First let me apologise for the long delay in providing this review. An eventful summer meant that I have had no time to work on it before mid-September. I have prioritised it on my return to work, but appreciate that this nevertheless has meant a considerable delay for which I am heartily sorry.

Considering the guidelines for evaluating Stage 2 manuscripts, I did not spot any substantial issues arising under 2A-2C. My comments focus mainly on the exploratory analyses and the Discussion (issues arising under 2D and 2E).

The researchers report an exploratory analysis (pp. 32 – 33) that is prompted by the fact that the mean for affective prejudice is nominally smaller in the threat-related self-affirmation condition than in the other conditions (contrary to hypothesis). Exploratory analysis shows that a contrast that compares this mean with the combined means for the other two conditions is statistically significant. (None of the other contrasts tested are statistically significant.) What this means, however, is hard to establish. It is always hard to know where the line is between being sensitive to a cue in the data about a possible effect that should be brought to the attention of researchers for further investigation and uncovering unstable “effects”. The fact that this effect is in the opposite direction to the one predicted strikes me as both good (it is not fishing to confirm the researchers’ hypotheses) and bad (how do we explain it?). The fact that the effect size is below their pre-specified criterion (p. 34) seems to me to be a factor against making too much of this finding.

Nevertheless, the researchers undertake further work (pp. 34-36) to tie this finding down by undertaking an exploratory content analysis of the affirmation texts. This reveals that participants in the secularism condition tended to explain its importance to them in national terms, whereas participants in the humour and physical endurance condition tended to explain its importance to them in individual terms.

The authors interpret this as indicating that “participants who self-affirmed on secularism have used this value as a shield to protect their national identity, while this is not the case in the other two conditions” (p. 35). They develop this argument both on p. 36 and in the Discussion (p. 37). While this is a possible interpretation, the problem is that the finding is necessarily derived from different values. The claim would be stronger if it could be shown that the different levels of abstraction do not emerge naturally from the values themselves. That is, it is possible that the values naturally lend themselves to thinking in more national or individual terms. It may be that this has important psychological consequences when it comes to self-affirmation, but this possibility needs to be entertained when evaluating the strong claim that participants “used secularism as a protective shield guaranteeing everything that the French Republic stands for, all of which are values that promote tolerance and living together in harmony” (p. 37). It is a subtle difference and I do not want to overstate the point, but wording such as “used … as a shield” implies to me some intent, whereas it may be that any benefits that accrue are incidental rather than intentional. It may even be that the difference is merely coincidental and immaterial to the findings. This possibility should at least be addressed in the Discussion, albeit while not over-elaborating a tentative finding that stems from a post-hoc exploratory analysis.

I also noted the following while reading the Discussion:

p. 36

The claim that “the results provide supporting evidence that the malleability of the

understanding of secularism as an ideology influences attitudes” is potentially misleading. These data are correlational and cannot provide evidence of causality. I appreciate that much can be read into the phrase “provide supporting evidence”, but my advice would be to avoid potential misunderstanding by using wording that avoided any implication of causality.

More generally, as someone who is not an expert on prejudice, I would like to have seen the authors spell out clearly what has been added by the principal findings concerning secularism and prejudice and discrimination. These comprise the positive findings from the study, given the lack of support for the hypothesised self-affirmation effects. Do they add to what we know and if so in what way? One reading of the section on pp 36-37 is that they duplicate previous work. I am not suggesting this is a bad thing – just I would like to know for sure whether they duplicate or add and, if the latter, in what way. I think it is there in the Discussion on p.36 and on pp. 38-39, but it could be spelled out for the reader so that they are clear about this contribution.

p. 39

It depends what you mean by “classical” self-affirmation, but in many self-affirmation studies participants are not asked to rank values and reflect on the highest ranked, but to rate them and reflect on the highest rated. Ranking is a much more onerous procedure. (See also p. 21 in the Method section.)

Although the value importance check did find humour to be more important than physical endurance, that isn’t really the issue being addressed here. In determining why the self-affirmation manipulation did not affect the outcomes the question is whether humour is sufficiently important to offset the threat to self-integrity. The risk of imposing a value is that it is not.

There are probably other limitations that could potentially have undermined the effectiveness of the self-affirmation manipulation. For instance, in reviewing the stage 1 submission in December 2023 I noted that the distributive matrices could be a problem: “I have some concerns about the extent to which the behaviour matrices will be sensitive to the self-affirmation manipulation. These matrices seem potentially quite complex and involved cognitively and I am not sure the manipulation, which is relatively subtle and potentially time limited, will be sufficiently robust to have an impact on the individual’s responses to it. I am not suggesting the researchers change this element of the design but they should bear this issue in mind when interpreting any null findings.” It seems to me at least worth considering this possibility.

Other issues

Abstract – the current draft has no conclusions.

Some suggested rewording:

Abstract

“for both affective (ηp2 =.004) and behavioral (ηp2 = .005) dimensions” > “for either affective … or behavioral …”

p. 14 The Present Research – should this now be in the past rather than the present tense?

Overall, the researchers seem to me to have done a good job of remaining faithful to the pre-registration in this draft and I commend them for the thoroughness their work and the clarity of their write-up.