© 2022 AIPR, Inc.
ISSN: 1445-2308

Australian Journal of Parapsychology
Volume 22, Number 1, pp. 7-28

The Case for Postmortem Survival from the Winners of the Bigelow Institute for Consciousness Studies Essay Contest: A Level of Evidence Analysis

PATRIZIO TRESSOLDI, ADAM J. ROCK, LUCIANO PEDERZOLI, & JAMES HOURAN

Abstract. We assessed the 29 winning essays of the Bigelow Institute of Consciousness Studies (BICS) contest using an evidence hierarchy approach adopted in many scientific fields. Two independent judges rated the target essays for their quality of scientific evidence, reproducibility, and replicability using an evidence hierarchy adapted from several published models that accommodate both qualitative and quantitative evidence. According to our criteria, six essays (20.7%) were categorized as the highest level of scientific evidence, four essays (13.8%) were categorized at a medium level of scientific evidence, whereas the remaining 19 essays (65.5%) were considered a low level of scientific evidence. The overall agreement of the essay rankings between the present authors' classifications of evidence quality and the rating system used by the BICS judges was only 44.8%, with a non-significant Spearman's *rho* correlation of .03. This result indicates extremely little concordance (overlap) of the two evaluation systems, which corroborates prior research on the critical shortcomings of evidence hierarchies. The essays representing the highest level of scientific evidence per our criteria involved near-death experiences and mental mediumship. For other anomalies that ostensibly support the survival hypothesis (e.g., physical mediumship or electronic voice phenomena), more studies with refined experimental designs are needed to improve their quality of evidence as defined in current scientific terms. Important considerations and future research directions are likewise discussed.

Keywords: anomalous experiences, consciousness; evidence hierarchy, postmortem survival, replication

INTRODUCTION

In June 2020, Robert T. Bigelow, a famous aerospace entrepreneur from the United States, founded the Bigelow Institute for Consciousness Studies (BICS) to support research on the prospect of postmortem survival of human consciousness and the associated nature of any such state. To these ends, BICS launched an international contest to award contestants who wrote the most compelling essays about the best available evidence for the survival hypothesis. According to the purposes of the BICS contest,[1] the intention was "to provide a public service by ... providing essays focused on scientific evidence as well as objective and subjective supported documentation as gathered:

- "From special cases, including older cases, from very credible witnesses;
- "From photographic or electronic data;
- "From all available literature;
- "From highly validated and authenticated human experiences;
- "From other relevant sources."

Evaluation of the essays was assigned to six judges, i.e., Christopher C. Green, Leslie Kean, Jeffrey J. Kripal, Harold Puthoff, Jessica Utts, and Brian Weiss.[2] They agreed "that true (or veridical) evidence includes a combination of a wide variety of forms; scientific, experiential, witnessed, repeatable, anecdotal and *otherwise persuasive far beyond rules of traditional evidence-based hypothesis tested research paradigms* [emphasis added]".[3] The panel members independently assessed all 205 submissions and sent their discrete rankings confidentially to the BICS headquarters so as not to influence the other judges. The majority votes determined the final rankings of the essays with corresponding cash awards. BICS announced the outcomes on November 1, 2021.[4]

The contest stoked much publicity, as well as some notable discussion and controversy among sympathetic and skeptical researchers alike. Accordingly, the present study sought to re-analyze the contents of the 29 winning essays following standardized criteria currently used to evaluate the quality of scientific evidence based on traditional rules. This alternative approach does not intend to discredit the BICS judging procedure, but rather to serve as a valuable comparison. In this way, we

1 https://www.bigelowinstitute.org/about.php

2 https://www.bigelowinstitute.org/judges.php

3 https://www.bigelowinstitute.org/docs/ApplicationForm2021.pdf

4 https://www.bigelowinstitute.org/News4.php

hope to (a) contextualize the outcomes and implications of the BICS essay contest, and (b) generate new information and knowledge to help future researchers assess and weigh the evidential value of different scientific methodologies used in this domain.

THEORETICAL BACKGROUND: LEVELS OF SCIENTIFIC EVIDENCE

'Evidence' takes many forms because the scholarly analysis of information can involve different qualitative and quantitative methodologies. The biomedical and clinical fields generally adhere to certain guidelines when evaluating the scientific support for specific claims. This assists in the identification and adoption of evidence-based or empirically supported therapeutic interventions or treatments (Balshem et al., 2011; Blunt, 2015; Guyatt, et al., 2008a, 2008b; Sakaluk et al., 2019). These evaluation criteria are customarily known as a 'hierarchy of evidence (or levels of evidence).'

The discipline of psychology has similarly used evidence hierarchies as tools to rate the quality of evidence for phenomena, including ego depletion (Friese et al., 2019) and the purported link between violent video games and physical aggression (Prescott et al., 2018). Furthermore, these hierarchical frameworks consider direct or conceptual replications (LeBel et al., 2009) as the gold-standard for scientific evidence given the current reproducibility or replicability crisis impacting many scientific fields (e.g., psychology, economics, and neuroscience) (Munafò et al., 2017; Nosek et al., 2022). Replicable evidence involves independent studies that use the same experimental design, sample characteristics, materials and procedures, or with variations testing the generalizability of previous findings (Schmidt, 2009).

The rationale underlying all evidence hierarchies is that they ostensibly describe an ascending weight of evidence from multiple studies corresponding to *increased methodological quality* and *decreased risk of bias* (e.g., randomized controlled trials testing multiple participants and no single cases). Though widely used, the application of evidence hierarchies also has been criticized on conceptual and practical grounds. Most notably, Blunt (2015) examined the facts and logic underlying the development, use, and interpretation of medical evidence hierarchies. He concluded that, "hierarchies in general embed untenable philosophical assumptions: principally that information about average treatment effects backed by high-quality evidence can justify strong recommendations, and that the impact of evidence from individual studies can and should be appraised in isolation" (p. 3). This clearly implies that such hierarchies can be a poor basis for the evaluation of evidence. To be sure, over 80 evidence hierarchies have been

proposed, and there is no obvious or objective way to judge which is most accurate or useful. Adding to the confusion, developers of different hierarchies have suggested different interpretation schemes and not all consistently agree on what constitutes or counts as 'evidence.' However, evidence hierarchies ultimately represent a method for ranking the quality of *methodologies* versus evidence *per se*.

THE PRESENT STUDY

Blunt's (2015) cautions and recommendations guided the selection of an evidence hierarchy for use here. It was important for the framework to accommodate both qualitative and quantitative research, so we conducted a scoping review of suitable models. We identified four published hierarchies with ample flexibility to address the different types of evidence presented in the BICS essays (i.e., Daly et al., 2007; Ho et al., 2008; LoBiondo-Wood & Haber, 2018; Steele & Tiffin, 2014). The overlapping components of these models were then used to create the hierarchy of scientific evidence shown in Figure 1.

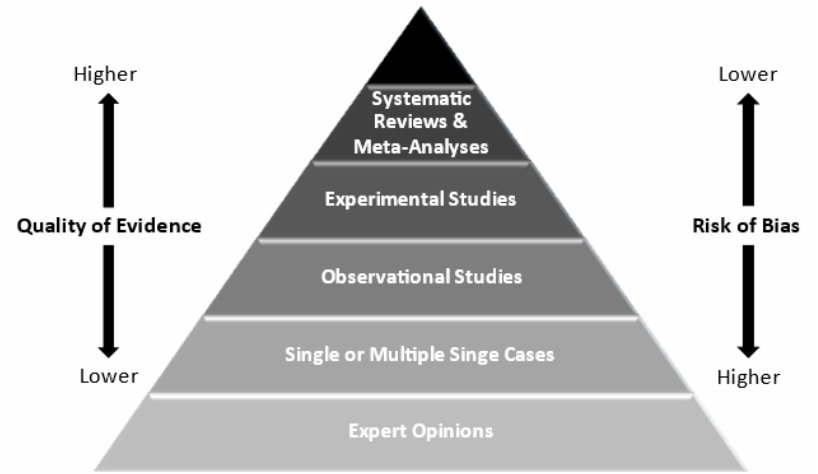

Figure 1. Author-modified 'hierarchy of scientific evidence' used in the present study.

Our modified hierarchy should be interpreted as other proposals. Namely, the five levels of evidence considered here are depicted by a pyramid as each level — from bottom to top — reflects the quality of research designs (*increasing*) and quantity (*decreasing*) of each study design in the published literature. For example, systematic reviews are higher quality and more labor intensive to conduct, so a lower quantity is typically published.

The five levels of scientific evidence considered here are further explained. The lowest level comprises *experts' opinions* that can be obtained from interviews or questionnaires. The second level contains discrete or multiple *single cases*, which are first or third-person descriptions of individual experiences. The third level of *observational studies* refers to investigations about specific phenomena carried out on random or specific samples of participants by using questionnaires, interviews, specific tasks, etc., but without control of independent or moderator variables.

The fourth level consists of *experimental studies*, that is, investigations with random or specific samples of participants where specific phenomena are observed under controlled conditions, which typically test two or more different hypotheses, e.g., "Are mediums' accuracy above or within chance expectations?" or "Are near-death experiences similar or different to other autobiographical experiences?"

The fifth and final level concerns *systematic reviews*. This involves the selection and discussion of all available evidence related to a specific phenomenon following precise inclusion and exclusion criteria (e.g., type of studies, date of publication range). The database of a systemic review might be further scrutinized with *meta-analysis*, i.e., a statistical approach that combines the results from multiple studies to increase power (over individual studies), improve estimates of the size of empirical effects, and to resolve uncertainty when reports disagree.

Of course, the *reproducibility* and *replicability* of investigations must also be considered when assessing levels of evidence. This is done by inspecting a study's 'open science practices,' or the various practices that allow independent researchers to reproduce a study or potentially replicate the results of a study (see e.g., Nosek et al., 2015). In order for a study to be 'reproducible,' the author(s) must accurately describe how the study was carried out (e.g., how participants were selected, the procedure, the materials used, how the data was analyzed).

Additionally, the author(s) must provide open access to all materials, raw data, and the codes necessary to allow the possibility to reproduce or reanalyze the original findings independently. On the other hand, results are 'replicable' when the author(s) accurately describe how the study was performed (e.g., how participants were selected, the procedure and materials used, how the data was analyzed) in order for independent researchers to

test the robustness and the generalizability of the results by replicating the reported findings with different samples, materials, and even with some modifications to the procedure (Dixon & Glover, 2000).

Based on the preceding, the present study had two aims: (a) to examine the *scientific strength*, *reproducibility*, and *replicability* of the evidence for the survival hypothesis cited by the 29 BICS-winning essays, and (b) to compare our final rankings of the reported evidence to those from the BICS judges based on the criteria above and operationalized by our evidence hierarchy.

METHOD

Materials

The 29 winning essays were open access and thus downloaded from the BICS webpage on 24 November 2021.[5] These works were classified into three categories: (a) the top three winners (BICS contest category 1), (b) the eleven winners of $50,000 (BICS contest category 2), and (c) the 15 winners of $20,000 (BICS contest category 3).

Procedure

Co-authors PT and AR designed a basic approach and agreed to independently rate each of the top four-ranked BICS essays as examples using the following scientific criteria:

- **Study materials**, that is, the source of information. We differentiated the scientific from the general literature. 'Scientific' pertains to all studies conducted using shared scientific methods (e.g., systematic observation of the phenomenon by using first and third-person sources, its correlates and causes) and experimental designs, mostly published in peer-reviewed journals. 'General' refers to all materials available online, in books, private correspondence, etc., mostly including personal opinions and interpretations, as well as data observed or collected with non-scientific methods, etc. For each essay, we computed the percentage of scientific literature with respect to all the cited references.

[5] https://www.bigelowinstitute.org/contest_winners3.php

- **Level of scientific evidence (LSE)** in decreasing rank order: (a) *Level 5*—quantitative and qualitative meta-analyses or systematic reviews of scientific studies; (b) *Level 4*—experimental studies; (c) *Level 3*—observational studies; (d) *Level 2*—single cases of first-person direct experiences with or without third-person concurrent validation; and (e) *Level 1*—experts or personal opinions.
- **Reproducibility**, that is, the information included in the essay (i.e., data, procedure to analyze them) was sufficient to independently reproduce the original findings. This was a simple 'Yes/No' criterion. This item was applied only to the essays including novel findings obtained by their authors and not to those referring to published scientific and non-scientific literature.
- **Replicability**, that is, independent authors had replicated the purported evidence for a phenomenon. This was also a 'Yes/No' criterion. It was sufficient for our evaluation that there was at least *one* replication.

Scoring. Unlike submissions to peer-reviewed journals, the BICS essays were not mandated to follow specific standards for their content or structure. This situation required several nuances in our evaluation of these disparate presentations:

First, we adopted a more liberal criterion to classify essays as 'systematic reviews,' as none of the authors satisfied international guidelines (e.g., Appelbaum et al., 2018; Page et al., 2021). This category was thus applied to essays that included most of the scientific studies available up to 2020 related to the phenomena discussed in the essays.

Second, we adopted a 'prevalence criterion' to compute the LSE metric, because most essays cited and discussed various sources of information (e.g., single cases, experimental studies, and meta-analyses). Specifically, we identified and scored the sources of information on which the authors most frequently relied in their arguments (e.g., their analysis and interpretation of the different phenomena focused on mostly single-cases or experimental studies with control conditions).

Third, the 'Level of Scientific Evidence Total Score' (LSETS) was obtained by summing the LSE score (range = 1 to 5) plus one point each if the 'reproducibility' and 'replicability' items were judged as "Yes." This yielded a total score that could range from 0 to 7, with higher scores denoting evidence of stronger scientific merit or weight. In order to reasonably equate the LSETS to the three winning categories of the BICS contest, we also devised three 'Level of Scientific Evidence Categories' (LSECat) with the following cut-offs:

(a) Cat 3 ('low quality of evidence') = LSETS ≤ 3;
(b) Cat 2 = LSETS > 3 to ≤ 5 ('moderate quality of evidence'); and
(c) Cat 1 = LSETS > 5 ('high quality of evidence').

Interrater reliability. The first-round percentage of interrater agreement for LSECat ratings of the BICS essays was 70%, which translates to a significant Spearman's rank-order correlation, $r_s(27) = .31$, $p < .05$ (two-tailed). Therefore, the use of a standardized evidence hierarchy helped to uniformly evaluate the evidence cited in the essays. The interrater agreement then reached 100% after the raters collectively discussed and resolved their discrepancies via an expert panel approach to enhance the reliability of the final classifications (see e.g., Bertens et al., 2013).

RESULTS

Table 1 lists the LSE category, LSECat metric that corresponds to ascending levels of scientific evidence, and the BICS contest category (BCat) for each of the 29 winning essays. The full database with all scores related to the LSE, reproducibility, and replicability, is available open access to interested readers.[6]

Crosstabulation between the three LSECat rankings and the three BCat classifications revealed a 13/29 (or 44.8%) agreement, which translates to a non-significant correlation, $r_s(27) = .03$, $p = .885$ (two-tailed). This indicates a vast difference between our quality rankings of the evidence cited in the essays and those by the BICS judges. Indeed, among the six essays in the first LSECat (authors' names in **bold**), only one was assigned to BCat = 1 (i.e., van Lommel, 2021), two were assigned to BCat = 2 (i.e., Beischel, 2021; Long, 2021), and the remaining three were assigned to BCat = 3 (Delorme et al., 2021; Parnia et al., 2021; Roe et al., 2021). These outcomes have important theoretical and practical implications as we next explore in detail.

DISCUSSION

We aimed to contextualize the BICS contest results by assessing the quality of the winning essays via a standardized 'levels of evidence' analysis. Our results arguably showed that the contest's criteria to accept "credible witnesses" and "highly validated and authenticated human

[6] https://doi.org/10.6084/m9.figshare.17211878.v3

experiences" as sources of evidence was quite disputable—even untenable—from a scientific standpoint. Indeed, serious problems with witness reliability are well-documented in the empirical literature (Głomb, 2021; Loftus, 2019) as well as particularly salient to reports of spontaneous cases of anomalous or altered perceptions (Houran et al., 2017; Lange et al., 2019).

Table 1
LSE, LSECat, and BCat Variables for Each of the 29 BICS-Winning Essays

Author(s)	*LSE*	*LSECat*	*BCat*
Beischel, J.	Systematic review	1	2
Delorme, A., et al.	Systematic review	1	3
Long, J.	Systematic review	1	2
Parnia, S., et al.	Systematic review	1	3
Roe, C., et al.	Systematic review	1	3
van Lommel, P.	Systematic review	1	1
Fenwick, P.	Observational studies	2	3
Kastrup, B.	Experimental studies	2	2
Kerr, C.	Observational studies	2	3
Rousseau, D. & Billingham, J.	Observational studies	2	2
Braude, S.	Multiple single cases	3	2
Carter, C.	Multiple single cases	3	3
Cook, N.	Multiple single cases	3	3
Krohn, E.	Single case study	3	2
Leininger, B.	Single case study	3	3
Mays, R. & Mays, S. B..	Multiple single cases	3	3
Meyer zu Erpen, W.	Multiple single cases	3	3
Mishlove, J.	Experts' opinions	3	1
Nahm, M	Multiple single cases	3	2
Neppe, V.	Multiple single cases	3	3
Rawlette, S.	Multiple single cases	3	2
Rocha, A.	Single case study	3	2
Rouleau, N.	Expert's opinion	3	2
Ruickbie, L.	Multiple single cases	3	1
Sommer, A.	Multiple single cases	3	3
Taylor, G.	Multiple single cases	3	3
Taylor, S.	Multiple single cases	3	3
Tymn, M.	Multiple single cases	3	2
Weerasekera, A.	Single case study	3	3

Note: LSE: Level of Scientific Evidence category; LSECat = Level of Scientific Evidence rating; BCat = BICS contest ranking.

The issue of evidence quality in the contest was further complicated by the acceptance of nonscientific literature ("other relevant sources") as sources of evidence. We therefore conclude that the winning essays grounded in scientific papers represented legitimate evidence of high scientific quality, whereas nonscientific literature only served to weaken the arguments for postmortem survival.

Not surprisingly given Blunt's (2015) caution about different evidence hierarchies, our scientific criteria produced different rankings of the 29 essays as compared to the conclusions of the six judges of the BICS contest. However, we would be remiss not to mention those points of limited agreement between the two systems. According to the data presented in Table 1, six out of 29 essays (20.7%), were included in the first ('best') LSECat ranking. This percentage rises to 34.5% by adding the four essays included in the second category. This concordance might be viewed as a sort of conceptual replication for some of the original BICS outcomes. As such, these essays that were consistently ranked as citing high quality evidence can help to pinpoint those research areas with arguably the most evidential value for the survival hypothesis.

Among the essays included in the first LSECat, three are related to *near-death-experiences* (NDEs) (i.e., Long, 2021; Parnia et al., 2021; van Lommel, 2021), one to *mental mediumship* (Beischel, 2021), and two to the above phenomena plus other experiences such as *after-death-communication, reincarnation, and haunt-type episodes* (Delorme et al., 2021; Roe et al., 2021). Interestingly, Laythe and Houran (in press) also found in their analysis of the available evidence for survival that 'haunt/poltergeist episodes' and "veridical anomalous experiences" like after-death-communication (e.g., ADCs) had the most promise for obtaining witness testimony that can withstand counter-arguments and cross-examination in a forensic sense.

NDE studies are routinely published in mainstream journals, so it is perhaps expected that three essays related to these experiences obtained the highest scientific strength scores. Though most of the scientific literature examined in the three essays is common, each included complementary and specific information. For example, van Lommel (2021) discussed the mind-brain relationship in NDEs relative to other phenomena (e.g., placebo, meditation, end-of-life experiences) as convergent support for the primacy of consciousness and its nonlocal characteristics. In contrast, Long (2021) discussed the frequency of some new NDE characteristics extracted from his large and probably unique database of first account NDEs; for example, the characteristics and vividness of sensorial information, the comparison of the consciousness level and alertness with that in the normal everyday condition, NDEs memory accuracy with respect to other personal life events, and so on. Parnia et al. (2021), given their expertise in resuscitation

procedures, presented interesting details about brain physiological processes that occur after cardiac arrest and how they differ in the transition from reversible to irreversible cell damage.

However, Beischel's (2021) essay was a systematic review of the evidence obtained from rigorous empirical investigations of 'mental mediumship.' Gauld (1982) described this phenomenon as communication with deceased persons that is experienced "through interior vision or hearing, or through the spirits taking over and controlling their bodies or parts thereof, especially … the parts required for speech and writing" (p. 4). Beischel was one of the first authors to devise experimental designs that guarded against conventional communication of information between the medium and the sitters who requested a mediumship consultation about a deceased person. In this way, it was possible to quantify the level of accuracy of the medium's reading and investigate whether the medium's reception of information involved the 'telepathic scanning' of the sitter's mind (i.e., the super-psi hypothesis) or communication with the deceased (i.e., the survival hypothesis). Beischel contended that the findings from this line of investigation clearly support the survival hypothesis.

The essays from Roe et al. (2021), and Delorme et al. (2021), represented two systematic reviews of the evidence related to many phenomena concerning the survival hypothesis. Whereas Roe et al. considered the evidence obtained from the scientific literature and described some specific cases, Delorme et al. presented an original hierarchy of the scientific evidence of such phenomena and the results of a survey about which experimental evidence is more convincing for the survival of human consciousness. For the examination of the scientific strength of the different phenomena, Delorme et al. (2021) devised a structured grading system ranging from 'Grade A' (strong evidence) to 'Grade F' (no evidence), similar to our rating approach used here. For Grade A, the scientific evidence must be obtained by prospective, blinded, pre-registered and meta-analyzed experimental studies, replicated by independent groups, not explainable by materialist science and not requiring statistical analyses to be observed. Grade F (no evidence) alternatively must be obtained either by systematic fraud or obvious documented and undocumented flaws (for further details, see their Table 1, p. 10). Following their criteria, mental and physical mediumship obtained the best grade, corresponding to B+, followed by NDEs and reincarnation studies with B-, electronic voice phenomena (EVP), instrumental transcommunications (ITC) and death-bed visions with C+, and apparitions, induced experiences of survival and after-death-communications with grade C. Their survey further found that two experiments were rated as being more convincing support for the survival hypothesis; that is: (a) the study testing the vision of randomly selected images by patients scheduled for a cardiac arrest, very similar to the

AWARE study (Parnia et al., 2014), and (b) a mental mediumship investigation in which hospice patients agreed to contact one or more mediums after they died.

LIMITATIONS AND CONSIDERATIONS

All hierarchies of scientific evidence have strengths, weaknesses, and nuances (Blunt, 2015). To be sure, different evidence hierarchies produce drastically different outcomes, and inconsistent application of a given hierarchy will likewise yield inconsistencies. It is thus not unexpected that our results would differ from those of Delorme et al. (2021) or the BICS judges. These discrepancies certainly highlight some important limitations of this study and any related process to evaluate the scientific quality of evidence. Though supported by a general consensus, we first acknowledge that our LSE hierarchy and level scores were arbitrary. A different hierarchy and scoring system for each LSE might have substantially changed the LSECat metric and hence the final results. Furthermore, an interrater evaluation using more than two raters—and perhaps with different opinions about the phenomena in question—would produce more balanced and robust ratings for subsequent analysis.

Most importantly, the LSE score only accounted for the *type* of study (e.g., single case study vs. systematic review) and not its internal *quality*. This same constraint perhaps applies to the evaluations by Delorme et al. (2021) and the BICS judges as well. A meaningful quality check of the proffered evidence would entail independent reviewers who critically examine each cited study in every winning essay. Such a laborious process would nevertheless deliver a maximally meticulous evaluation of the primary sources of evidence and the overall trends that emerge from their collective consideration. In fact, some authors advocate dividing the hierarchy levels into sub-levels based in part on study methodology and how rigorously it was conducted, while others suggest that poor methodology will knock a study down a level (Atkins et al., 2004). From this perspective, a well-executed investigation at the bottom of an evidence hierarchy can produce more reliable or valid evidence (and hence be rated as higher quality) than a carelessly conducted study towards the top. On this point, major factors that can decrease the quality of evidence include: (a) study limitations, (b) inconsistency of results, (c) indirectness of evidence, (d) imprecision, (e) publication bias, and (f) plausible confounding, which would reduce a demonstrated effect (Guyatt et al., 2008a, p. 996).

Study quality is not a trivial issue or confound. For instance, some researchers in the biomedical sciences have found a huge array of low-quality systematic reviews that were more likely to report ‘positive

findings' than higher-quality reviews (e.g., Płaszewski & Bettany-Saltikov, 2014). Or consider meta-analyses that show statistically significant effects but are based on studies with poor quality of evidence that negates their practical relevance (Cleminson & McGuire, 2016). To make matters worse, it can be tricky when examining the methods section of a study to see how it was conducted—that is, if a task or action was not reported as having been done (such as the method of randomization), this does not necessarily mean it was neither conducted nor effectively so (Devereaux et al., 2002). We therefore agree with those authorities (e.g., Blunt, 2015; Golden & Bass, 2013; Guyatt et al., 2008a) who caution that the demonstration of significant effects is different than our confidence level in these effects.

FUTURE RESEARCH DIRECTIONS

The popularity of the BICS contest showed that the survival hypothesis is a lively and burgeoning scientific domain. Understanding levels of evidence helps investigators to contextualize and prioritize information, but this does not imply that Level 1 evidence should be ignored, and all Level 5 evidence accepted as fact. Obviously, a grounded hierarchy provides a guide but readers must be ultra-cautious when interpreting the results. But echoing Blunt (2015), our study strongly suggests that it will not be enough to work from a well-defined evidence hierarchy to assess ongoing evidence rigorously and consistently. To our way of thinking, the quality of survival-related evidence arguably has yet to be explored comprehensively using structured methods advocated in the clinical and biomedical sciences (for a discussion, see for example Yetley et al., 2016, pp. 250S-251S).

In this respect, the areas of agreement between our rankings and the BICS outcomes are noteworthy for identifying those phenomena that might be the most promising for future research. In contrast to Delorme et al. (2021), we contend that end-of life, shared death and after-death communication experiences, as well as reincarnation phenomena are currently investigated with 'good' scientific methodologies (e.g., Elsaesser et al., 2021; Masayuki, 2017; Moraes et al., 2021; Penberthy et al., 2021; Shared Cross Initiative, 2021; Tucker, 2008), whereas physical mediumship, EVP, and ITC require more advanced experimental designs in order to support their ontological reality (Drinkwater et al., 2020; Williams et al., 2021; Wiseman & Greening, 2005).

Taken altogether, we hope that the BICS contest has raised greater or wider scientific interest in possibly the most fundamental existential question aside from the inherent nature of human consciousness, and that more researchers become involved in this line of scientific exploration. On

this point, there is one last feature of the BICS judging panel that new studies can profitably emulate. It has been noted that analysis and interpretation is increasingly being conducted by cross-disciplinary or transdisciplinary teams, which are arguably best equipped to address complex challenges (Tebes et al., 2014). Accordingly, the best chance for exponential progress on tackling the complexities of the survival hypothesis might not come from lone studies or laboratories, but rather pooled resources, coordinated efforts, and adversarial collaborations among biomedical experts, parapsychologists, and other researchers in consciousness studies.

ACKNOWLEDGMENTS

We thank Lance Storm and reviewers for their comments and suggestions. We thank the Proof Reading Service for English revisions. In full disclosure we note that the last author was a co-entrant in the BICS contest (cf. Laythe & Houran, in press) albeit their submission did not place. However, our study was structured such that this factored neither in the rankings nor the associated conclusions presented here.

CREDIT AUTHOR STATEMENT

PT: Conceptualization; Methodology; Validation; Formal analysis; Investigation; Data curation; Visualization; Writing - original draft preparation; Writing-Review & Editing; AR: Methodology; Validation; Formal analysis; Investigation; Data curation; Visualization; Writing - original draft preparation; Writing-Review & Editing; LP: Conceptualization; Writing - Review & Editing; JH: Visualization; Writing-Review & Editing.

THE AUTHORS

Patrizio Tressoldi (Ph.D.) is Head of the Science of Consciousness Research Group (https://socrg.org), and Senior Scientist affiliated with the Studium Patavinum of Padova University, Italy (www.patriziotressoldi.it). He conducts research in science of consciousness, nonlocal mind functions, and non-ordinary states of consciousness.

Adam J. Rock (Ph.D.) Adam J. Rock is a Senior Lecturer in Psychology at the University of New England (Australia) and an Associated Distinguished

Professor of Integral and Transpersonal Psychology at the California Institute of Integral Studies.

Luciano Pederzoli (M.Eng.) obtained a doctorate in electronic engineering in Bologna in 1971. He is Head of the EvanLab in Florence, Italy (https://www.evanlab.org/en). He has used hypnosis techniques to investigate past-life regression, psychic archaeology and out-of-body experience.

James Houran (Ph.D.) holds a Master's Degree in Clinical Psychology (1996) from the University of Illinois at Springfield, USA, and a Doctorate in Medicine (Psychology) (2004) from the University of Adelaide, Australia. He serves as a Research Director at Integrated Knowledge Systems, Research Professor at the Laboratory of Statistics and Computation, ISLA in Vila Nova de Gaia, Porto, Portugal, and the Editor-in-Chief of the *Journal of Scientific Exploration*.

REFERENCES

(References marked with an asterisk indicate BICS contest essay winners)

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist, 73*(1), 3–25. https://doi.org/10.1037/amp0000191

Atkins, D., Eccles, M., Flottorp, S., Guyatt, G. H., Henry, D., Hill, S., et al. (2004). Systems for grading the quality of evidence and the strength of recommendations I: Critical appraisal of existing approaches The GRADE Working Group. *BMC Health Services Research*, *4*(1), 38. https://doi.org/10.1186/1472-6963-4-38

Balshem, H., Helfand, M., Schünemann, H. J., Oxman, A. D., Kunz, R., Brozek, J., Vist, G. E., Falck-Ytter, Y., Meerpohl, J., Norris, S., Guyatt, G. H. (2011). "GRADE guidelines 3: Rating the quality of evidence - Introduction.. *Journal of Clinical Epidemiology*, *64*(4), 401–406. https://doi.org/10.1016/j.jclinepi.2010.07.015

Bertens, L. C. M., Broekhuizen, B. D. L., Naaktgeboren, C. A., Rutten, F. H., Hoes, A. W., van Mourik, Y., Moons, K. G. M. & Reitsmaet, J. B. (2013). Use of expert panels to define the reference standard in diagnostic research: A systematic review of published methods and reporting. *PLOS Medicine*, *10*(10), e1001531. https://doi.org/10.1371/journal.pmed.1001531

*Beischel, J. (2021, November 24). Beyond reasonable: Scientific evidence for survival. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/Julie_Beischel.pdf

Blunt, C. (2015). Hierarchies of evidence in evidence-based medicine. Doctoral dissertation, London School of Economics and Political Science. http://etheses.lse.ac.uk/3284

*Braude, S. (2021, November 24). A rational guide to the best evidence of postmortem survival. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/Stephen_Braude.pdf

*Carter, C. (2021, November 24). The case for the afterlife. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/1_Chris_Carter.pdf

Cleminson, J., & McGuire, W. (2016). Topical emollient for preventing infection in preterm infants. The Cochrane database of systematic reviews, *Cochrane Database of Systematic Reviews*, *2016*(1), CD001150. https://doi.org/10.1002/14651858.CD001150.pub3

*Cook, N. (2021, November 24). What is the best available evidence for the survival of human consciousness after permanent bodily death? Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/2_Nick_Cook.pdf

Daly, J., Willis, K., Small, R., Green, J., Welch, N., Kealy, M., & Hughes, E. (2007). A hierarchy of evidence for assessing qualitative health research. *Journal of Clinical Epidemiology*, *60*(1), 43–49. https://doi.org/10.1016/j.jclinepi.2006.03.014

*Delorme, A., Radin, D. & Wahbeh, H. (2021, November 24). Advancing the evidence for sSurvival of consciousness. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/3_Dean_Radin_et_al.pdf

Devereaux, P. J., Manns, B. J., Ghali, W. A., Quan, H., & Guyatt, G. H. (2002). The reporting of methodological factors in randomized controlled trials and the association with a journal policy to promote adherence to the Consolidated Standards of Reporting Trials (CONSORT) checklist. *Controlled Clinical Trials*, *23*(4), 380–388. https://doi.org/10.1016/S0197-2456(02)00214-3

Dixon, P., & Glover, S. (2020) Assessing evidence for replication: A likelihood-based approach. *Behavior Research Methods*, *52*(6), 2452–2459. https://doi.org/10.3758/s13428-020-01403-6

Drinkwater, K., Denovan, A., Dagnall, N., & Parker, A. (2020). Predictors of hearing electronic voice phenomena in random noise: Schizotypy, fantasy proneness, and paranormal beliefs. *Journal of Parapsychology*, *84*(1), 96–113. https://doi.org/10.30891/jopar.2020.01.09

Elsaesser, E., Roe, C. A., Cooper, C. E., & Lorimer, D. (2021). The phenomenology and impact of hallucinations concerning the deceased.

BJPsych Open, *7*(5), e148. https://doi.org/10.1192/BJO.2021.960

*Fenwick, P. (2021, November 24). To be and not to be. This is the answer: Consciousness survives. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/4_Peter_Fenwick_et_al.pdf

Friese, M., Loschelder, D. D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2019). Is ego depletion real? An analysis of arguments. *Personality and Social Psychology Review*, *23*(2), 107–131. https://doi.org/10.1177/1088868318762183

Gauld, A. (1982). *Mediumship and survival: A century of investigation*. William Heinemann.

Głomb, K. (2021). How to improve eyewitness testimony research: theoretical and methodological concerns about experiments on the impact of emotions on memory performance. *Psychological Research*, *86*, 1-11.

Golden, S. H., & Bass, E. B. (2013). Validity of meta-analysis in diabetes: Meta-analysis is an indispensable tool in evidence synthesis. *Diabetes Care*, *36*(10), 3368–3373. https://doi.org/10.2337/dc13-1196

Guyatt, G. H., Oxman, A. D., Kunz, R., Vist, G. E., Falck-Ytter, Y., Schünemann, H. J., & GRADE Working Group (2008a). What is "quality of evidence" and why is it important to clinicians? *BMJ* (*Clinical research ed.*), *336(*7651), 995–998. https://doi.org/10.1136/bmj.39490.551019.BE

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008b). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, *336*(7650), 924-926. https://doi.org/10.1136/bmj.39489.470347.AD

Ho, P. M., Peterson P. N., & Masoudi, F. A. (2008). Evaluating the evidence: Is there a rigid hierarchy? *Circulation*, *118*(16), 1675–1684. https://doi.org/10.1161/CIRCULATIONAHA.107.721357

Houran, J., Lynn, S. J., & Lange, R. (2017). Commentary on Stokes' (2017) quest for "white crows" in spontaneous cases of psi. *Australian Journal of Parapsychology*, *17*(1), 61-88.

*Kastrup, B. (2021, November 24). A rational, empirical case for postmortem survival based solely on mainstream science. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/Brnardo_Kastrup.pdf

*Kerr, C. (2021, November 24). Experiences of the dying: Evidence of survival of human consciousness. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/5_Christopher_Kerr.pdf

*Krohn, E. (2021, November 24). The eternal life of consciousness. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/Elizabeth_Krohn.pdf

Laythe, B., & Houran, J. (in press). Adversarial collaboration on a *Drake-S* Equation for the survivial question. *Journal of Scientific Exploration*.

LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A brief guide to evaluate replications. *Meta-Psychology*, *3*, 1–9. https://doi.org/10.15626/MP.2018.843

*Leininger, B. (2021, November 24). Consciousness survives physical death definitive proof of reincarnation. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/6_Bruce_Lehninger.pdf

LoBiondo-Wood, G., & Haber, J. (2018). *Nursing research: Methods and critical appraisal for evidence-based practice*. Mosby Elsevier.

Loftus, E. F. (2019). Eyewitness testimony. *Applied Cognitive Psychology*, *33*(4), 498-503.

*Long, J. (2021, November 24). Evidence for survival of consciousness in near-death experiences: Decades of science and new insights. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/Jeffrey_Long.pdf

Masayuki, O. (2017). Same-family cases of the reincarnation type in Japan. *Journal of Scientific Exploration*, *31*(4), 551–571.

*Mays, R. & Mays, S. B. (2021, November 24). There is no death: Near-death experience evidence for survival after permanent bodily death. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/7_Robert_Mays.pdf

*Meyer zu Erpen, W. (2021, November 24).Pursuit of best evidence for survival of human consciousness after permanent bodily death. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/8_%20Walter_Meyer_zu_Erpen.pdf

*Mishlove, J. (2021, November 24). Beyond the brain: The survival of human consciousness after permanent bodily death. Bigelow Institute. https://www.bigelowinstitute.org/docs/1st.pdf

Moraes, L. J., Barbosa, G. S., Castro, J. P. G., Tucker, J. B., & Moreira-Almeida, A. (2021). Academic studies on claimed past-life memories: A scoping review. *EXPLORE*. https://doi.org/10.1016/j.explore.2021.05.006

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., … Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

*Nahm, M. (2021, November 24). Climbing mount evidence: A strategic assessment of the best available evidence for the survival of human consciousness after permanent bodily death. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/Michael_Nahm.pdf

*Neppe, V. (2021, November 24). What is the best available evidence for the survival of human consciousness after permanent bodily death?. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/9_Vernon_Neppe.pdf

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/SCIENCE.AAB2374/

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Almenberg, A. D., ... & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology*, *73*(1), 719-748. https://doi.org/10.1146/annurev-psych-020821-114157

Page, M. J., Mckenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. https://doi.org/10.1136/bmj.n71

Parnia, S., Spearpoint, K., de, V. G., Fenwick, P., Goldberg, D., Yang, J., … Schoenfeld, E. R. (2014). AWARE-AWAreness during REsuscitation-A prospective study. *Resuscitation*, *85*(12), 1799–1805.

*Parnia, S. & Keshavarz Shiraz, T. (2021, November 24). What is the best available evidence for the survival of human consciousness after permanent bodily death? Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/10_%20Sam_Parnia_et_al.pdf

Penberthy, J. K., Pehlivanova, M., Kalelioglu, T., Roe, C. A., Cooper, C. E., Lorimer, D., & Elsaesser, E. (2021). Factors moderating the impact of after death communications on beliefs and spirituality. *Omega - Journal of Death and Dying*, advance online publication. https://doi.org/10.1177/00302228211029160

Prescott, A. T., Sargent, J. D., & Hull, J. G. (2018). Meta-analysis of the relationship between violent video game play and physical aggression over time. *Proceedings of the National Academy of Sciences*, *115*(40), 9882-9888. https://doi.org/10.1073/pnas.1611617114

Płaszewski, M., & Bettany-Saltikov, J. (2014). Non-surgical interventions for adolescents with idiopathic scoliosis: An overview of systematic reviews. *PLOS ONE*, *9*(10), e110254. https://doi.org/10.1371/journal.pone.0110254

*Rawlette, S. (2021, November 24). Beyond death: The best evidence for the survival of human consciousness. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/Sharon_Rawlette.pdf

*Rocha, A. (2021, November 24). Mediumship as the best evidence for the afterlife: Francisco Candido Xavier, a White Crow. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/Alexandre_Roch_et_al.pdf

*Roe, C., Cooper, C., Lorimer, D., & Elsaesser, E. (2021, November 24). A critical evaluation of the best evidence for the survival of human consciousness after permanent bodily death. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/11_Chris_Roeet_et_al.pdf

*Rouleau, N. (2021, November 24). An immortal stream of consciousness. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/Nicolas_Rouleau.pdf

*Rousseau, D. & Billingham, J. (2021, November 24). What would have to be true about the world? On evidence for the possibility of consciousness surviving death. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/David_Rousseau_et_al.pdf

*Ruickbie, L. (2021, November 24). What is the best available evidence for the survival of human consciousness after permanent bodily death? Bigelow Institute. https://www.bigelowinstitute.org/docs/3rd.pdf

Sakaluk, J. K., Williams, A. J., Kilshaw, R. E., & Rhyner, K. T. (2019). Evaluating the evidential value of empirically supported psychological treatments (ESTs): A meta-scientific review. *Journal of Abnormal Psychology*, *128*(6), 500–509. https://doi.org/10.1037/abn0000421

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, *13*(2), 90–100. https://doi.org/10.1037/a0015108

Shared Cross Initiative. (2021). Shared death experiences: A little-known type of end-of-life phenomena reported by caregivers and loved ones. *American Journal of Hospice and Palliative Medicine*, *38*(12), 1479–1487. https://doi.org/10.1177/10499091211000045

*Sommer, A. (2021, November 24). What is the best available evidence for the survival of human consciousness after permanent bodily death?. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/12_Andreas_Sommer.pdf

Steele, R., & Tiffin, P.A. (2014). 'Personalised evidence' for personalised healthcare: Integration of a clinical librarian into mental health

services: A feasibility study. *The Psychiatric Bulletin*, *38*(1), 29–35. https://doi.org/10.1192/pb.bp.112.042382

*Taylor, G. (2021, November 24). What is the best available evidence for the survival of human consciousness after permanent bodily death? Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/13_Greg_Taylor.pdf

*Taylor, S. (2021, November 24). The next phase of the journey: The evidence for the survival of human consciousness after permanent bodily death. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/14_Steve_Taylor.pdf

Tebes, J. K., Thai, N. D., & Matlin, S. L. (2014). Twenty-first century science as a relational process: From eureka! to team science and a place for community psychology. *American Journal of Community Psychology*, *53*(3–4), 475–490. https://doi.org/10.1007/s10464-014-9625-7

Tucker, J. B. (2008). Children's reports of past-life memories: A review. *Explore*, *4*(4), 244-248.

*Tymn, M. (2021, November 24). Long concealed, now revealed: Overwhelming evidence for life after death. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/Michael_Tymn.pdf

*van Lommel, P. (2021, November 24). The continuity of cvonsciousness: A concept based on scientific research on near-death experiences during cardiac arrest. Bigelow Institute. https://www.bigelowinstitute.org/docs/2nd.pdf

*Weerasekera, A. (2021, November 24). What is the best available evidence for the survival of human consciousness after permanent bodily death?. Bigelow Institute. https://www.bigelowinstitute.org/Winning_Essays/15_%20Akila_Weerasekera_et_al.pdf

Williams, J. M., Carr, M., & Blagrove, M. (2021). Sensory processing sensitivity: Associations with the detection of real degraded stimuli, and reporting of illusory stimuli and paranormal experiences. *Personality and Individual Differences*, *177*, 110807. https://doi.org/10.1016/j.paid.2021.110807

Wiseman, R., & Greening, E. (2005). 'It's still bending': Verbal suggestion and alleged psychokinetic ability. *British Journal of Psychology*, *96*(1), 115–127. https://doi.org/10.1348/000712604X15428

Yetley, E. A., MacFarlane, A. J., Greene-Finestone, L. S., Garza, C., Ard, J. D., Atkinson, S. A., Bier, D. M., Carriquiry, A. L., Harlan, W. R., Hattis, D., King, J. C., Krewski, D., O'Connor, D. L., Prentice, R. L., Rodricks, J. V., & Wells, G. A. (2017). Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: Report from a

joint US-/Canadian-sponsored working group. *American Journal of Clinical Nutrition*, *105*(1), 249S–285S. https://doi.org/10.3945/ajcn.116.139097

Patrizio Tressoldi
Science of Consciousness Research Group
Padova University
Padova
ITALY

Email: patrizio.tressoldi@unipd.it

Adam J. Rock
School of Psychology
University of New England
Armidale, NSW
AUSTRALIA

Email: arock@une.edu.au

Luciano Pederzoli
EvanLab
Via dei Ricci, 22
Impruneta, Firenze 50023
Tuscany
ITALY

Email: pederzoli@inwind.it

James Houran
Integrated Knowledge Systems
7041 Briarmeadow Drive
Dallas, Texas 75230
USA

Email: jim_houran@yahoo.com