26-May-2020  
  
Dear Dr. Gervais:  
  
Manuscript ID SPPS-20-0136 entitled "The Origins of Religious Disbelief: A Dual Inheritance Approach" which you submitted to Social Psychological and Personality Science, has been reviewed.  The comments of the reviewer(s) are included at the bottom of this letter.  
  
Let me add a little context to the reviews. First and foremost, I think the most fundamental challenge is to respond to the issues raised by Reviewer 1 in the relatively brief space you're allotted in SPPS. This is not to say that I disagree with any of Reviewer 2's points, as I don't, but that (s)he has requested less of you and hence it will be more manageable to address these concerns.   
  
As you consider how to respond to these reviews, let me add that there were points at which I struggled to follow exactly what you were saying - not because the writing is unclear, but because I'm not expert in the area and you have (by necessity) laid out the problem rather quickly. From my perspective, the intro seemed overly brief and the discussion overly long. If you could provide a clearer description of the problem up front, and why your approach is informative and important, then I think less will be demanded of you in the discussion section.  
  
I was also concerned about the large rate of participant exclusion in the study, as it is a potential threat to the study's validity. Perhaps you could address this concern by footnoting analyses that include everyone, whether they passed the attention check or not.  
  
Finally, let me add that if you choose to revise and resubmit the manuscript, I might run it by Reviewer 1 again. Additionally, I will probably run it by an expert in Bayesian analyses, as that was remiss of me not to address that issue in the first submission.  
  
To revise your manuscript, log into <https://mc.manuscriptcentral.com/spps> and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions."  Under "Actions," click on "Create a Revision."  Your manuscript number has been appended to denote a revision.  
  
When submitting your revised manuscript, you will be able to respond to the comments made by the reviewer(s) in the space provided.  Please also use this space to document any changes you made to the original manuscript.  In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the reviewer(s).  
  
You will be unable to make your revisions on the originally submitted version of the manuscript.  Instead, revise your manuscript using a word processing program and save it on your computer.  Once the revised manuscript is prepared, you can upload it in step 5 of the revision submission process.  
  
IMPORTANT:  Your original files are available to you when you upload your revised manuscript.  Please delete any redundant files before completing the submission.  
  
Because we are trying to facilitate timely publication of manuscripts submitted to Social Psychological and Personality Science, your revised manuscript should be uploaded as soon as possible.  If it is not possible for you to submit your revision in a reasonable amount of time, we may have to consider your paper as a new submission.  
  
Once again, thank you for submitting your manuscript to  Social Psychological and Personality Science. I look forward to receiving your revision.  
  
Sincerely,  
Bill von Hippel  
Guest Action Editor  
  
Reviewer(s)' Comments to Author:  
  
Reviewer: 1  
  
Comments to the Author  
(Apologies for the formatting here. I had hoped to upload a document, but I need to paste into a textbox).   
  
  
The present article quantifies the relationship between religiosity (operationalised as low scores using Jong et al 2013) and several proxy-measures associated with theories of belief/disbelief. The authors argue that they test several competing models that purport to explain disbelief, and find strong support for the [absence of] CREDs, corresponding with the dual-inheritance model of disbelief. Analytic disbelief (operationalized using a measure of cognitive reflection) also predicted disbelief, but to a lesser extent. Reported (male) gender identity predicted disbelief. The strongest predictor in the model was being socially liberal. The authors sampled a nationally representative group of people living in the US.   
  
The authors employ bayesian statistics. I will admit here and now that I am not sufficiently expert to review the specifics of their analysis, and I would caution against publication until a competent reviewer can evaluate. (Though I will make some comments regarding analysis).   
  
Overall, this is a simple and elegant paper, and represents a meaningful contribution that extends beyond the disbelief literature.  
  
As it stands, the paper can be read and largely understood as is (with the caveat that not everyone understands Bayes). However, the brevity in explaining and justifying some decisions poses a challenge, though I concede this is likely a product of the word count associated with the journal rather than a lack of thought on the part of the authors. Though I believe some of the claims  made in the discussion over-state the work’s significance, these do not detract from the findings (and in this spirit, I would much prefer words to be spent on explaining certain decisions rather than contextualising the results as a contribution to formal theories and non-WEIRD psychology - two things this paper is not).   
  
Major points:  
  
All the measures used to operationalize various models, with the exception of CREDs, are indirect. Though they are treated as though are direct. This occurs in figure 1, where model-names are substituted for measurement-values. This occurs throughout the discussion. This is to confuse the map for the terrain, even though they must correspond. Please be consistent and refer to measures as measures when technically appropriate.   
  
Relatedly, the CREDs items ask about religion, but the other three scales measure something trait-like about the individual generally. That cognitive reflection predicts atheism less than CREDs is not surprising, given that reflection was measuring nothing so direct as CREDs. I don’t think at present, on the strength of the evidence, that one can claim one model of disbelief is better than another model of disbelief, because the measures themselves are not equivalent. This critique may, however, be unavoidable, as the models work at different levels of specificity to the individual. This will need to be addressed in some depth.   
  
The SBS measure, and the CREDs measure had high reliability. The other scales all had low reliability (usually < .8). While heuristically satisfactory, the low reliabilities likely suggest that the true degree of predictive power is an under-estimate. This blog entry illustrates the issues with reliability, and implicitly demonstrates how the measures in the present paper - having quite varied differences in reliability - are likely to predict different amounts of variance in the outcome simply by virtue of increased measurement error. <https://medium.com/@Sam_D_Parsons/ignoring-measurement-reliability-is-a-real-life-horror-story-b98a2517db26>.   
  
I’m not sure exactly how the authors can address this. The data has already been collected, and has obvious merit. However the conclusions about the relative value of each model of disbelief do not seem as firm to me as they appear presented in the paper.   
  
This fact, coupled with how (in)direct the measures are of the purported construct will need some dedicated focus.   
  
To what extent are the measures correlated and co-linear? Probably not a great deal, but since the study really only has 5 focal variables, I would argue that the authors should include a correlation matrix.    
  
Is there a reason none of the regressions use a person’s own attendance/participation in religious events as a predictor of belief? Determining a causal arrow is impossible in the present context, but it’s not clear why the data was collected if it wasn’t really going to be used (was it only for the readers interest in the demographics? Surely that can’t be the case, because you specifically recruited a nationally representative sample.). I can’t think of a sensible justification for why the authors did not use religious-participation to predict religious-(dis)belief. Certainly their pre-reg document doesn’t include it, but what justification is there for ignoring such a relevant measure at that point in the planning? Certainly even atheists attend religious events at some non-0 rate (and all your variables are continuous, anyway. Except for the binary belief in god, but even so). If the authors so-choose to not run this analysis, I think they need to defend that decision. We must keep in mind that CREDs not only predicts how viewers ought to respond to perceiving others performing costly actions, but speaks to one’s own motivation, ability and willingness to participate in such costly action. Relative frequency of participation itself (once a year vs. one a week) should be accounted for a feature that maintains (dis)belief.   
  
The SBS regression, and the logistic regression are not independent analysis. Nor is the ‘individual zero-sum replication analysis’. These cannot sensible be regarded as converging evidence. (Unless somehow bayes accounts for this?). What is the Point-Biserial Correlation between SBS scores and the binary scores? Does it justify running both analysis? And I don’t understand the logic of the ‘replication analysis’. You’ve demonstrated in a more-complete model that some variables are predictive; what value is there is running simpler models that account for less variance? (See Middling point 1)  
  
I won’t impugn the authors for using Bayes. Though it must be admitted that far fewer people understand Bayes than they do frequentist statistics, and not only is it unclear why bayes was chosen here, it’s not sufficiently explained. That’s not to say it’s not justifiable, nor understandable. Just that it isn’t justified, and not easily understood by a naive reader. Certainly, scholars who use bayesian analysis must get sick of explaining bayes, but the burden of clear communication and understanding is not solely on the reader, but on the person with the message. At a minimum the authors should link to a primer for the reader. The authors also introduce terms without explanation (such as ‘golems’).    
  
Being socially liberal is the strongest predictor in the model. If I’m reading HPDI correctly (presuming analogy to CI’s), the effect of being socially liberal doesn’t even overlap with the effect of CREDs. This cannot simply be ignored. It might suggest that adoption of ‘socially liberal values’ is incompatible with religious belief. To the extent that any causation is inferred in the present analyses, this inference (or others of the author’s choosing) needs to be discussed. Particularly because CREDs correspond with social learning, and the other features are cognitive. And yet the single biggest predictor in the model is a social factor (presumably capturing a constellation of beliefs and values). What if being social liberal (and everything that broadly means) just doesn’t gel with US-style religion? (Certainly, being religious might ipso facto make someone identify as liberal, but even if this is the case, it needs to be discussed. Also lending support for correlation tables - how dissociable are the other predictors from being liberal? It might be the case that CREDs are orthogonal, but cognitive reflection is not).   
  
  
Middling points:  
  
The authors are pretty loose with the term ‘replication’. In part III they double-dip on analysis. Had the individual analyses been presented first (demonstrating validity of your measures and generalizability from the literature), then combined into your focal analyses… then maybe this would be persuasive. But right now you present the focal analysis, make inferences about which models of disbelief are valid, then fail to demonstrate previously published effects. Either you replicated the individual effects and found them valid, thus allowing you to reject them as plausible explanation for disbelief, OR, you failed to replicate the effects and deem them invalid primae facie, but if so, you can’t then justify their contribution to your focal analysis. It strikes me that a) they need to be in your focal analysis, because precedent demands it, and b) failure to find an effect doesn’t mean the variables aren’t contributing to the overall model in important ways. In part III, the authors claim: That one of the candidate factors culled from existing literature did not appear as a robust predictor may suggest tempered enthusiasm for its utility as a predictor of individual differences in religiosity more broadly. But this claim is not more valid for being tested in isolation than it was when tested in a more complete model (assuming no major issues of colinearity).   
  
In the discussion the authors conflate ‘cognitive reflection’ with ‘analytic atheism’. Surely one could only make the claim that it’s analytic atheism if one measured the actual reasons for why one holds their belief (God is [not] real, and I am convinced by the arguments; God is real, and I know this as a matter of faith). The measures used are mostly just logical/maths questions.   
  
The terminology about ‘replication-plus’ seems… unnecessary? Is this an established idea and name? It feels as though you’re rebranded the tried-and-true method of reading the literature, synthesizing it, and extending it by a small, sensible, incremental step. What am I missing here that makes ‘replication plus’ different from the ordinary process of reflection and extension? Even if this is the case, I would strongly urge the authors to use their limited word count to better justify their choice of measures, or the clarity of their analysis.  
  
I generally refrain from commenting on framing, because the paper is not my own, and I believe the authors should have the right to whatever frame they want (and the reader can ignore those things anyway, paying attention to the data and analysis alone if they so wish). That said, I raise the following cautiously, and I mean no offence, but I cannot see a clear justification for these decisions.  
  
The authors discuss both ‘veories and WEIRD psychology, lauding the value of formal theories and non-WIERD samples. Yes this research is both informal (with typical elements of statistical inference) and WEIRD.  In the discussion the authors write: If this general pattern holds across societies [of our findings], we predict that—beyond religion—veories developed by WEIRD researchers to explain the weird mental states of WEIRD participants can only aspire to ever more precisely answer a mere outlier of an outlier of our most important scientific questions about human nature.   
  
Are the authors claiming their own findings, presented here, are a ‘mere outlier of an outlier’? This is the penultimate paragraph of your paper. The paper celebrates non-WEIRD research but concludes with a call for other scholars to do the harder parts of the research? The impression I received when reading this - which I fully acknowledge wouldn’t have been the wilful intention of the authors - is to associate the virtues of formal, and non-WEIRD, on their own work which is is neither.   
  
  
Minor points:  
  
The SBS paper is cited as Jong 2012, but that’s the ‘fox hole’ paper. The Jong et al paper with the scale is 2012, I believe.  
  
The authors mask themselves on the title page, but their links to their OSF in the MS include Will Gervais’ name in the MS, and the OSF files list all other authors. It is the case I already knew who the authors on this MS were prior to accepting the review (and declared this with the editor). So I suppose I raise this for future reference (the authors may look into toggling how to blind their osf profiles for such materials, and using a link-masking service for the addresses).   
  
Concerns about the pre-reg  
I commend the clarity of the pre-reg document. However, there are some incongruencies.   
  
I.1 says that linear and quadratic terms will be analysed. Assuming the individual replication analysis are bayesian regressions, this has not been done. ‘Religion’ for I.1, I.2, I.3, and I.4 is not clearly specified (i.e., jong vs. binary measure. Please clarify).   
  
The authors appear to have entirely skipped several hypotheses pertaining to the interactions. If this paper is the paper the pre-reg is associated with, these need to be run and reported.  
  
The authors also talk about machine learning and split-half cross-validation techniques. These are not done. Is there another paper in which they are being reported?   
  
Were exploratory analysis conducted? If so, they should at least be acknowledge in the paper (e.g., We conducted additional exploratory analysis but do not report them here). I know that if this were my data, I would have explored it, and I would have included as a predictor in the full model religious-participation. Since the intention to conduct exploratory analysis was registered, would the authors be willing to run their full model with the inclusion of the religious-participation variable? If not exploratory analysis were conducted, this should also be acknowledged (since the intention to do so was declared).   
  
A final point about the pre-reg - to the best of my knowledge, and I’m afraid I dn’t have a convenient citation - split-half datasets with exploration and confirmation are not demonstrations of validity, but only reliability. I hope this is kept in mind when the technique is employed. It is also the case that the use [of] machine learning techniques to squeeze even more predictive power out of our dataset and we will use machine learning to suck out even more exploratory predictive power does not constitue a pre-registered analysis (though of course, such details of training will be discussed if such analyses are reported elsewhere).   
  
I enjoyed this paper, and believe it makes a meaningful contribution to the literature. However, there are several issues the authors need to address in some depth before I can recommend publication.  Good luck. I am happy to read this paper again.   
  
  
Reviewer: 2  
  
Comments to the Author  
This is a theoretically elegant and exceptionally well written paper. This paper will be of interest to a wide range of social science scholars and is highly consequential for evolutionary theory of religion. A few minor questions and comments for further consideration.  
  
1) The supernatural belief scale used in this study represents only a time fraction of supernatural beliefs. It is probably fine for a Judeo-Christian belief in God scale, but I think it is worth commenting more about the limitations (from a generalizability perspective across faiths and cultures). I appreciate the care that went into the measures chosen (all theoretically relevant) but wish a more wide-ranging measure of supernatural belief was included (to better capture spiritual, non-religious folks).  
  
2) Belief in God (or gods) is a tiny fraction of the supernatural belief spectrum. What would you predict (and what are the implications for the theories tested in this paper) for spiritual/supernatural beliefs more broadly?  
  
3) The authors mention the need for non-WEIRD (non-US) data in passing, but I think this  is a major limitation of this dataset and should be discussed in the context of theoretical implications. In some ways the US is the least representative country in the world for studying atheism (which surely the authors would acknowledge).