

## **Reply to decision letter: RR #559**

Dear Recommender,

We would like to thank you and the reviewers for taking the time to read our manuscript and for your constructive and positive feedback.

We begin by summarizing the main changes made to the experimental procedure, then we provide detailed responses to each point below.

First, after asking participants to choose their representation of secularism, we included a continuous measure to find out the extent to which participants agree with the prohibition of religious practices and symbols in public spaces (the main difference between the new and the historical representations of secularism). This measure from 1 (not at all) to 7 (completely) will be used as a manipulation check for the representation of secularism.

We have also modified the manipulation of self-affirmation in line with the experts' proposals. We included a new experimental condition of self-affirmation on a value unrelated to the threatened domain (humor). We also modified our control condition: participants explain how the value of physical endurance can be important to someone else. Thus, the experimental design becomes 2 (secularism representation: historical vs new; between-subjects) by 3 (self-affirmation: threat-related value vs threat-unrelated value vs control; between-subjects) by 2 (religious group: Muslims vs Christians; within-subjects). We also modified the self-affirmation manipulation check by including a question about the importance of the values used in the procedure (secularism, humor and physical endurance) for participants. For example, we expect participants in the self-affirmation on the threat-related value (i.e., secularism) to consider the value of secularism as more important compared to those in the other two conditions (see more details in our response to point 2 of Expert #2).

Finally, we reconsidered our dependent variables. We use a feeling thermometer for the affective dimension of prejudice and distributive matrices for the behavioral dimension of prejudice, and we examine relative negative feelings and relative de-favoritism (see detailed description in response to point 6 of Expert #1). We argue that the cognitive dimension of

prejudice (stereotypes) would be too negative for Muslims and therefore difficult to compare between both target groups. That is why we do not include this dimension in our measures.

Please note that nothing has changed in terms of theoretical reasoning or hypotheses (except for the inclusion of the threat-unrelated value modality in the self-affirmation procedure). We would also like to point out that these are proposals made with the aim of improving the method in line with the suggestions of the experts, whom we thank very much. We remain open to further remarks and comments.

Below, we have copied your comments and those of the two reviewers and explained the changes that we have made to our manuscript. Please note that the recommender's and reviewers' comments are in normal script, while our answers are underneath in bold.

We hope to submit a stronger Stage 1 manuscript. We are ready to answer any other questions or comments if necessary.

All the best,

Yara, Constantina & Béatrice

**A track-changes comparison of the previous submission and the revised submission displayed side by side can be found on: <https://draftable.com/compare/FVVBoINotUig>**

**A track-changes manuscript is provided:**

**Manuscript\_RR\_Self\_Affirmation\_track-changes**

## **Response to Recommender: Anna Elisabeth Fürtjes**

Thank you for your submission to PCI RR. I have now received the evaluation of two expert reviewers which are largely positive. In addition to their comments below, I would like to draw your attention to the following:

1) In a Registered Report, the inferential chain for every analysis and potential conclusion must be nailed down precisely. Please include a row in the design table (page 5) for every planned analysis to be tested, including manipulation checks. The paper should mention no other analyses; exploratory analyses are not mentioned at this point, but can be analysed in their own results section at Stage 2. (The abstract will be based on only the planned analyses, and the planned analyses should be centre stage in the discussion.) Each row of the design table should precisely indicate which statistical analysis will test the claim in its respective row. Thus stating a multiway ANOVA will be done is too broad; it allows too much analytic flexibility. For example, if a claim is to be tested by an interaction, state specifically that the interaction will be used to test the claim. If simple effects are of interest to a certain claim, indicate this, stating which simple effect and what conclusion will follow from each one. In sum: The analyses and the inferential chain leading to a conclusion must be so well specified that anyone reading it is guaranteed to come to the same conclusion with the same data. Just as you need power for every row, you need to specify the claim tested in the final column for each row. State a proposition, i.e., the precise claim that could be found wrong by the test conducted in that row. Each row will test a slightly different claim, so have a separate claim for each row. Instead of giving a noun phrase like the "affirmation theory and its modifiability", spell out the full claim that might be found wrong.

**RE/ We have added the specific analysis plan, power analysis, interpretation, and theory impact for each row as suggested (pages 5-8)**

2) A power analysis is needed for every test in the design table. This is often usefully done by considering the raw effect size for each test considered. Remember power is the way to control Type II errors, namely the long term rate of missing effects that would be relevant to the theoretical claim tested. To control that error rate, therefore, power should be calculated with the minimal effect just of theoretical interest. This should be scientifically justified for each test. See <https://doi.org/10.1525/collabra.28202> for how to approach this thorny problem.

**RE/ Power analysis has been added to each row in the design table.**

3) I appreciate the efforts you have invested to simulate study results at this stage. As the simulated results do not indicate any significant effects, you have only reported two main effects (secularism <-> prejudice & prejudice <-> religious group), but not the results of the interaction effect which would answer the central hypothesis of this project, testing whether the effect of self-affirmation on prejudice is moderated by an individual's representation of secularism. This exclusion implies that the key hypothesis can only be tested and interpreted in the presence of these main effects which may not be trivial. It would be good to make your reason for this explicit by including a short explanation of this in the abstract (1 sentence).

**RE/ Thank you for your observation. We have now added the missing part about the interaction effect, in the abstract and in the results section.**

## **Response to Reviewer #1 : Sauro Civitillo**

Thank you for the opportunity to read and review the Registered Report entitled 'Self-affirmation and negative attitudes towards minority groups: The role of ideological malleability.' The main goal of this study is to test the effectiveness of a brief self-affirmation in mitigating affective prejudice towards Muslims in France. Additionally, the study aims to assess the moderating role of secularism malleability. I appreciate the Authors' efforts to make the study as transparent as possible, including the results of a simulation study. I have a few comments that could help to strengthen their work.

1) The Authors pointed out that ideologies (for example secularism and colorblindness) can be malleable. When reading the example of colorblindness, it came to my mind that this ideology can be applied in different domains. Individuals may be strongly endorse colorblindness in one domain (e.g., education) but be less in other spheres such as religion. Just like colorblindness, also secularism could be domain-specific, but not necessarily malleable. I think this aspect (malleable vs domain-specificity) could be explained in the introduction because as we know from the literature on intergroup ideologies, ideologies tend to stable and difficult to change.

**RE/ Thank you for your comment.**

**We specify now that ideological malleability implies a motivational component. If individuals are motivated to protect the status quo, they may interpret an available ideology in a way that reinforces intergroup hierarchy, even if that ideology at its core is constructed to challenge social inequality (Knowles et al., 2009). This can be accomplished through two complementary ways: (a) the change in the intensity of adherence to the ideology and (b) the modification of the ideology's content. Individuals can show a greater adherence to the ideology in certain specific domains. However, this is possible because they modify the content of the ideology, its cognitive representation, to make it more consistent with the maintenance of social hierarchy. Taking the example of the colorblindness ideology, Knowles and collaborators (Knowles et al., 2009) have shown how this principle that individuals should not be judged by the color of their skin, can be variably used to reduce versus to maintain the status quo of racial inequality in the United States. Specifically, Knowles and colleagues (2009) identified two meanings of the colorblindness ideology: a distributive-justice versus a procedural-justice dictates. Distributive justice refers to the equal treatment of people in resource allocation, regardless of the color of their skin. People with an egalitarian view of intergroup relations agree more with this principle of distributive justice than people with an anti-egalitarian, hierarchical view of social groups. However, to achieve equal treatment of Black people and White people, Martin Luther King's dream, it might sometimes be necessary to consider people's skin color through positive actions that help to equalize opportunities and reduce social hierarchy. Anti-egalitarian people can, in these circumstances, claim to be attached to the application of the colorblindness principle in procedures such as employment or university admission. Thus, individuals motivated by the protection of racial hierarchy interpret the ideology of colorblindness as a procedural justice dictate. Studies by Knowles and colleagues (2009) showed that, when faced with an intergroup threat, anti-egalitarian participants' support for the ideology of colorblindness increased compared to a non-threat condition. However, this is explained by the fact that anti-egalitarian participants shifted their representation of colorblindness from an ideology of distributive justice to one of procedural justice.**

2) From reading the title of the manuscript, and throughout the introduction, I thought that the focus would be on attitudes towards minority groups (DV). However, as reported on page 18, the DV refers to affective prejudice towards two religious groups (Muslims and Christians) with prejudice towards Muslim being the main focus. Thus, I would encourage the Authors to

be consistent throughout the manuscript and change the wording accordingly, e.g., affective prejudice towards Muslims.

**RE/ We agree with your remark. We propose to change the measure of intergroup emotions and to use a more general measure of affective prejudice: the feeling thermometer. We also propose to measure behavioral prejudice using matrices inspired from Rubin et al. (2014) and Anier et al. (2018). Using these measures, we propose to examine relative negative feeling and relative de-favoritism, comparing the target group (Muslims or Christians) to the ingroup (French people) - see also the detailed presentation in the response to point 6. We also changed the wording throughout the manuscript to prejudice against religious groups.**

3) Could the Authors clarify if they will be recruiting participants regardless of the religious beliefs, or specifically targeting non-Muslim and non-Christian French participants? This important information is not reported in the participants section.

**RE/ We aim to recruit participants regardless of their religious beliefs and we report this information now in the participants section (p.19). We examine whether participants' religiosity impacts the effects of self-affirmation and of secularism representation on prejudice (see multiple regression analysis in the part "Analysis Including Participants' Religiosity").**

4) It is stated that the calculation of the required sample size is based on an unpublished study (Nugier et al., 2023). I wonder if the Authors could explain why they chose this study. Also, I wonder if the expected effect size could be more in line with the literature on self-affirmation and prejudice reduction interventions.

**RE/ We based our power calculation on a study recently carried out in France, with an experimental design of 3 (Affirmation procedure: self-affirmation, group-affirmation, control) x 2 (salient social norm: new secularism vs historical secularism), between subjects. We observed a significant and small interaction between variables,  $f = .10$ . Participants in the self-affirmation condition displayed higher prejudice compared to control, when the norm of new secularism was salient.**

**To our knowledge, this is the study that comes closest to our experimental design. However, this study manipulates the perception of secularism norms (new vs.**

**historical), and does not examine participants' personal adherence to these norms, nor their own representation of secularism. We decided to take a cautious approach in which we need to have a sufficiently powerful sample to detect small effects.**

5) I'm not quite sure about the manipulation check for the secularism display. If I understand correctly, two visions of secularism are presented, and then participants are presented with four items about the rights of religious groups. I wonder if filling out a scale about the rights of religious groups might already serve as a way of expressing prejudice against religious groups, in sum, as a manipulation, but not as a manipulation check. I would like to hear the Authors' thoughts on this concern.

**RE/ Thank you for this important remark. We propose to change the manipulation check and to use a continuous measure assessing the agreement with the following affirmation: “Religious practices and symbols should not be allowed in public institutions”. Participants in the new secularism condition are expected to display a higher score on this question than participants in the historical secularism condition. Indeed, the main difference between “historical” and “new” representations of secularism is related to the prohibition of displaying religious symbols or practices in the public space.**

6) I have also some concerns regarding the measure of affective prejudice towards Muslims. The plan is to use a differential measurement approach by comparing emotions towards (only) two groups, Muslims and Christians. Such differential prejudice measure may mischaracterize in-group favoritism as out-group bias. Some people may consider outgroup members as less good but may not necessarily harbor hostility against them. On a related note, I would encourage to make a stronger argument for why affective prejudice is the only DV of interest, and not for example cognitive prejudice (stereotypes) or behavioral prejudice (discrimination), or all three dimensions of prejudice.

**RE/ We have decided to follow your advice and adjust our DV. First, we propose to replace the measure of intergroup emotions with a more general measure of affective prejudice : a feeling thermometer. This would allow us to add several other groups, including “French people”, as a distraction to the target groups. Moreover, we propose to consider relative negative feelings, comparing the target group (Muslims or**

Christians) to the ingroup (French people), instead of the simple feeling towards each group.

For the feeling thermometer, we propose to calculate two scores of relative prejudice (see also Hugenberg & Bodenhausen, 2003). This can lower the sensitivity of our measures to social desirability concerns :

**Affective prejudice against Muslims = feeling towards French - feeling towards Muslims. Higher score means higher relative prejudice towards Muslims**

**Affective prejudice against Christians = feeling towards French - feeling towards Christians. Higher score means higher relative prejudice towards Christians**

In addition, we propose to add a measure of behavioral prejudice using distributive matrices (Rubin et al., 2014; Anier et al., 2018). Participants will be asked to imagine that they are responsible for allocating subsidies at Paris City Hall, and that they have to indicate their intention to divide a sum of money between two high impact associations that need support. Four fictitious student associations will be used. One matrix will include a Muslim vs. a French association. The other matrix will oppose a Christian vs. another French association. These matrices will allow us to measure the relative de-favoritism of Muslims and that of Christians in comparison with French people.

Then, we will calculate two scores of behavioral prejudice:

**Behavioral prejudice against Muslims = money attributed to the French association (Students of France) - money attributed to the Muslim association (the Muslim students' association). Higher score means higher de-favouritism towards Muslims**

**Behavioral prejudice against Christians = money attributed to the French association (French students united) - money attributed to the Christian association (the Christian students' association). Higher score means higher de-favouritism towards Christians**

We did not include the cognitive dimension of prejudice (stereotypes) because the stereotypes about Christians (focused on lower scientific competence, e.g. Riois et al., 2015) are different in content compared to stereotypes about Muslims (focused on terrorism, e.g. Saleem & Anderson, 2013), see also Erentzen et al. 2022. In contrast,



**examining affective and behavioral prejudices allows the use of the same measures for both groups.**

7) A final note on causal language. Although the present study proposed to carry on an experiment, I would be careful to use impact, cause, etc.

**RE/ Thank you for this comment. We carefully reformulated the sentences using the terms “impact”, “cause”, etc.**

## **Response to Reviewer #2 : Pete Harris**

The paper presents an interesting study that will attempt to test a potential moderator of the effect of self-affirmation manipulations on prejudice reduction. The goal is to understand more about when self-affirmation can be expected to be effective as a means of reducing prejudice. The moderator is “ideological malleability”, operationalised as the participant’s representation of secularism.

My expertise is in self-affirmation and in issues such as the uptake of health-risk messages, rather than the reduction of prejudice. I therefore have some hesitation in evaluating the theoretical origins and operationalisation of the proposed moderator. Nevertheless, I felt the issue was set up well and clearly in the introduction and gave a persuasive account of the underlying rationale for choosing this moderator.

As I describe below, I had several queries and issues as I read through the draft manuscript. I must stress that these are not criticisms but as things to think about as the study is developed. It may well be that these issues have already been considered and the researchers are happy enough with their existing decisions. Indeed, several of the issues that bothered me while reading the paper I found were eventually dealt with somewhere later on. If this paper proceeds to stage 2 it would be useful if the authors indicate earlier in the paper that they are aware of these issues and will deal with them in due course. This includes the manipulation checks, especially for value importance in the self-affirmation condition, the impact of religiosity and own religion on the findings, and the risk of systematic effects involving exclusions (in some ways addressed by the fact that analyses will be reported both with and without exclusions).

Issues

**1) Interaction hypothesis for self-affirmation: impact on those holding a “new” representation of secularism.**

As I say, my expertise in self-affirmation is in issues such as the uptake of health-risk messages, rather than the reduction of prejudice. Perhaps as a result, I was not clear about the reasoning and evidence underlying the prediction that self-affirmation would increase prejudice among those with a “new” representation of secularism. This is asserted, rather than clearly explained, on p.10 and p. 11. I felt this prediction would benefit from more explicit derivation and explanation.

**RE/ We have adjusted this part of the introduction (p. 12 and 13) to explain more clearly the derivation of this interaction hypothesis.**

**We argue that self-affirmation on secularism will increase its importance for participants and their own support for it\*. Previous research shows that support for “historical” secularism is associated with lower prejudice against Muslim immigrants, while support for “new” secularism is correlated with higher prejudice (Roebroek & Guimod, 2016). Consequently, self-affirmation using the secularism value can increase negative intergroup attitudes among participants with a new representation of secularism.**

**\*The inclusion of a new manipulation check for the self-affirmation manipulation (i.e. the measure of the importance accorded to the secularism value immediately after the self-affirmation task), will allow us to test this interpretation.**

**In line with our reasoning, other research using group-affirmation shows an increase of negative intergroup attitudes following the affirmation procedure when a discriminatory norm is salient. The argument was that group affirmation increases the salience of social identity used in the affirmation procedure, and the content of this collective identity including social norms. If social norms are discriminatory, group-affirmation increases the conformity to these norms, and consequently prejudice against minority groups (Badea et al., 2021).**

**2) Manipulation of self-affirmation.**

A central tenet of self-affirmation theory is that the self-attribute that forms the basis for

affirming the self (e.g., the affirmed value) must be at least as important to self-integrity as the threat is detrimental to it. One issue that confronts anyone who constrains the individual's choice of attribute – as is the case here with the focus on affirming secularism – is whether that meets this requirement. A second consideration is that there can be specific consequences when a chosen value is related, rather than unrelated, to the self-integrity threat, again as is the case here. At the very least, these considerations need to be borne in mind when interpreting the findings.

The wording of the self-affirmation instructions assumes that secularism is an important value (“Please explain why this vision of secularism is important for you personally”, p. 42) and it would be useful to be able to assess whether, and indeed how, important it is. In many studies researchers place a brief 7-point rating of value importance immediately after the manipulation, which can be used as a manipulation check, compared between conditions (not possible here, given the control condition), or used as a covariate or moderator in analyses.

Instead, the researchers opt to undertake a content analysis of the affirmation and code the responses (p. 45). (As I say above, it would be useful to mention this when describing the affirmation procedure in the paper at p. 17.) This is a good idea, but a lot will depend on what can be gleaned from the statements.

The researchers currently envisage using this to exclude participants. However, only those who receive a rating of 1 will be excluded (p. 45); yet arguably only values 4 and 5 of their coding scheme indicate the value is important to the individual. A rating of 3 (general justification) does not seem to me to necessarily indicate that the value is important personally.

At the risk of expanding the exploratory analyses, I wondered if the researchers had also considered exploratory analyses in which this rating is included as a covariate or even a moderator (split at 3 or lower v 4 and 5)?

Also, with this, as with all exclusions, it will be important to test whether these are systematically related to the conditions. However, analyses both with and without exclusions (p. 45) will help with that issue; again, however, it would be useful to mention this earlier in

the paper.

**RE/ We agree with your comments and we propose to ameliorate the self-affirmation procedure. We first propose to adjust the control condition to make it more equivalent to the experimental condition. Similarly to Lehmiller et al. (2010), we will ask participants to explain why another irrelevant value, physical endurance, would be important for another person, and to give an example.**

**Second, we thank you for the point you raised about the affirmed value being related or unrelated to the threat domain. We thought about that and included a new experimental self-affirmation condition on a value unrelated to the threatened domain (humor). The experimental design becomes 2 (secularism representation: historical vs new; between-subjects) by 3 (self-affirmation: threat-related value vs threat-unrelated value vs control; between-subjects) by 2 (religious group: Muslims vs Christians; within-subjects).**

**Then, we followed your advice and replaced our manipulation check for the self-affirmation procedure. Specifically, we included a question concerning the importance of the values used in the procedure (secularism, humor and physical endurance) for participants, immediately after the self-affirmation manipulation. Participants in all conditions will answer the question for each of the three values presented in a randomized order from 1 to 7 (1 = not at all important, 7 = very important): “ Before we move to the next part of the survey, we would like to know how important are some values to you:**

- secularism**
- humor**
- physical endurance”.**

**We expect participants who self-affirm on the threat-related value (i.e., secularism) to consider the value of secularism as more important compared to those in the other conditions. We expect participants who self-affirm on the threat-unrelated value (i.e., humor) to consider the value of humor as more important compared to those in the other conditions. Finally, we do not expect significant differences in the importance of physical endurance between conditions.**

### **3) No value control group.**

I appreciate the difficulty the researchers have in finding an appropriate control, given the experimental condition focusses on secularism. However, the chosen control inevitably means that the conditions differ in both self-affirmation and value salience, which cannot be disentangled by this design. This will need to be borne in mind when interpreting the findings.

More mundanely, looking at the materials, the researchers do not provide the participants with any cover story (rationale) for such a task, which may therefore seem odd to the participant. Is there anything they can do to smooth this element of the procedure (e.g., by way of a cover story that enables the task to make some sense without undermining the experimental design)?

**RE/ We proposed to change the control task as mentioned above. Additionally, we made the secularism value salient in all conditions, by reminding participants of their chosen representation of secularism and by asking about the importance of this value to them (see point 2). This way, the experimental and the control conditions will only differ in self-affirmation but not in value salience.**

**We also added a rationale for the self-affirmation on the threat-unrelated value and for the control task. Please see below the detailed procedure for the control condition and for the self-affirmation on the threat-unrelated value:**

## **SELF\_AFFIRMATION**

*Participants are randomly assigned to one of the three following conditions (self-affirmation on threat-related value vs. self-affirmation on threat-unrelated value vs. control)*

### ***Condition 1: Self-affirmation on threat-related value***

You have chosen the following vision of secularism as being closest to your personal convictions : [*The vision that corresponds to the participant's choice is displayed below*]

-----

Please explain why this vision of secularism would be an important value for you personally.

\_\_\_\_\_

Give an example where this value has guided your behavior (for instance, in your daily life, in your interactions with others...).

\_\_\_\_\_

Before we move to the next part of the survey, we would like to know how important are some values to you on a scale from 1 to 7 (1 = not at all important, 7 = very important).

*[The presentation order of items is randomized]*

- Secularism
- Physical endurance
- Humor

***Condition 2: Self-affirmation on threat-unrelated value***

You have chosen the following vision of secularism as being closest to your personal convictions : *[The vision that corresponds to the participant's choice is displayed below]*

-----

Now we would like to focus on another aspect of life, this time social, that is humor.

Please explain why humor would be an important value for you personally

\_\_\_\_\_

Give an example where this value has guided your behavior (for instance, in your daily life, in your interactions with others...).

\_\_\_\_\_

Before we move to the next part of the survey, we would like to know how important are some values to you on a scale from 1 to 7 (1 = not at all important, 7 = very important).

*[The presentation order of items is randomized]*

- Secularism
- Physical endurance
- Humor

***Condition 3: Control***

You have chosen the following vision of secularism as being closest to your personal convictions : *[The vision that corresponds to the participant's choice is displayed below]*

-----

Now we would like to focus on another aspect of life, this time corporal, that is physical endurance.

Please explain why physical endurance would be an important value for another person (not for you personally)

---

Give an example where this value could guide the behavior of this person (for instance, in their daily life, in their interactions with others...).

---

Before we move to the next part of the survey, we would like to know how important are some values to you on a scale from 1 to 7 (1 = not at all important, 7 = very important).

*[The presentation order of items is randomized]*

- Secularism
- Physical endurance
- Humor

#### **4) Sensitivity of the dependent measure**

As I say, I am not an expert in prejudice, but I did wonder about the sensitivity of the dependent measure (6 items assessing emotional reactions). Clearly this measure has been used successfully in previous research, so my concerns may be misplaced, but the study has a fairly transparent design (questions about secularism > a manipulation > measures of emotional reactions to Muslims and Christians) and I wondered about the potential impact of



socially desirable responding.

**RE/ We have adjusted our dependent variables to make them less sensitive to social desirability concerns (see our detailed response to point 6 of Expert #1)**

### **5) Own religion and religiosity**

Again, like the manipulation check for self-affirmation, this issue is discussed rather late in the paper (pp. 24-25). It is highly likely, isn't it, that one's own religion will heavily influence responses to the dependent measure, to the potential detriment of the between-subjects manipulations? In the current draft this issue is touched on only briefly and again rather late on.

The proposed remedy is to run exploratory analyses with participants' self-rated religiosity as a covariate (p. 25). I wondered if the researchers had considered also, or instead, running analyses controlling for, or assessing the impact of, own religion (especially Muslim or Christian) as this may well be a powerful influence on responding? I appreciate this is not straightforward, but it does seem important.

Likewise, I wondered if own religion will be correlated with secularism (and therefore confounded with it)? It is not hard to imagine that people who practice some form of religion, perhaps particularly Muslims, may hold a more historical perspective.

Finally on this issue, I wondered why the researchers differentiated the type of Christian (Catholic, Protestant) but not any other religion, including (and especially) Muslim? (See comments on wording below.)

**RE/ We agree that religious belief can have a great influence on responses. We had mentioned the measurement of participants' own religiosity in the supplementary materials and in the exploratory analyses, but for greater clarity we have now included this measure directly in the method section. We will measure the extent to which participants practice a religion and control for this variable in a multiple regression. As suggested, we will also investigate associations between reported religion and the chosen secularism representation.**

**We will not differentiate the subcategories of Christianity nor of any other religion.**

#### **6) Attention checks**

I have mixed feelings about attention checks. On the one hand, it is obviously good to detect inattentive responding, but on the other it could risk the questionnaire coming across as somewhat bizarre and potentially undermine engagement. On that front, it was good to see there was advanced warning in the introduction about the use of attention checks.

In this I think balance is key and I think I probably misunderstood the statement about this on p 18 – that there would be two attention items per prejudice measurement – to mean that each participant would see four attention check items. If I did misread it, it would be good to make clear exactly how many each participant saw. If there are four attention check items, this seems too many given there are only six emotion items.

**RE/ Participants will see one attention check item in the feeling thermometer to detect careless responding: “Please, put the cursor on ten”.**

#### **7) Data quality**

Several of the issues I have raised have implications for data quality, so it was good to see the researchers had comprehensive diagnostics in place to assess data quality (p. 25). It was, however, not always self-evident what each element was controlling for. Perhaps this section could be reworded to assist the reader in this regard. For example, “To control for problem X, we used options A, B and C ...”

As I say, a related issue concerns systematic effects of exclusions that might undermine the randomisation, so it would be good to have an explicit description of how the researchers will assess this or control for it.

**RE/ We have adjusted this section for greater clarity. We modified our manipulation check questions. This allows us to include all participants. The exclusion criteria are failing the attention check item and reporting careless responding (less than 4 out of 5 when asked if they answered the questionnaire seriously). We will also run analyses with and without exclusions.**

## 8) Power analysis

The power analysis is explicitly described on p 13 and p31. I was not sure, however, whether the fact that the analyses were derived from an unpublished study raised issues.

**RE/ To our knowledge, this is the study that comes closest to our experimental design. It was recently carried out in France, with an experimental design of 3 (Affirmation procedure: self-affirmation, group- affirmation , control) x 2 ( salient social norm: new secularism vs historical secularism), between subjects. (Please see our detailed response to point 4 of Reviewer #1)**

## 9) Analysis strategy

I was puzzled as to why there was a separate ANOVA to test the main interaction hypothesis concerning self-affirmation and secularism (p.19). Couldn't this be picked up by tests of simple effects at the level of muslims following the test of this 2-way interaction as part of the three-way ANOVA? I appreciate the researchers are closer to this issue than am I and may have spotted a complication that I am missing.

**RE/ We agree and will run a 2 (secularism representation: historical vs new; between-subjects) by 3 (self-affirmation: threat-related value vs threat-unrelated value vs control; between-subjects) by 2 (religious group: Muslims vs Christians; within-subjects) ANOVA. We will examine the following simple effects, in each condition of secularism representation : self-affirmation on the value related to the threatened domain vs. control; self-affirmation on the value unrelated to the threatened domain vs. control. This is now adjusted in the results section.**

## 10) Pre-test

In principle, pre-tests are a good idea. The draft is very clear about what will be done and the fact that the data from the 100 participants will not be analysed at the time of the pre-test but will nevertheless form part of the final sample considered for analysis (p. 15). However, I wondered whether there was a Plan B. What will the researchers do if they do not meet the requirements of the pre-test they set out on p. 14?

**RE/ Having modified our experimental procedure as well as our recruitment method, the pretest is now irrelevant. In consequence, we choose to not do it. Our new recruitment method is through a survey institution, which allows us to have a representative sample of the French population, and guarantees that all participants complete the survey to the end. We were also asked by the survey institution to remove early exclusion questions, thus the option of choosing neither of the secularism representations is now unavailable. However, the continuous item used as the secularism representation check allows for a more nuanced answer. We have thus minimized the risks of participant exclusion, and we have waived the pre-test.**

### **11) Manipulation checks**

There are manipulation checks for secularism (measuring opposition to religious groups' rights (p. 17) and for self-affirmation (a content analysis of the writings, mentioned in Table 1 but not in the main text).

**RE/ We added a description of both new checks (see the detailed description above, in response to your point 2 and to point 5 of Reviewer #1)**

**12) The materials. When I looked at the materials themselves, I had some wording issues:**

#### **12a) *Definition of new secularism***

I understand that the definitions employed here are based on previous studies (pp. 16-17), so there may be nothing the researchers feel they can do or wish to do about this, but as an old experimenter it did strike me that the descriptions were less equivalent than they might be. In particular, the wording of the new secularism condition, at least in its English version, came across as somewhat more strident than the historical one. It could easily be reworded so that the wording and structure of these two definitions more closely mirrored each other. For example, in English:

“Individuals are free to practice their religion in private but not in public. Citizens do not have the right to show their religious affiliation in the public sphere. The state is not totally neutral with regard to/must regulate these religious practices.”

**RE/ As suggested, we have modified the definition to make it more equivalent to the historical secularism definition: “Individuals are free to practice their religion in private but not in public. Citizens do not have the right to show their religious affiliation in the public sphere. The state must regulate these religious practices.” (p.21).**

### **12b) Wording of the experimental question**

I also wondered whether the wording of the experimental question could be reworded to avoid begging the question concerning whether this is an important personal value. As I say above, the current wording (“Please explain why this vision of secularism is important for you personally”, p. 42) assumes it is personally important. A more neutral question could ask:

“Is this vision of secularism important to you?” “If so, please explain why.”

Indeed, if there was a yes/no/uncertain answer available to the first question, this could be used as an index of value importance and manipulation check for comparison in the two conditions (in addition to the codings).

Or a self-rating could be used: “How important to you is this version of secularism” 7 points, not at all/extremely.

**RE/ As stated above, we added a question about the importance of the value of secularism to participants in all conditions. In addition, the wording of the experimental question was changed to “Please explain why this vision of secularism would be important for you personally”.**

### **13) Funneling and debriefing**

I wondered whether “How seriously have you answered this questionnaire?” might be less loaded than “Have you answered this questionnaire seriously?” (p. 43). Again, it might be something that is more of an issue in English than French.

**RE/ We believe that, as raised above, it is more of an issue in English than in French, but we modified the question to “How seriously have you answered this questionnaire?”**

#### 14) Own religion

Was there some reason for differentiating the type of Christian (Catholic, Protestant) but not any other religion, including (and especially) Muslim?

Would it make more sense for the general question “To what extent do you practice a religion” to precede rather than follow the more specific question about which religion or beliefs you feel closest to?

**RE/ We will measure the extent to which participants practice a religion before asking about the specific religious belief. Additionally, we will not differentiate the subcategories of any religion.**

#### 15) Debrief

Will participants know what self-affirmation means in this context?

**RE/ We have added a definition of self-affirmation in the debriefing: “Self-affirmation techniques help individuals reflect on values they consider significant in their life. Studies show that this simple activity can help people feel better about themselves and have a higher self-esteem, which makes them exhibit fewer negative attitudes towards others”.**

#### 16) Some items by page number

16.1) P 17 I am not sure the fourth item, about French TV, is tapping into the same issue as the other three items. I could imagine holding an historical view of secularism but having no desire to see TV programmes made by or for religious people!

**RE/ We agree with this potential difference and have replaced the manipulation check for secularism’s representation with one item (see our response to the point 5 of Reviewer #1)**

16.2) P 22 I was puzzled by the use of post-hoc tests when testing hypotheses (e.g., p. 22). Again, I may be missing some important subtlety here.

**RE/ We will instead use t-tests to investigate the specific simple effects to test specific hypotheses.**