose interest in such matters runs beyond that likely to be satisfied by a broad and popular introduction, and here I am ambivalent.

In fact this is a book that never really comes to terms with its subject. The ghost is always present, but what Morton writes about is our reaction to it — literary, cultural, folkloric, religious, cinematic, even briefly in terms of video games. Yes, there are cases discussed in passing, but from the introduction onwards the book deals with what ghosts mean to us, and I suspect some ambiguity on Morton's part as to whether they actually exist in an objective sense at all.

The book passes with great speed over the history of ghosts, early ghost literature and the Classical World (drawing extensively on Daniel Ogden's works) and dedicates intriguing chapters to ghosts in the East (India, China, Japan, Korea) and South America. These chapters are certainly of interest, and I learned a great deal from them, but they lack the depth of *Beyond the Threshold* (Moreman 2012) or other academic reviews of various cultures' after life beliefs.

However, the next chapter shows why I am wary of any superficial and general overview; it is entitled 'The Quest for Evidence: The Ghost and Science' yet it managed to go without once mentioning the SPR, or names as seminal to modern apparitional research as G. N. M. Tyrrell, Celia Green, Charles McCreery, Hilary Evans, or even the Sidgwicks and *The Census of Hallucinations*. There is a short, critical piece on Michael Persinger's 'Godhelmet' experiments, and references to James Houran, Rense Lange and Richard Wiseman on suggestibility, but pretty much nothing on any theoretical or experimental work on apparitions beyond that. Instead, most of the chapter is dedicated to pieces of equipment used by amateur ghosthunting groups, who might be regarded as considerably further from the purported subject of this chapter than the parapsychologists who are ignored. There is a short and welcome reference to Nandor Fodor, but overall there is pretty much nothing in the chapter about real research into apparitional experience.

I turned at this point to the bibliography — which is rather short — and the more copious and eclectic notes, and found that barring Nandor Fodor (1968) there is no sign of familiarity with the work of the SPR or indeed the ASPR, ASSAP, the Ghost Club or most of the authorities in the field. I sense this is because Morton does not wish to actually tackle ghosts — this is, as I

say, about our reaction to them. But while reading about items like the Ovilus and Frank's Box proved amusing, the chapter is badly misnamed. We do find listed a book by Harry Price on Borley Rectory, and Jay Anson's on the so-called Amityville Horror; so popularity rather than evidential importance appears to be the determining factor.

The next chapter deals with the cinematic and media representations of ghosts. I know little on the subject so found it informative and interesting; I suspect, however, if I were an expert on the subject I would find it disappointing. And here is the problem — Morton's book is a book for the public, and a very good one, but it makes no serious attempt to engage with the heart of the supposed subject, as opposed to, say, Clarke (2012) or Roach (2008), popular books that make a worthwhile contribution in themselves. At times I felt the coverage of ghosts in this book was similar to what this reviewer might produce if asked to write a book on football: I may know something of football, and I may well write a good history of football fans, and football violence, and films about football, and historical ball games, but the fact I have hardly ever kicked a ball or even watched a match might show through.

So this is a well written, amusing and enjoyable book. At times it is erudite — it cites a passage from the *North China Herald* on the Hornby case that I have not seen before from a book by Weatherly and Maskelyne that I did not know. Little gems like this crop up from time to time, while slight irritations like the obvious bias towards the USA culturally irritate; for example, one would infer from Morton that *Ghost Hunters* sparked the 'reality ghosthunting' media boom of the early 21st century, but in fact MTV's *Fear* and *Scariest Places on Earth* were probably the first such mainstream shows and it was the UK's *Most Haunted* that first reached mass audiences and inspired much that followed. Morton rather implies the reverse, which may simply be an editing issue.

There is much to like about this book, but it is probably not for the readers of this *Journal*. What it does it does well, but it is a history with strange omissions, not least almost all the serious work ever done on the subject. Pick it up in paperback?

3 Baker Street
Cheltenham
GL51 9HG
chrisjensenromer@hotmail.com

CHRISTIAN JENSEN ROMER

REFERENCES

- Clarke, R (2012). *A natural history of ghosts: 500 years of hunting for proof*. London: Penguin.
- Fodor, N., & Carrington, H. (1968). *Haunted people: The story of the poltergeist down the centuries*. New York: Dutton.
- Moreman, C. (2008). *Beyond the threshold: Aterlife beliefs and experiences in world religions*. Lanham, MA: Rowman & Littlefield.
- Roach, M. (2008). *Six feet over: Adventures in the afterlife*. London: Canongate Books.

PARAPSYCHOLOGY: A BEGINNER'S GUIDE by Caroline Watt. Oneworld Publications, London, 2016. xiv + 226 pp. £ 9.99. ISBN 978 1 78074 887 0

This book aims to introduce the general lay reader to the field of parapsychology in a manner that is concise, non-technical, and fairly entertaining. In that respect, it is rather reminiscent of the approach taken by Richard Broughton (1991) in his book *Parapsychology: The Controversial Science*, which has served as one of the first (and arguably, one of the best) introductions to the field that the general public has had over the past 25 years. Could it be considered a worthy and up-to-date follow-up to Broughton's book?

As many readers of this *Journal* are likely to be aware, the author is more than qualified to write an introductory book of this type: Caroline Watt has devoted a good part of her academic career to parapsychology, having been one of the founding members of the Koestler Parapsychology Unit at the University of Edinburgh, where she studied and worked with the influential Robert L. Morris, and where she continues to be a member of the Psychology Department (from August 2016 as the second holder of the Koestler Chair of Parapsychology, after Morris). She served as President of the Parapsychological Association from 2004 to 2005, and held the Perrott-Warrick Senior Research Fellowship from 2010 to 2014. Most notably in relation to this book, she has taught an introductory course on parapsychology through Edinburgh's online teaching portal since 2008, and had previously co-authored the fifth edition of *An Introduction to Parapsychology* with Harvey Irwin (Irwin & Watt, 2007).

Oftentimes, introductory presentations on parapsychology will open with discussions of very basic phenomena (e.g., telepathy) and progressively work up to more complex ones (e.g., macro-PK, survival). Watt actually does the opposite in this book, opening the first main section with a chapter on macro-PK. Most apparent in the discussion is a focus on high-profile demonstrations of alleged macro-PK that have appeared in the media, as well as a continual emphasis on trickery, misperception, and deception as the probable explanations. While this approach offers a useful message to the reader about how one should rightfully be cautious in approaching claims of macro-PK, it does not seem particularly balanced when one realizes that the chapter largely neglects to address any of the more careful studies of macro-PK that parapsychologists have conducted over the years, such as the semi-controlled tests of RSPK conducted in a Miami warehouse by Roll and Pratt (1971), and the laboratory-based PK studies of Felicia Parise conducted by Graham and Anita Watkins (1974). While they may not be considered too highly evidential by some, studies like these suggest that there may be more to macro-PK than simple trickery and misperception in certain instances. Astute readers might recognise that this three-faceted approach of focusing on high-profile cases, emphasising trickery, and drawing little attention to more rigorous studies with notable results is very similar to the one often taken by Watt's partner and collaborator Richard Wiseman (2011), who is well-known for his avowedly sceptical views on psi phenomena.

The same general approach is extended to the next chapter, in which Watt discusses psychic detection, psychic readings, remote viewing, and reports of psi in animals. High-profile examples of public demonstrations are again

placed mostly at the forefront, with more wisely stated cautions about cold reading, sensory cues, subjective biases, and trickery. But again, discussion doesn't seem too balanced for certain example cases — one of them being the return anticipation studies with the dog Jaytee by Sheldrake and Smart (1998), and by Wiseman et al. (1998). Although Watt points out that these two research teams differed in their approaches and in how they interpreted each other's data (pp. 51–52), no details are given on how their study designs actually differed, and no mention is made of how Wiseman et al.'s (1998) results appeared to be consistent with Sheldrake and Smart's when evaluated using the latter's original analytical approach (Sheldrake, 1999). While this may not definitively resolve the issue of return anticipation effects, it obscures the possibility that there may have been something to Sheldrake's findings that was worth studying further. More to the point, it leaves the discussion of this issue rather incomplete.

Similarly, the coverage of mediumship and reincarnation in the following chapter well highlights certain methodological limitations to research on these topics, but when certain important cases are mentioned, their discussion is seemingly put off to the wayside. For instance, one of the notable mental mediums that Watt briefly mentions by name is Eileen Garrett (p. 64). Yet nothing else is said about the various controlled studies that were conducted with her (e.g., Birge & Rhine, 1942), some of which were notably significant. And although Watt does mention in passing reincarnation-type cases involving birthmarks (p. 73), not much in-depth consideration is given to them. But to her credit, the overview that Watt gives in this chapter on such things as subjective evaluation of chance correspondence in mediumship readings is quite valuable for general readers to take careful heed of.

The second main section of Watt's book begins with two chapters respectively on out-of-body experiences and near-death experiences. Primary attention is devoted to neurological and psychological theories in each instance, which is consistent with Watt's recent viewpoint arguing that these phenomena can mostly be reduced to conventional factors, and that there is nothing inherently paranormal about them (Mobbs & Watt, 2011). One important distinction that is not made, however, is the difference between spontaneous OBEs and the 'OBE-like' experiences that can be artificially induced through brain stimulation or perceptual illusions. Arguably, this distinction is important because spontaneous OBEs sometimes have a veridical component that is often lacking in artificial OBE-like experiences. Watt does briefly touch on veridical OBEs with a discussion of Tart's (1968) well-known study with Miss Z., although nothing is ever said about the studies with more experienced OBE participants such as Ingo Swann, who produced statistically significant results on more than one occasion (e.g., Harary & Solfvin, 1977; Mitchell, 1973). Watt is right about neurological and psychological theories having relevance to OBEs in general, but whether they will be able to adequately explain all such experiences — especially if additional veridical cases arise in further controlled research — remains to be seen.

A similar argument could be made for NDEs: Watt does briefly mention anecdotal accounts of NDEs which seemingly involved veridical perception,

though their significance is mainly downplayed through her observation that the relatively few studies conducted so far to test this aspect of NDEs have produced little of note. But one thing Watt does recognise — and which probably should have been emphasized more to general readers — is that this line of research is still mostly in its infancy, largely owing to the rarity of NDE occurrence (p. 107). A larger database is needed in order to make a fairer assessment of whether or not some NDEs could indeed be veridical.

These chapters on OBEs and NDEs are then followed by a chapter devoted to apparitions and hauntings, with the focus mainly being on possible psychological and geophysical explanations for these phenomena. Veridical cases are briefly discussed at the beginning of the chapter, although the discussion again seems somewhat muted. For instance, although Watt discusses the early field investigations pioneered by Schmeidler (1966) to test the impressions of mediums at reputedly haunted locations using statistical methods, she makes no mention of the meta-analytic results subsequently obtained by Maher (1999) across five separate haunting investigations and conducted using these same methods. Particularly notable in Maher's results is the finding that mediums were able to correctly pinpoint the specific areas where witnesses had reportedly seen ghosts or experienced other haunting phenomena to an overall significant degree. In contrast, control groups (comprised of sceptical-minded individuals who tried to pinpoint these areas through environmental sensory cues and logical inference) were largely unable to do so. It is curious that no mention of this finding is made in the chapter, as it goes some way toward addressing the issue mentioned by Watt that "the similarity between the mediums' and eyewitnesses' reports may have been due to both groups responding to the same [sensory] cues" (pp. 190–191). Had sensory cues been involved, one would expect the control groups to have shown significant overall correspondence with the eyewitnesses, as well.

The following chapter on the psychology of psychic experiences is perhaps the best one of the book in terms of detail, mainly because it is based on recent studies Watt has done on the possible psychological mechanisms which may influence people's perceptions of having had precognitive dreams (e.g., Valášek & Watt, 2015). The coverage given here of such mechanisms as dream intrusion, selective memory, and misjudging the probability of event outcomes provides useful insight for general readers into some of those conventional psychological phenomena identified by Morris (1986) as being "not psi, but looks like it".

The last main section of Watt's book looks at experimental research, beginning with a chapter on laboratory studies of telepathy and clairvoyance. Good descriptive summaries are given of the procedures used in the ESP card tests by Rhine and his colleagues, the Maimonides dream experiments, and the ganzfeld studies. But when results are discussed, a subtle change occurs: whenever the results of a given experimental database seem to have only modest significance and/or are somewhat inconclusive, details such as hit rates or effect sizes are given. But when the results are highly significant and seemingly less ambiguous, these details tend to be curiously omitted. This is most apparent in the discussion of the ganzfeld results, in which Watt cites the recent meta-analysis by Storm et al. (2010). No mention of the overall hit

rate or its associated significance level is given for this meta-analysis, nor does Watt make any mention of the even more recent meta-analyses by Rouder et al. (2013) and Storm et al. (2013), both of which found significant results for the ganzfeld database (though these two research groups disagreed over certain methodological details in interpreting the results). This seemingly follows in line with a trend of selective and incomplete coverage seen in previous chapters.

A little more balanced than the others are the two chapters that follow on precognition and psychokinesis. Early forced-choice tests and presentiment studies are covered quite well, and although the recent series of ‘feeling the future’ experiments by Daryl Bem are discussed, no mention is made of the latest meta-analysis of all experiments of this type that have been conducted to date (Bem et al., 2016), which has been openly available on the Social Science Research Network online database since at least 2014. Micro-PK tests with RNGs and DMILS studies are described in pretty good detail, although only passing mention is made of the PEAR programme with no real in-depth coverage of its methodology and results, and no discussion is devoted at all to field RNG studies and the Global Consciousness Project.

The last chapter of Watt’s book concisely summarises the various lines of research covered in previous chapters and considers the value that this research has had. Watt highlights the important points she made in her Presidential Address at the 2005 Parapsychological Association Convention (Watt, 2005), noting that several of the research techniques which are now commonplace in psychological studies (e.g., randomisation, blinding, statistical analysis) were first put to use in parapsychology. She also emphasizes the value of pre-registering studies as a way to minimize experimenter bias, and to encourage this practice, she and James Kennedy have recently established a study registry on the Koestler Parapsychology Unit website — an effort for which they deserve to be applauded.

Although much attention has been drawn here to the selective coverage of research in Watt’s book, this should not keep it from being used as a resource in introductory courses on parapsychology, as one must also realize that not all useful books are perfect in every way. Watt has written the text in a way that is easy to read and she conveys many key concepts admirably well, making them very accessible to general understanding for lay readers. In that respect, her book should be considered (and recommended) as a good supplemental companion to Broughton’s (1991) own introductory book, best being read alongside or subsequent to his in order to help give the reader a more fairly balanced overview of the field as it stands today.

Psychical Research Foundation

Albuquerque, NM, USA

bryanwilliams@psychicalresearchfoundation.com

BRYAN J. WILLIAMS

REFERENCES

- Bem, D.J., 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*, 4, 1188. doi: 10.12688/f1000research.7177.2

- Birge, W.R., & Rhine, J.B. (1942). Unusual types of persons tested for ESP: I. A professional medium. *Journal of Parapsychology*, 6, 85–94.
- Broughton, R.S. (1991). *Parapsychology: The controversial science*. New York: Ballantine Books.
- Harary, S.B., & Solfvin, G. (1977). A study of out-of-body experiences using auditory targets. In J.D. Morris, W.G. Roll, & R.L. Morris (Eds.) *Research in Parapsychology 1976* (pp. 57–59). Metuchen, NJ: Scarecrow Press.
- Irwin, H.J., & Watt, C.A. (2007). *An introduction to parapsychology* (5th Ed.). Jefferson, NC: McFarland & Company, Inc.
- Maher, M.C. (1999). Riding the waves in search of the particles: A modern study of ghosts and apparitions. *Journal of Parapsychology*, 63, 47–80.
- Mitchell, J. (1973, April). Out of the body vision. *Psychic*, pp. 43–47.
- Mobbs, D., & Watt, C. (2011). There is nothing paranormal about near-death experiences: How neuroscience can explain seeing bright lights, meeting the dead, or being convinced you are one of them. *Trends in Cognitive Sciences*, 15, 447–449.
- Morris, R.L. (1986). What psi is not: The necessity for experiments. In H.L. Edge, R.L. Morris, J. Palmer, & J.H. Rush (Eds.) *Foundations of parapsychology: Exploring the boundaries of human capability* (pp. 70–110). London: Routledge & Kegan Paul.
- Roll, W.G., & Pratt, J.G. (1971). The Miami disturbances. *Journal of the American Society for Psychical Research*, 65, 409–454.
- Rouder, J.N., Morey, R.D., & Province, J.M. (2013). A Bayes factor meta-analysis of recent extrasensory perception experiments: Comment on Storm, Tressoldi, & Di Risio (2010). *Psychological Bulletin*, 139, 241–247.
- Schmeidler, G.R. (1966). Quantitative investigation of a 'haunted house'. *Journal of the American Society for Psychical Research*, 60, 137–149.
- Sheldrake, R. (1999). Commentary on a paper by Wiseman, Smith and Milton on the 'psychic pet' phenomenon. *Journal of the Society for Psychical Research*, 63, 306–311.
- Sheldrake, R., & Smart, P. (1998). A dog that seems to know when his owner is returning: Preliminary investigations. *Journal of the Society for Psychical Research*, 62, 220–232.
- 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.
- Storm, L., Tressoldi, P.E., & Utts, J. (2013). Testing the Storm et al. (2010) meta-analysis using Bayesian and frequentist approaches: Reply to Rouder et al. (2013). *Psychological Bulletin*, 139, 248–254.
- Tart, C.T. (1968). A psychophysiological study of out-of-the-body experiences in a selected subject. *Journal of the American Society for Psychical Research*, 62, 3–27.
- Valášek, M., & Watt, C. (2015). Individual differences in prophetic dream belief and experience: Exploring demographic and sleep-related correlates. *Personality and Individual Differences*, 87, 65–69.
- Watkins, G.K., & Watkins, A.M. (1974). Apparent psychokinesis on static objects by a 'gifted' subject. In W.G. Roll, R.L. Morris, & J.D. Morris (Eds.) *Research in Parapsychology 1973* (pp. 132–134). Metuchen, NJ: Scarecrow Press.
- Watt, C. (2005). Parapsychology's contribution to psychology: A view from the front line. *Journal of Parapsychology*, 69, 215–232.
- Wiseman, R. (2011). *Paranormality: Why we see what isn't there*. London: Macmillan.
- Wiseman, R., Smith, M., & Milton, J. (1998). Can animals detect when their owners are returning home? An experimental test of the 'psychic pet' phenomenon. *British Journal of Psychology*, 89, 453–462.

PSYCHOLOGY AND THE NEAR DEATH EXPERIENCE: A SEARCH FOR GOD by Roy L. Hill.
White Crow Books, Hove, 2015. 312 pp. £12.99. ISBN 978-1910121429

Balancing scholarship while discussing exceptional experiences in a book meant for a 'lay audience' can be difficult. Roy Hill attempts this balance with *Psychology and the Near Death Experience: A Search for God*. Hill introduces many examples of near-death experiences (NDEs) that have been abstracted from larger personal narratives published on the Near Death Experience Research Foundation and After Death Communication Research Foundation websites. Intermingled with these testimonies are references to different psychological theories (such as psychoanalytic/Jungian), scientific paradigms (such as materialism), philosophy (such as existentialism), neuroscience and religious and spiritual underpinnings. Existentialism and materialism are discussed briefly and pushed to the side to shine the spotlight on Jungian theory and Hill's personal theologically-oriented perspective on NDEs.

Hill states that his book is a "spiritual work" (p. xxii) but discusses the topic of NDEs from a "psychologist" view point (p. xxii). He also states that he "writes objectively" (p. xiii). By the end of Chapter 1, however, the precarious balance of scholarship and objectivity while discussing NDEs has been compromised in several ways. The first crack that I noticed in the foundation is Hill's statement, "If this were a pure psychology text, the content would be limited to phenomena that are definable and measurable" (p. xxii). This indicated to me that there may be no reference to phenomenological perspectives of psychology, which turned out to be a valid assumption. Also absent were references one should expect in a literature review of psychological research on NDEs (such as works published by Bruce Greyson, Susan Blackmore, Craig Murray). The main psychological perspective offered is Jungian, largely taken out of context.

This book could have provided more psychological perspectives while maintaining its accessibility to the general public. It could have retained its personal narrative components, while thoroughly exploring the psychology of them. It could have easily resonated on a personal/spiritual level while enhancing scholarship of NDEs. Instead, Hill unfortunately uses a 'buckshot' approach to frame the NDE testimonies, with faulty psychological application and biased opinion. The promise for a balance of psychology, objectivity and the "search for God" is not maintained.

Hill tries to use Jungian psychology to correlate with and expand the NDE testimonies provided. However, Hill's usage of Jungian references is shallow and typically out of context. For instance, on p. 24, Hill states, "Not only does Jung suggest that we make up the Whole of God, but that individuality derives from the Whole". He also states that Jung "believed that every soul exists as an integrated element of the Whole". His statement is based on Jung's quote (also on p. 24), "The collective unconscious, and we describe it as objective because it is identical in all individuals and is therefore one. But out of the universal one, there is produced in every individual a subjective consciousness". Jung is talking about the psyche and not the soul, in this quote. This difference is explained by Jung in his compendium *Psychological Types*:

I have been compelled, in my investigations into the structure of the unconscious, to make a conceptual distinction between soul and psyche. By psyche, I understand the totality of all psychic processes, conscious as well as unconscious. By soul, on the other hand, I understand a clearly demarcated functional complex that can best be described as a "personality". (Jung, 1971, def. 48 par. 797)

Furthermore, When Jung uses the term "God" in his psychoanalytic orientation, he is describing an archetype or psychic energy, such as when discussing God-consciousness. Hill does not use the term God in this way but in an Abrahamic-oriented way. Regarding God, Jung states in *Letters*, "When I say 'God' I mean an anthropomorphic (archetypal) God-image and do not imagine that I have said anything about God" (Jung, 1952, p. 52).

Hill's Christian preferences are obvious in this book, which tarnish his claim of objectivity when discussing spirituality and NDEs. He makes many assertions about God, such as on p. 234 with his statement, "God does not want each person to develop an individualized moral compass". This is but one example of many. His term of choice, and sole reference to a "higher power", "divinity", "universal energy", etc., is "God". There are no references to NDEs and other religions or spiritualities. The preface by Jody Long is also very exclusive: "Many people are completely transformed when they meet God or Jesus. Instead of being a two-dimensional being talked about in Church, the NDE makes God, Jesus, and love more real than we would ever experience on earth" (p. xiv).

Not everyone who identifies as spiritual or as having had an NDE also identifies the higher power they experienced (if they experienced one at all) as God. Just as Jesus is not the only religious figure that NDErs have reported. This is not only important regarding the lack of objective scholarship, but it is important because it is disrespectful to NDErs who are reading this book and feel they have been promised an objective stance on their experiences and spirituality. When one writes a book and claims a psychological, objective perspective then that is what should be delivered. Hill does not present an accurate, objective, nor psychological perspective; his title should not be *Psychology and the Near Death Experience: A Search for God* but something like *The Near Death Experience: One Perspective on a Search for God*.

Psychology Division
University of Northampton
Park Campus, Northampton NN2 7AL
eapratte@gmail.com

ERIKA A. PRATTE

REFERENCES

- Jung, C.G. (1952). *Letters, vol. 2: 1951–1961*, Princeton, Princeton University Press.
Jung, C.G. (1971). *Psychological types, collected works of C.G. Jung*, Volume 6, Princeton, Princeton University Press.

THE *SELF* DOES NOT DIE: VERIFIED PARANORMAL PHENOMENA FROM NEAR-DEATH EXPERIENCES by Titus Rivas, Anny Dirven and Rudolf H. Smit. Translated from Dutch by Wanda J. Boeke. Durham, NC, 2016: IANDS Publications. 367 pp., £14.99. ISBN 978-0-9975608-0-0

*'Girl fell from tree, visited Jesus in heaven
and awoke cured, mom claims in new book.'*

Headlines like this example from the *Huffington Post* (April 16, 2015) have tended to create misleading impressions in the general public as to what a near-death experience (NDE) actually is, and how convinced can those of us who have not had one be of their reality?

The authors define it (pp. 322–3) as: “a profound psychological experience typically occurring to persons who are close to death or who are in intense physical danger or emotional distress”. It is a syndrome, additional symptoms of which include feelings not of fear or panic as might be expected, but of love and peace, moving around out of the body, seeing a bright light, having a life review and being sent back to earth, sometimes very reluctantly, it is not always clear by what or whom.

The NDE is typically accompanied by such pleasant side-effects as absence of fear of death, and increased concern for the wellbeing of others. It is not surprising that such a generally pleasant and often transcendental experience is often interpreted, especially by children, as a visit to heaven and meeting with a deity of their choice. It is also worth noting that NDEs, unlike dreams, do not ‘die at the opening day’, but can live for a lifetime and have a powerful and lasting effect on those who experience them.

While fully aware of the perils of claiming absolute certainty in scientific matters nowadays, the authors feel able to state that “extrapolation from an NDE during a clinical death to personal survival after death is completely justifiable”. Moreover, “the presence of personal consciousness during a cardiac arrest implies that humans can expect personal survival after irreversible death”. This, they insist, is not irrational or wishful thinking, but rather “a rational conclusion from the type of cases that we presented in Chapter 3”.

The hundred-plus cases they refer to include some of the most widely publicised recent examples such as those of musician Pam Reynolds, best-selling author Eben Alexander, and other lesser-known patients of such pioneer NDE specialists as Raymond Moody, Michael Sabom, Melvin Morse, Penny Sartori and Sam Parnia. When the authors come across an exceptionally persuasive case, such as that of the anonymous hypothermia survivor and NDE percipient immortalised in the literature as ‘The man with the dentures’, they extract the last drop of relevant information from all the witnesses they could track down, making this an invaluable sourcebook for future students of the early days of NDE research.

There are several examples of sudden recovery during an NDE of serious medical conditions without any conventional treatment, some recalling cases of ‘miraculous cure’ at religious shrines such as Lourdes, which some prefer to label ‘spontaneous remissions’. A particularly impressive example is the case reported by author and experienced intensive care nurse Penny Sartori,

of a patient whose right hand had been deformed since birth, but seemed to heal itself after an NDE in which there was no verbal suggestion of healing. "It remains unexplained how it is possible for the patient to open and use his previously contracted hand", she noted.

Not surprisingly, such claims, however well documented, have attracted the attention of organised groups of self-proclaimed sceptics, some medically qualified, others not. The latter include mathematician and CSI (formerly CSICOP) Fellow Jan Willem Nienhuys, who has described NDE experiencers as "people who attach so much value to something that ultimately does not differ from dreams and chemical hallucinations", and who have "a loose screw in their head".

An NDE can provide a wealth of evidence for a variety of psi phenomena, and as the subtitle suggests, these can be of special interest provided that they have been properly verified, as all too often they have not been in the past. They include precognitive dreams, psychokinetic abilities and apparent encounters with discarnates. There are also several accounts of instantaneous healings even of supposedly incurable conditions, as in the case mentioned above of Annabel Beam, the girl who fell from a tree when she was nine years old and met both Jesus, (she assumed), and her deceased grandmother. The meetings may not have been fully verifiable but the subsequent cures undoubtedly were.

The authors' conclusion (p. 296) is concise and fully justified:

Having presented in this book over 100 cases of NDEs involving confirmed paranormal aspects, and having demonstrated the inadequacy of arguments that pseudoskeptics have proffered in an attempt to discredit paranormal NDEs, we are left with the conclusion that mind and brain are essentially independent, though closely associated during physical existence, and that consciousness continues when the brain and body have ceased to function.

7 Earls Court Square
London SW5 9BY
UK

glplayfair@waitrose.com

GUY LYON PLAYFAIR

NEAR DEATH EXPERIENCES: UNDERSTANDING VISIONS OF THE AFTERLIFE by John Martin Fischer & Benjamin Mitchell-Yellin. Oxford University Press. UK. 2016. 191 pp. £10.99. ISBN 978-0-19-046660-2

Dr John Martin Fischer is Distinguished Professor of Philosophy, University of California, who has written extensively on the subjects of free will, moral responsibility, and metaphysical and ethical issues pertaining to life and death. Dr Benjamin Mitchell-Yellin is Assistant Professor of Philosophy, Sam Houston University, who has written extensively on ethics, philosophy of action, modal logic, death and immortality. They have joined forces to try to understand the implications of NDE visions of an apparent afterlife and apparent veridical knowledge obtained when an Out-of-Body Experience (OBE) is part of a Near-Death Experience (NDE). They express their purpose by saying, "in these pages we wish to address the *implications*

of near-death experiences for questions about the fundamental nature of reality and the relationship between our minds and our brains. Do near-death experiences establish that there is an afterlife? Do they show that our minds can function separately from our brains? Our aim is to explore these questions with the seriousness and rigor they deserve". Later they add that "Our method, as intellectual descendants of Socrates, will be to consider the arguments on all sides and get to the bottom of things as best we can, using informed, compassionate reason and logic".

Despite their laudable intentions, their inquiry has made them very unpopular with those — particularly in the medical profession — who have committed themselves publicly to the position that NDEs clearly indicate brain/mind dualism in which death represents a transition to a brain-free, self-aware, afterlife in heaven. As far as they are concerned the *implications* of NDEs have already been done and dusted. Dr Pim van Lommel, for example, has even questioned Fischer and Mitchell-Yellin's credentials, asking them whether they have personally interviewed anyone who has had a NDE or personally corroborated veridical perceptions that occurred during a NDE, or attended anyone in cardiac arrest and assessed their clinical state. Fischer and Mitchell-Yellin's response is that they have not, nor do they need to do so. They question the implication in van Lommel's questions that being a medical doctor (MD) somehow qualifies one to answer the metaphysical questions posed by the occurrence of NDEs. To imply that speaking personally with people who have had a NDE somehow puts someone who is an MD in a position of categorically stating that NDEs '*verify the existence of the afterlife*' (their italics) is, in their opinion, nonsense. What you obtain from a personal interview is that person's account of their memory of a NDE, not an ontological answer regarding their occurrence simply because the interviewer is an MD. The NDE literature is there for all to read.

I have written this book review because someone of a similar opinion to van Lommel was approached before me, but responded by saying that they did not think the book merited a review in the *JSPR* or any other scholarly journal on the grounds that it adds nothing to the literature but simply repeated poorly-grounded philosophical arguments against survival, while studiously avoiding any meaningful discussion of the empirical evidence. If that refusal was based upon reading the book, then why not provide a review so as to point out its deficiencies and restate the grounds for one's firmly established supernaturalist position? If that refusal was based upon dismissive hearsay then the decision would seem to be based upon prior prejudice.

The authors point out that NDEs take two main forms. There are the purely heavenly ones, as famously experienced by Eben Alexander (2012, 2014), that describe visions of beautiful vistas, heavenly music and deceased friends and relatives living with other heavenly inhabitants including, in his case, "a beautiful girl with high cheek bones and deep blue eyes" wearing "peasant style clothes" and having "golden-brown tresses framing her lovely face" who "looked at me with a look that, if you saw it for a few moments, would make your whole life up to that point worth living". Speaking to Alexander without words, her encouraging message was:

“You are loved and cherished for ever”

“You have nothing to fear”

“There is nothing you can do wrong”

As, rather disappointingly in this case, we are unable to check such heavenly visions for ourselves, we can only take the experient's word for it that they had really entered a heaven that exists independently of themselves. The other form of NDE includes an associated OBE containing apparently veridical details that can be checked against consensual reality. On recovering consciousness the patient recalls seeing and hearing details of what was happening around them that can be confirmed by those who were there at the time. In chapter 2, ‘Two Famous Near-Death Experiences’, the authors consider the case of Pam Reynolds and the ‘Dentures in the trolley drawer’ case. Reynolds had a giant aneurysm near her brainstem that could burst at any moment. Following what must have been a detailed description by the surgeon of the three stages of the operation before she signed her consent form (a foreknowledge which, surprisingly, is not discussed by these authors, nor by Sabom [1998] who first reported her case, nor by van Lommel [2010] in his review of the case) she was conscious and aware when she entered the operating theatre and was greeted by the surgeon and clinical team.

The first stage involved being anaesthetised with her eyes covered while wearing ear plugs making loud clicks that were recorded as brainwaves on a screen and faded away to indicate unconsciousness. During this stage a catheter was inserted into a groin artery and the surgeon shaved part of her head and cut through her skull with a buzz saw shaped like a dentist's drill. She reported that she had floated out of the top of her head, saw that only part of her head was shaved, that she heard a woman doctor saying her groin arteries were small and catheter insertion was difficult, and saw the general theatre scene and the bone saw, looking rather like a dentist's drill, which buzzed in D natural. Later she experienced a bright light, had a heavenly NDE and returned very reluctantly to her body. In the second stage, when her body temperature was lowered and her brain was drained of blood, the surgeon excised the aneurysm, and in the third stage blood flow was re-established, her body re-warmed to normal temperature and she eventually regained consciousness.

There is much debate as to when the heavenly NDE occurred. The supernaturalist explanation is that during the first stage she separated from her brain, as a fully conscious self, perched on or near the surgeon's shoulder, and briefly saw and heard what was going on before entering heaven, meeting relatives and returning to rejoin her brain, possibly after she had been re-warmed and her brain was functioning again. It was a real-time OBE. The physicalist explanation favoured by the authors is that her apparently unconscious brain was still in a state where it registered meaningful speech and sound, but not the meaningless clicks which she does not mention, and these memories resurfaced into an imagined reconstruction then referred back by her and the medical staff to that time during the operation.

In the ‘Dentures in the trolley drawer’ case (for which we have only the male nurse's account but which is not in question) the Dutch patient had suffered a heart attack, collapsed in a field in winter, was brought in cold and

apparently lifeless, was resuscitated and his dentures removed and placed in a trolley so that an endo-tracheal tube could be inserted. He later recovered consciousness as an inpatient, saw the male nurse who had removed the dentures and reminded him where they were. He said he had floated above the scene and saw what was happening but did not say that he saw the nurse removing his dentures or putting them in the adjacent medical trolley. The authors propose that he partially 'came to' during the very vigorous resuscitation procedures that restored cerebral blood flow, felt the nurse open his mouth and remove his dentures, heard his very distinctive voice which was why he recognised him again several days later, and maybe heard his dentures drop onto the trolley shelf. Their physicalist explanations in both cases are in general agreement with those advanced by the anaesthetist G. M. Woerlee, whose very detailed review of each case can be found online. The difference is that in Reynold's case Woerlee suggests that she was partially conscious during the time of her apparent OBE.

Visual OBEs in the blind seem to offer strong support for the supernaturalist position because it is difficult to explain how someone blind can see the visual components of a cardiac resuscitation attempt? The authors point out that there are two main types of blindness: the no visual input blindness, where no visual stimuli from the retinas reach the cortical visual centres that convert them into conscious visual imagery; and cortical blindness, where the visual cortex is completely non-functional so visual imagery of any kind is impossible. In the former group — which is by far the larger — subjective visual imagery is very common as in what is termed Charles Bonnet syndrome, where people 'see' three dimensional fully animated scenes, as if 'out there' (cf. Sachs, 2012). The authors consider that some form of physicalist explanation seems far more likely than a supernaturalist one of a blind self leaving the unconscious brain, 'seeing' an actual outside scene, and re-engaging with the brain on recovery of consciousness to, presumably, download the memory.

Childhood NDEs are considered to offer strong support for the supernaturalist argument on the grounds that they are not preconditioned by adult experience and beliefs. They examine the case of Colton Burpo (2010) who, aged nearly four, suffered a burst appendix requiring two operations in quick succession and nearly died. He had a visionary NDE in which he visited heaven, met God, the Holy Spirit, sat on the knee of Jesus, saw angels, met John the Baptist, met 'Pop', his deceased grandfather, and a sister whom his mother had miscarried, saw a rainbow horse, and Todd, his Christian pastor father, praying for him in another room and his mother praying and talking on the phone. These details trickled out over months, sometimes spontaneously and sometimes in answer to questions, such as when his father asked him if he had ever seen God's throne, showed him a Bible story picture of a throne, and Colton replied "Oh yeah! I saw it a bunch of times". Brought up in a devout Christian family, his imagination was filled with Christian imagery, church services, sermons, Bible readings and conversations about his illness, miraculous recovery and his parents praying for him. The authors argue that he had a NDE but much of the vivid detail was 'remembered' in retrospect as an outcome of questions and hearing his

parents and others discussing his NDE. It had become a composite NDE filled out with later details as a completely guileless false memory. As for the miscarriage, the parents had discussed this with his older sister and he may have heard the conversation, or she may have told him. They also discuss the famous James Leininger case (Leininger, 2009) often presented as a model example of reincarnation, but as this was not a NDE their discussion will not be reviewed here. They point out that accounts of childhood NDEs are not primary sources described in their own words, but secondary accounts written by adults over months and sometimes years who inevitably reframe them according to their understanding and belief systems. Children are highly imaginative and can be very suggestible, particularly when their story means that they are the centre of adult attention.

They open their chapter entitled ‘A Strategy for Explaining “Near-Death Experiences”’ by saying that

It often seems as if certain phenomena must be ‘paranormal’ and therefore outside the reach of science. When we scratch below the surface, however, alternative explanations emerge — explanations that are well within the physicalist framework of science. We have contended in this book that near-death experiences appear to fit this pattern.

They advocate a “multi-factor physical explanation of near-death experiences that is sensitive to all their complexity, wonder and transformational power”. In effect, what we do not know today we will know tomorrow because physicalist interpretations allow for testable hypotheses. This includes the cognitive neurosciences with their ever-increasing understanding of brain functioning during blood loss through ever more sensitive brain scanning technologies. Rather surprisingly, although earlier in the book they mention and reference Borjigin et al.’s findings (2013a) as an example of such new information, they do not describe and discuss these findings which seem central to their argument. Borjigin and colleagues made the completely unexpected discovery that dying rat brains generate a period of intense, coordinated brainwave activity characterised by Borjigin et al. (2013b) as “demonstrating the presence of electrical fingerprints of consciousness in the near-death brain”. If further confirmed, this finding may carry implications for all dying brains, including the ‘electrical fingerprints’ of hyperconscious NDEs and OBEs of dying human brains occurring before successful resuscitation.

Their argument against what they term the “single strategy supernaturalist interpretation” is that it solves nothing, but instead raises unanswerable questions such as how the non-physical can possibly interact with the physical. How does an immaterial consciousness exit and re-enter a material brain before and after a NDE and what does the locality of ‘exit’ and ‘re-enter’ mean for a non-local non-physicality? When ‘outside’ their body, how does the non-physical self see the physical world during an OBE? Without eyes and visual cortex upon which the self is dependent during life, what does its disembodied self see with? How does it ‘download’ its NDE or OBE into the brain on its return to create the neural memory that seems essential for recall after regaining consciousness? Appealing to the sheer vividness and awesomeness of NDEs is not a good argument for supernaturalism. Similar emotions can be experienced when seeing a glorious sunset.

It must be stressed that Fischer and Benjamin Mitchell-Yellin do not argue against the possibility of an afterlife, but argue that such a possibility needs to be posited on grounds other than NDEs and OBEs. They discuss confirmation bias whereby we selectively seek evidence to support our comforting beliefs, as in meeting up again in an afterlife, but of course supernaturalists will argue that this goes both ways. What has raised the ire of MDs who support the supernaturalist interpretation is that the authors have not reviewed the clinical reasons for this interpretation. If brains after cardiac arrest are as moribund as present clinical signs seem to indicate, they are not capable of generating the neural correlates of NDEs. The only way they could be generated by is the mind leaving the brain, entering heavenly reality and returning with this experience for later downloading recall, or popping out of the brain to see what is going on around their recumbent body as in an OBE. This assumption of an almost functionless brain during NDEs and OBEs could be interpreted as support for dualism given that materialist accounts equate conscious experience with brain activity. This is not a very encouraging message for those whose NDE was not one of heaven but of a terrifying Hell and offers no explanation of how the returning mind downloads its NDE and/or OBE experience to create the neural correlates necessary for later conscious recollection. However, if Borjigin's 'electrical fingerprints of consciousness' interpretation is confirmed and generalises to all dying brains, then this could support a physicalist interpretation as this activity would create the necessary neural correlates and would bring NDEs and OBEs that occurred during cardiac arrest in line with the findings of neuroscientific research. In support of Borjigin et al., similar 'electrical fingerprints of consciousness' have been recorded from dying human brains (Chawla et al., 2009). What remains puzzling is that these important findings seem curiously absent from both sides of the debate, including the 'Near Death Experiences' entry in the SPR's Psi Encyclopedia (Sartori, 2016).

Should this book receive a review in scholarly journals such as the *JSPR*? Definitely. While it does not examine in depth the present clinical case for a supernaturalist interpretation, it presents an argument that should be heard and is well referenced. The book was one of many projects funded by the three year *Immortality Project: Philosophical and Theological Implications* (2012-2015), set up by the University of California and financed by the John Templeton Foundation, with Fischer as Project Leader. This source of funding by a Foundation that tends to seek a spiritualist interpretation of the sciences seems to have angered critics even more than the book itself which, in contrast, offers a physicalist argument.

9 Mountbatten Close, Cardiff CF23 5QG.
bigbobcharman@yahoo.co.uk

ROBERT A CHARMAN

REFERENCES

- Alexander, E. (2012). *Proof of heaven: A neurosurgeon's journey into the afterlife*. New York: Simon & Schuster.
- Alexander, E. (2014). *The map of heaven: How, science, religion and ordinary people are proving the afterlife*. New York: Simon & Schuster.
- Borjigin, J., et al. (2013a). Surge of neurophysiological coherence and connectivity in the dying brain. www.pnas.org/content/early2013/08/08/1308285110.abstract

- Borjigin, J., Wang, M.M., Mashour, G.A. (2013b). Reply to Greyson: Experimental evidence lays the foundation for a rational understanding of near-death experiences. *Proceedings of the National Academy of Science*, 110(47): E4406 Published online 2013 Nov 6.doi: 10.1073/pnas.1317581110
- Burpo, T., & Vincent, L. (2010). *Heaven is for real: A little boy's astounding trip to heaven and back*. New York: Thomas Nelson.
- Chawla, L.S., Akst, S., Junker, C., Jacobs, B., & Seneff, M.G. (2009). Surges of electroencephalogram activity at the time of death: A case series. *Journal of Palliative Medicine*, 12, 1095–1100.
- Leininger, B., Leininger, A., Gross, K. (2009). *Soul survivor: The reincarnation of a World War II fighter pilot*. New York: Grand Central.
- van Lommel, P. (2007). *Consciousness beyond life: The science of the near-death experience*. New York: Harper One.
- Long, J., & Perry, P. (2010). *Evidence of an afterlife: The science of near death experiences*. New York: Harper Collins.
- Sabom, M. (1998). *Light & death: One doctor's fascinating account of near-death experiences*. Grand Rapids, MI: Zondervan.
- Sachs, O. (2012). *Hallucinations*. New York: Alfred A. Knopf.
- Sartori, P. (2016). Near-death experience. <http://psi-encyclopedia.spr.ac.uk/articles/near-death-experience>.

GHOSTLY ENCOUNTERS: THE HAUNTINGS OF EVERYDAY LIFE by Dennis Waskul with Michele Waskul, Temple University Press, Philadelphia, 2016. Xi + 166pp. References, Notes, Index, Plates, Appendix. £20.99 (paperback). ISBN 978-1-4399-1289-8.

Dennis Waskul, a Professor of Sociology at Minnesota State University, and his wife, Michele, an independent scholar, state clearly in their Preface that *Ghostly Encounters* is intended as an academic book based on “reflexive ethnographic fieldwork” of reported experiences of ghosts and hauntings. They travelled extensively in the American Midwest interviewing people who reported first hand encounters with ghosts, and where possible visiting the sites of the alleged hauntings. In a few cases where participants preferred not to meet face-to-face, written accounts were accepted. Their university did not permit snowballing techniques, in which participants can recommend others who are then approached directly, so they were dependent on advertising for respondents. The results are related to existing academic literature and then packaged very much for a general, rather than solely academic, audience.

For those familiar with writer and historian Ian Wilson’s 1995 book *In Search of Ghosts*, the ground covered is similar (and even the cover images are almost identical). Both start with descriptions of initial scepticism followed by the story of a personal ghostly encounter. In Wilson’s case he and his wife were staying with friends in Abercrombie House, New South Wales, and didn’t know at the time that the room they were sleeping in had a reputation for being haunted. During the night both Wilson and his wife became aware of someone standing beside their bed, breathing audibly, as if trying to attract their attention. When they turned on the light the sound faded, only to return when they switched it off again. Tired from their journey and wishing to sleep, Wilson remembered the story of a haunting he had read about in which a ghost was banished by invoking the Holy Trinity. He thought

he would try something similar and mentally said a prayer releasing the ghostly presence. Much to his surprise, this seemed to work. In the Waskuls' case they were in bed at home when Dennis saw a wispy-white cloud come through the window and blinds into the corner of the room, and then send out tentacles towards him. Two short sentences came into his mind as he watched the apparition, "I will tell the truth. I will tell the story right". In Wilson's case, he had already been commissioned to write a book on ghosts when the experience occurred; the Waskuls were well into their research at this point. In neither instance did the experimenter feel afraid. The result of the encounters for both writers was to subtly shift their perspective on the subject of ghosts. It is perhaps significant in both cases that the authors were already engaged with the topic when these experiences took place.

Ghostly Hauntings is divided into five, fairly brief, chapters with an Appendix describing the methodology used. The first chapter presents ghosts as a "cross-cultural and trans-historical" phenomenon and simultaneously as "uniquely modern" (p. 18). What is meant here is that while stories of ghosts are universal the notion of the supernatural depends on a post-Enlightenment definition of the natural order. They make the point that while ghosts in popular imagination are a largely visual phenomenon, the spectre is only one way that spirits of the dead (if that is what they are) can make their presence known. They are as likely to be audible, either directly as in the Wilsons' experience, or through manipulating objects, making knocks and bangs, scratching on walls, or turning electrical appliances on and off. They might also make the living feel as if they are being touched or even choked, manipulate the temperature or otherwise give the impression of being watched. The experience may be individual or shared, lending the haunting a perception of veridical objectivity. In the second chapter, first-hand accounts of ghostly encounters are interwoven with interviewees' interpretations of what has occurred, placing the experience within a North American cultural context in which popular interest is combined with scepticism. Several interviewees were nervous of being perceived as mad or deluded or sought ordinary rational explanations for their encounters. Chapter Three, with numerous quotes and examples from the research data, attempts a typology of ghosts, listing intelligent hauntings, residual hauntings, anniversary and historical hauntings. Forms of ghosts are divided into apparitions (visible ghosts), phantasms (a visual appearance in a dream or altered state of consciousness), wraiths (a person who visits the living around the time of his or her death), poltergeists (noisy or restless ghosts), spectres (a threatening or menacing ghost) and phantoms (a spectre occurring in a dream). Each chapter ends with an extended case study, in this case of a child ghost named Madison who communicated with both the informant and her younger sister independently when they were a similar age to the ghostly girl. The ghost identified herself as a previous occupant of the house, and could be manipulative and jealous, a disturbing 'friend' for a young child.

Chapter Four takes a rather different direction with an account of Loon Lake Cemetery in Minnesota, a sad story in which the reputation of the site as haunted gave ghost hunters *carte blanche* to destroy and desecrate it. A Nineteenth Century inhabitant of the cemetery (unjustly) acquired a

reputation as a witch. Legends of ghostly and supernatural goings on at the site were promulgated and repeated in local legend and in print. Descendants of the accused are still working to clear her name, and the chapter is included as a cautionary tale against revealing the locations of supposed hauntings. Chapter Five tackles the central questions of epistemological and ontological relativity and certainty. It is generally taken for granted in the social sciences that truth is contingent on perspective, culture, history, symbolic frameworks, and so on. In other words, truth is epistemologically relative. What about the ontological ‘reality’ of ghostly phenomena? The study is not directly studying ghosts but people’s accounts of ghosts. The personal experience recounted at the start of the book did, however, have the effect of opening the author’s mind to the possibility that the accounts they collected are based on actual occurrences, and that they may involve the spirits of the dead, rather than being psychological projections or hallucinations.

Waskul, like Ian Wilson before him, ends the book with a conversation with a medium and an unusual occurrence which suggests that some people may have privileged access when it comes to communication with the dead. Dennis Waskul was warned that he was in danger of inviting a spirit attachment, and subsequently experienced some inexplicable phenomena, including his wedding ring disappearing from his hand and reappearing some hours later on the step of his office, and a poltergeist taking control of his laptop, deleting messages before he could read them. Waskul mentions mediums, religious specialists and paranormal investigators as a resource for communicating with troublesome ghosts (but not spirit release therapists). I also wonder if Waskul was aware of the SPR and ASPR archives, as he claims that with the possible exception of Diane Goldstein’s *Scientific Rationalism and Supernatural Experience Narratives* (2007), his is probably the first empirical study of reported first-hand experiences with ghosts (p. 150).

Despite an apparent lack of awareness of earlier research, *Ghostly Encounters* is an interesting and lively read. As it is based almost exclusively on first-hand accounts (unlike the Wilson volume) some of the problems involved in checking the veracity of second-hand accounts are avoided (although the first-hand accounts are only as good as the memory and narrative ability of the interviewees). For those who are sceptical it is hard to argue with first-hand experience. The possible interpretations of these experiences are set in their cultural context, while acknowledging the similarity of ghostly narratives across time and culture. For those familiar with the topic there is little that will come as a surprise or seem particularly original, but the volume can serve as a very useful introduction to ghosts and hauntings for the inquisitive and discerning general or student reader.

5 Castleford Gardens
Tutshill, Chepstow NP16 7LF
fiona.bowie@kcl.ac.uk

FIONA BOWIE

REFERENCES

- Goldstein, D., Grider, S., & Thomas, J. (2007). *Haunting experiences: Ghosts in contemporary folklore*. Logan: Utah State University Press.
Wilson, I. (1995). *In search of ghosts*. London: Headline.

CORRESPONDENCE

To the Editor:

On Robert Charman's Research Note: Seeing an Apparition at Glasgow Airport

Robert A. Charman's thoroughgoing "Research Note" on the Robert Macleod Glasgow airport case states that "apart from some local publicity at the time ... the only current account of the case" can be found in Tricia Robertson's 2013 *Things You Can Do When You Are Dead: True Accounts of After Death Communication* (cited in Charman, 2016, p. 149). Ron Halliday also writes up the case in his *Haunted Glasgow* (2008, pp. 61–62), but here Hambleton-Jones — who witnessed the apparition of his former colleague, Macleod, on the airport concourse — was not retired, as Charman has it (p. 149), but a cargo pilot about to depart Glasgow for Moscow (p. 61). Unlike the painstaking research evinced by Charman, Halliday does not cite his sources for the Macleod story.

I also cover the Macleod case in my collection of found poetry that focuses on ghosts and poltergeists that have the ability to speak and/or write (Thistleton, in press). My own research indicates that Glasgow, incidentally, has a good claim to be inducted as the world's capital of these speaking and writing entities. Of the hundred or so cases that I survey worldwide, seven of them are from Glasgow.

"Lachlan's"
314 Glen Oaks Road
Greendale NSW 2550
AUSTRALIA
claythistleton@gmail.com

CLAY THISTLETON

REFERENCES

- Charman, R.A. (2016). Research note: Seeing an apparition at Glasgow airport—The Robert Macleod and Robert Hambleton-Jones case reconsidered. *Journal of the Society for Psychological Research*, 80(3), 149-157.
- Halliday, R. (2008). *Haunted Glasgow*. Ayr, UK: Fort Publishing.
- Thistleton, C. (in press). *Gef the talking mongoose and other ... poems*. London: Blart Books.

To the Editor:

Cesare Lombroso's "Psychiatric Explanation" of Eusapia Palladino

In papers published some years back in this journal I discussed ideas of radiations from the body of mediums proposed to explain physical phenomena (Alvarado, 2006; Alvarado & Nahm, 2011). More recently, while working on a paper about Cesare Lombroso (Alvarado & Biondi, 2017) I studied in more detail Lombroso's ideas of mediumistic forces not covered in the *JSPR* papers,

which I would like to present briefly in this letter. I hope these ideas are of interest to readers of the *Journal* because they help us to learn more about Lombroso and about the various ideas of concepts of force proposed to explain physical mediumship in relation to the medium Eusapia Palladino. Lombroso's ideas are particularly interesting because they combined both belief in the actual existence of physical phenomena and pathological process. This was different from the better known conceptual tradition in which mediumship was seen as completely pathological, without the acceptance of veridical phenomena (for overviews see Alvarado & Zingrone, 2012; Le Maléfan, 1999).

Lombroso was well-known for his ideas of abnormality, and atavistic features in criminals, the insane, and women. Furthermore, he also had a strong interest in psychic phenomena, and further developed his ideas of pathology in the case of Palladino. The medium had convinced him that her phenomena were real (Lombroso, 1892).

Lombroso (1892) characterized Palladino as a neuropath that, in addition to showing lack of tactile sensitivity, was "subject to the accesses of epilepsy, catalepsy and hysteria, which occur mainly during mediumistic phenomena" (p. 146, this, and other translations, are mine). The following addressed the production of a mediumistic force, which he believed were connected to cerebral anomalies of the medium:

I am not going to deny as inadmissible the excitation of some centers in hysterics and the hypnotized which become powered through the paralysis of all the other [centers], and also cause a transposition and a transmission of psychic forces which can also bring a transformation to luminous or to motor force. Thus we understand how the force of a medium, which I call cortical or cerebral, can, for example, raise a table, pull a beard, beat, caress someone, phenomena quite frequent in these cases (p. 146)

In some conditions, Lombroso also argued, thought could be projected out of the body. This could change into a motor force, and in turn, into light and heat.

Such was, in his view, the explanation of the medium's physical phenomena, but he did not present details about how the force could leave the body and what principle guided it to produce, in different situations, the various phenomena of Palladino, many of which showed intelligence and responded to the request of some sitters.

In a book published in the year of his death Lombroso (1909) continued to mention mediumistic forces, but he believed that some phenomena may have a spirit origin.

He also continued to write about Palladino's pathologies. Among other things he pointed out: "She has the hyperaesthetic zone, especially in the ovary. She has the bole in the oesophagus that women with hysteria have, and general weakness, or paresis, in the limbs of the left side" (p. 107).

Others also commented on Palladino's pathology, and related it to the projection of a force from the body (e.g., Morselli, 1908). This tendency to relate what was believed to be genuine mediumship to pathology was not limited to Lombroso or to Palladino (e.g., Fairfield, 1875). Similarly, others speculated on veridical mental phenomena related to bodily unbalances, and

presented such observations with hysterical individuals, a topic I cannot discuss here but that points to a theoretical tradition that deserves further discussion.

Parapsychology Foundation
P.O. Box 1562
New York, NY, 10021
carlos@theazire.org

CARLOS S. ALVARADO

REFERENCES

- Alvarado, C.S. (2006). Human radiations: Concepts of force in mesmerism, spiritualism and psychical research. *Journal of the Society for Psychical Research*, 70, 138–162.
- Alvarado, C.S., & Biondi, B. (2017). Classic Text No. 110: Cesare Lombroso on mediumship and pathology. *History of Psychiatry*, 28, 225–241.
- Alvarado, C.S., & Nahm, M. (2011). Psychic phenomena and the vital force: Hereward Carrington on “Vital Energy and Psychical Phenomena”. *Journal of the Society for Psychical Research*, 75, 91–103.
- Alvarado, C.S., & Zingrone, N.L. (2012). Classic Text No. 90: ‘The Pathology and Treatment of Mediomania’, by Frederic Rowland Marvin (1874). *History of Psychiatry*, 23, 229–244.
- Fairfield, F.G. (1875). *Ten years with spiritual mediums: An inquiry concerning the etiology of certain phenomena called spiritual*. New York: D. Appleton.
- Le Maléfan, P. (1999). *Folie et spiritisme: Histoire du discours psychopathologique sur la pratique du spiritisme, ses abords et ses avatars (1850–1950)*. Paris: L’Harmattan.
- Lombroso, C. (1892). Le spiritisme et la psychiatrie: Explication psychiatrique de certains faits spirites. *Annales des Sciences Psychiques*, 2, 143–151.
- Lombroso, C. (1909). *After death – what? Spiritistic phenomena and their interpretation*. Boston: Small, Maynard.
- Morselli, E. (1908). *Psicologia e “Spiritismo.” Impressioni e note critiche sui fenomeni medianici di Eusapia Paladino* (2 vols.) Turin: Fratelli Bocca.
-

NOTICE

Prof. Roe has given notice that he intends to step down from his role as Editor of the *Journal* after the October 2017 issue. The SPR Council should therefore like to receive expressions of interest for this position. Those wishing to be considered should send a CV that includes particular reference to prior experience of journal editing and reviewing and a short (up to 500 words) personal statement that outlines why the applicant would be a suitable candidate. Applications or enquiries should be sent by the closing date, 1 July 2017, to Prof. Bernard Carr, Chair, Education and Publicity Committee, The Society for Psychical Research, 1 Vernon Mews, London W14 0RL or emailed to journal@spr.ac.uk