**Peer Review and Communication History**

**MS Title**: Hypothesis Testing Preferences in Research Decision Making

**Author Names**: Stephanie M. Anglin, Caitlin Drummond Otten, Stephen Broomell

**Submitted:** Jul 22, 2022

**Editor First Decision**: Revise & Resubmit

Nov 2, 2022

Dear Anglin,

I have now received 2 reviews of your manuscript, “Motivated Reasoning in Research Decision Making”, from researchers with special expertise in judgment and decision making. I also independently read the manuscript before consulting these reviews. The reviewers had mostly positive reactions to your manuscript. I agree that your manuscript has important strengths and also that there are some issues that need to be addressed. I therefore encourage you to submit a revised version for further consideration at Collabra: Psychology.

The reviewers did an outstanding job in their reviews. I will highlight issues I think are particularly salient here. In your resubmission, please include a document with a point-by-point response to both the points I list here and the reviewers’ comments, outlining each change made in your manuscript or providing a suitable rebuttal.

Please pay special attention to issues of clarity, highlighted by both reviewers, and testing the simple effects.

In summary, I think this is a promising manuscript and, I hope you will revise it for further consideration at Collabra: Psychology. I look forward to receiving your revision. Please see the instructions below for submitting your revision.

Please ensure that your revised files adhere to our author guidelines, and that the files are fully copyedited/proofed prior to upload. Please also ensure that all copyright permissions have been obtained. This may be the last opportunity for major editing, therefore please fully check your file prior to re-submission.

If you have any questions or difficulties during this process, please contact the editorial office at [editorialoffice@collabra.org](mailto:editorialoffice@collabra.org).

We hope you can submit your revision within the next six weeks. If you cannot make this deadline, please let us know as early as possible.

Sincerely,

John Flournoy

# Reviewer 1

##### Rating scale questions

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  | ✔ |  |

##### Open response questions

### Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

I would like to thank Editor and the Authors for the opportunity to read this manuscript, it was quite informative and interesting. Below I list several issues I think might help to improve the readability and comprehensibility of the manuscript.

1. Some of the information in the Introductory parts are repetitive; and conversely, some of the concepts are not described in sufficient details when mentioned for the first time. For example, there is a good distinction between hot (motivated reasoning) and cold (cognitive biases) cognitive mechanisms (p. 4), but I miss more elaborate explanation how it relates to the research conducted. Authors mention myside bias and motivated reasoning interchangeably (are they treated as synonyms?) and they do not specify how their design aims to measure motivated reasoning, i.e. distinguish between hot and cold cognitive mechanisms in deciding about which study should be performed. This needs to be highlighted more in the Introduction.
2. I would recommend highlighting the novelty of this research, i.e. that in previous studies participants reasoned about “results” consistent or inconsistent with their beliefs, while this study examine the support for “hypotheses” consistent with their beliefs. Also, at least short review about previous findings on how framing of hypotheses might affect the choice of participants could be mentioned.
3. On the other hand, section “Confirmation and Disconfirmation Bias” does not seem to follow logically after previous section, where RQ1 was specified, as it seemed to elaborate more on rationale behind this research question. I would recommend to streamline introduction more, so that information provided is not repetitive (description of the current research is on p. 4 as well as p. 10, etc.).
4. I also did not quite understand why authors posed research questions, although they provide a rationale for hypotheses and make predictions. If they choose to go with hypotheses, I would recommend building stronger case for why they expected reversal of preferences for conflicting results. If the participants use positive testing strategy, it still makes sense to repeat what they believe to be true. If the participants see both studies as equivalent, then they should have no preference regardless of their beliefs. And even if they are motivated to see particular result, it seems that they could believe that repeating belief-consistent study would yield stronger evidence for their case.
5. In contrast, it seemed to me that review of literature regarding individual differences allowed for more specific predictions than posed by RQ3. Authors could reasonably expect no preference for participants with better scientific reasoning skills (p. 10).
6. As mentioned in point 1, it is not clear whether Authors mean the same thing when they talk about motivated reasoning and myside bias. I recommend clarifying and uniting the terminology. Moreover, it is not clear whether myside bias is one of the variables and how it is operationalized (or what will be counted as motivated reasoning).
7. I appreciate analyzing also qualitative responses and I think it deserves highlighting in the Introduction as one of the contributions of the current research.
8. In the Methods section, I would recommend explicitly stating what was your main dependant variable.
9. I did not understand why results for some of the individual measures are reported in the main text and some in supplementary materials. Could you explain or include them all in the main results?
10. If I understand it correctly, three questions were used to measure participants’ beliefs (related to politicized topic of gun control): political orientation (liberal vs. conservative), belief about efficacy of gun control policy and position about gun control, but only two of them were used in analyses (political orientation and gun control belief). Why? Or if they are highly correlated, maybe you could use composite measure?
11. Table 1 – I found the title a little confusing – what does a “forced choice” mean? In materials Authors describe that participants could expressed “no preference”.
12. P. 17: “responses to each question were coded…” – please, specify to what question. Similarly, on p. 18: “The condition x gun control belief interaction became stronger” – please, specify which condition. Also, p. 23: “responses were coded so that higher scores indicate…” – please, specify responses to what.
13. Moreover, I did not understand how were responses coded so they were belief-consistent to both liberals and conservatives – can you specify the procedure? I also wondered if the chi-square or t-test were not more appropriate to see the differences between liberals and conservatives and their choice and differences in scores.
14. Table 4 – to be consistent with your terminology, Scenario should be renamed as Study.
15. Discussion to Experiment 1 – Do you consider your results to provide support for PTS explanation rather than motivated reasoning? (Because the results were similar for politicized as well as control study?) You found different rationales for participants´ choices – some of them more in line with motivated reasoning account, some of them more in line with PTS. What are the implications of these findings? Can you elaborate a bit more?
16. Discussion to Experiment 2 – Authors wrote (p. 34-35): “people have a tendency to favor research with belief-consistent hypotheses and do not significantly alter their decision-making strategy from before the outcomes are known to after competing outcomes are obtained. These preferences do not appear to be driven by the assumption that the researchers are manipulating the outcome, as participants showed the same belief-consistent preference when an independent research group conducted the study.” Do these results imply that people hold their opinions regardless of the evidence? And again, because there was no difference whether researchers or independent research group conducted the study – can it be concluded that people rely on PTS?
17. It would be interesting to see the results for ambiguous results (after the research in the scenario is conducted, but inconclusive results are found) or with different polarizing topic. (This is just an idea; you do not need to respond to this.)
18. Lastly, I found the last sentence vague and generic. The research reported in the manuscript is interesting, and I believe that the paper should end with stronger and more specific message than generic conclusion that more research is needed.  
    In general, I enjoyed reading the paper, I found the methodology and analysis adequate and I highly appreciate pre-registration and sharing the materials, which help to understand some details better.

# Reviewer 2

##### Rating scale questions

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  | ✔ |  |  |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  | ✔ |  |  |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  | ✔ |  |  |

##### Open response questions

### Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

The current manuscript tests several hypotheses related to the way motivated reasoning impacts preference for scientific methods and findings. For example, the author(s) first research question is “Do people exhibit a preference for which of two scientifically equivalent studies to conduct based on the framing of the hypotheses?” Overall, I think that this research is interesting and could be very important to understand the conditions that lead to the funding (or not) of important scientific work. My major comments about the paper have to do with structure and clarity both in terms of generating the hypotheses based on theory and clearly and cleanly reporting the results.

Big Picture Comments

Overall, I think the authors are tackling an interesting and important set of questions. They have several research questions, and I did find the overall theoretical perspective/rationale being used to derive the predictions to be a bit difficult to ascertain. The author(s) provide a thorough review of literatures relating to confirmation bias, motivated reasoning, etc., but there does not seem to be a clear thread throughout that informs the hypotheses and contextualizes them in previous theory. Furthermore, it did not seem to me that the hypotheses are necessarily related and each one could almost be fleshed out to serve as a standalone paper. I think it would be helpful to establish a central perspective that justifies and links together the study hypotheses. One way to do this might be to lay out the broad literatures of interest, describe how they may relate to each of the three research questions, and then have subsections for each of the three hypotheses.

It also did not seem to me that the hypotheses were directional in nature. For example, the hypothesis given above in this review does not seem to establish an expected direction of results but is a general question. Directional hypotheses might help to clarify what it is that the researchers expected. RQ3 is similar and does not feel strongly grounded in specific previous work. Generally, I think the authors test three interested and related questions but it takes a lot of bandwidth to derive which perspectives inform which hypotheses, how the methods test the hypotheses, and how the findings map onto the predictions.

In terms of the general discussion, there is quite a lot of speculation about why the findings differed from expectations starting on page 37 of the manuscript. I think it’s great to report unexpected findings, especially if they replicate across experiments. I also found this section a bit difficult to parse given that the predictions had seemed unclear to me from the outset. I think a grounding of all three hypotheses in the literature, directional statements, and tie-ins to which patterns map onto which expectations in the results would help with this.

Comments on the Studies

Study 1

In Study 1 skin cream being ineffective is used as a control condition (i.e. a non-polarizing topic). However, I do not see a measure of people’s agreement with skin-cream effectiveness like I do for agreement with gun control. How do we know these two conditions differ in the degree of polarization induced by the stimuli?

I appreciate the materials being posted online. In the Study 1 Method it might be helpful to describe the predictions made by research groups A and B. It was hard for me to map the stimuli onto the hypotheses throughout the manuscript, and it can be difficult to search through supplements to make sense of methods. I think more clarity and a more specific roadmap describing how the materials tested specific directional hypotheses would be helpful.

Similarly, I think it would help to give more of a roadmap in the results to link the findings to what was expected (or not). For example, language like: “we found that Liberals showed higher (insert outcome) in condition A than condition B, but this pattern reversed for conservatives. This provided support for the (insert name of the hypothesis) because (insert brief description).

On page 18 it appears that the authors split the data to examine simple effects. I do not think this is an appropriate strategy, as all simple effects can be probed in the context of the original model. Comparing differences from split datafiles to an original is essentially comparing apples to oranges in terms of variance explained.

Study 2:

My comments for Study 2 are similar as for Study 1. The authors describe how the study differs (and ostensibly builds) on the findings of Study 1, but it was hard for me to tell why the specific predictions were made and how the findings relate to the research questions in the introduction in a specific fashion.

I have the same comment here about data splitting as above.

**Author Response**  
Jan 9, 2023

John Flournoy, Editor

Simine Vazire, Editor-in-Chief

*Collabra: Psychology*

Dear Dr. Flournoy:

We have enclosed a revised version of our submission to *Collabra: Psychology,* previously entitled “Motivated Reasoning in Research Decision Making” (revised to “Hypothesis Testing Preferences in Research Decision Making” to more fully capture the scope of the work). We are grateful for the opportunity to revise and resubmit our paper, and we believe the reviewers’ feedback has substantively strengthened our paper. Please find the reviewers’ comments pasted below; our responses are in *italics*.

As before, we have read the Transparency and Openness policy and aimed to adhere to it in our submission. We report all variables, treatment conditions, observations, and exclusions, and distinguish between confirmatory and exploratory analyses. All data, code, and materials for this research are available on OSF at <https://osf.io/854kq/?view_only=c2abbc117fd249a288ddec9529462dac>, along with the preregistered protocols and data analysis plans posted prior to data collection (Experiment 1: <https://osf.io/tywuf/?view_only=c2abbc117fd249a288ddec9529462dac>; Experiment 2: <https://osf.io/qr2da>). We have uploaded revised data files and code to OSF to include the additional analyses requested by the reviewers and correct a few minor errors in the descriptive statistics from the previous submission. The research reported in this article was conducted in compliance with the standards of the American Psychological Association. The manuscript is not currently under consideration at another journal, nor have the data been reported in the literature previously.

Thank you for your consideration of our revised manuscript.

Sincerely,

Stephanie Anglin

Department of Psychological Science

Hobart and William Smith Colleges

Geneva, NY 14456, USA

[anglin@hws.edu](mailto:anglin@hws.edu)

Caitlin Drummond Otten

School of Human Evolution and Social Change

Arizona State University

Tempe, AZ 85287, USA

[caitlin.drummond@asu.edu](mailto:caitlin.drummond@asu.edu)

Stephen Broomell

Department of Psychological Sciences

Purdue University

West Lafayette, IN 47907, USA

[broomell@purdue.edu](mailto:broomell@purdue.edu)

Dear Anglin,

I have now received 2 reviews of your manuscript, “Motivated Reasoning in Research Decision Making”, from researchers with special expertise in judgment and decision making. I also independently read the manuscript before consulting these reviews. The reviewers had mostly positive reactions to your manuscript. I agree that your manuscript has important strengths and also that there are some issues that need to be addressed. I therefore encourage you to submit a revised version for further consideration at Collabra: Psychology.

The reviewers did an outstanding job in their reviews. I will highlight issues I think are particularly salient here. In your resubmission, please include a document with a point-by-point response to both the points I list here and the reviewers’ comments, outlining each change made in your manuscript or providing a suitable rebuttal.

Please pay special attention to issues of clarity, highlighted by both reviewers, and testing the simple effects.

*We’ve thoroughly edited the paper in response to the helpful comments from both reviewers regarding ways to improve clarity. Our revisions include restructuring the Introduction to streamline our presentation of prior research and how it informed our research design and hypotheses. Please see our responses to Reviewers #1 and #2, below.*

*We’ve also clarified the presentation of the simple effects in Experiments 1 and 2. Please see our response to Reviewer #2, below.*

In summary, I think this is a promising manuscript and, I hope you will revise it for further consideration at Collabra: Psychology. I look forward to receiving your revision. Please see the instructions below for submitting your revision.

Please ensure that your revised files adhere to our author guidelines, and that the files are fully copyedited/proofed prior to upload. Please also ensure that all copyright permissions have been obtained. This may be the last opportunity for major editing, therefore please fully check your file prior to re-submission.

If you have any questions or difficulties during this process, please contact the editorial office at [editorialoffice@collabra.org](mailto:editorialoffice@collabra.org).

We hope you can submit your revision within the next six weeks. If you cannot make this deadline, please let us know as early as possible.

Sincerely,

John Flournoy

**Review #1**

**Open response questions**

I would like to thank Editor and the Authors for the opportunity to read this manuscript, it was quite informative and interesting. Below I list several issues I think might help to improve the readability and comprehensibility of the manuscript.

1. Some of the information in the Introductory parts are repetitive; and conversely, some of the concepts are not described in sufficient details when mentioned for the first time. For example, there is a good distinction between hot (motivated reasoning) and cold (cognitive biases) cognitive mechanisms (p. 4), but I miss more elaborate explanation how it relates to the research conducted. Authors mention myside bias and motivated reasoning interchangeably (are they treated as synonyms?) and they do not specify how their design aims to measure motivated reasoning, i.e. distinguish between hot and cold cognitive mechanisms in deciding about which study should be performed. This needs to be highlighted more in the Introduction.

*Thank you for this feedback regarding improving the clarity of the Introduction. We have edited and restructured the Introduction to reduce repetition and more clearly describe how prior research informed our research design and hypotheses (please see the section titled, “The Present Research” (starting page 8)). In addition, we have sought to more clearly describe how both hot and cold cognitive mechanisms may influence research decision-making in our experiments. We have also carefully edited the Introduction to clarify our usage of myside bias and motivated reasoning.*

1. I would recommend highlighting the novelty of this research, i.e. that in previous studies participants reasoned about “results” consistent or inconsistent with their beliefs, while this study examine the support for “hypotheses” consistent with their beliefs. Also, at least short review about previous findings on how framing of hypotheses might affect the choice of participants could be mentioned.

*We thank the reviewer for this suggestion, and we have edited the Introduction to highlight these points in the section titled, “The Present Research” (starting page 8).*

*We now state in the Introduction: “*Further, our research design also extends prior work on motivated scientific reasoning, which has typically examined perceptions of *results* consistent or inconsistent with beliefs (e.g., Drummond & Fischhoff, 2019; Lord et al., 1979; Munro & Ditto, 1997; Taber & Lodge, 2006), to examine preferences for *hypotheses* consistent or inconsistent with beliefs.” (page 8)

*We have also edited the Introduction to add a sentence about how our research builds on prior research on framing effects: “*Our research design thus builds on prior research in judgment and decision-making, which finds a robust *framing effect*, such that participants may display systematic preferences across two identical options when those options are framed differently (e.g., choosing different policies when an identical decision is framed in terms of lives saved vs lives lost; Tversky & Kahneman, 1981), to examine preferences for new research based on the framing of their hypotheses.” (page 8)

1. On the other hand, section “Confirmation and Disconfirmation Bias” does not seem to follow logically after previous section, where RQ1 was specified, as it seemed to elaborate more on rationale behind this research question. I would recommend to streamline introduction more, so that information provided is not repetitive (description of the current research is on p. 4 as well as p. 10, etc.).

*We have streamlined the introduction, following this helpful feedback as well as feedback from Reviewer #2. Our introduction now begins with sections detailing literatures relevant to our research, and concludes with a section (“The Present Research”; page 8) that describes how these literatures motivate our specific hypotheses.*

1. I also did not quite understand why authors posed research questions, although they provide a rationale for hypotheses and make predictions. If they choose to go with hypotheses, I would recommend building stronger case for why they expected reversal of preferences for conflicting results. If the participants use positive testing strategy, it still makes sense to repeat what they believe to be true. If the participants see both studies as equivalent, then they should have no preference regardless of their beliefs. And even if they are motivated to see particular result, it seems that they could believe that repeating belief-consistent study would yield stronger evidence for their case.

*We have edited our Introduction to include a section, “The Present Research,” that presents our hypotheses and how they follow from prior literature. We’ve also sought to build a stronger case for our hypothesis regarding a reverse of preferences when results are conflicting. Please see the text following Hypothesis 2:* “Scientific research can be performed when there is very little evidence (e.g., for emergent polarized topics such as how to mitigate COVID-19) or in the presence of a lot of conflicting evidence (e.g., nutrition topics). Previous research on disconfirmatory cognitive strategies focuses on scrutinizing evidence, which can only be done in contexts where prior evidence supporting a belief-inconsistent hypothesis exists. We predicted above, in Hypothesis 1, that without prior evidence, participants will rely on confirmatory strategies, as there is no evidence to scrutinize. However, in the presence of conflicting evidence, participants now have two options: they can either continue to apply a confirmatory strategy, or they can pivot toward a disconfirmatory strategy by scrutinizing the belief-inconsistent evidence. We therefore predict that preferences for running the belief-inconsistent study will increase in the presence of conflicting evidence.” (p. 10).

1. In contrast, it seemed to me that review of literature regarding individual differences allowed for more specific predictions than posed by RQ3. Authors could reasonably expect no preference for participants with better scientific reasoning skills (p. 10).

*We have edited the Introduction to form more specific hypotheses with regards to each of our individual different measures. Please see Hypothesis 3 in the newly added section of the Introduction entitled, “the Present Research” (page 8). We have added a footnote to highlight that although these hypotheses were not specified in the preregistration, they were guided by prior research and theory.*

1. As mentioned in point 1, it is not clear whether Authors mean the same thing when they talk about motivated reasoning and myside bias. I recommend clarifying and uniting the terminology. Moreover, it is not clear whether myside bias is one of the variables and how it is operationalized (or what will be counted as motivated reasoning).

*As mentioned in our response to comment #1, we have carefully edited the paper to clarify our terminology and distinguish between motivated reasoning and myside bias. We’ve also edited our paper to clarify the potential mechanisms driving our findings. As we state in the Discussion section of Experiment 1, “*Our quantitative findings are consistent with either “cold” (positive test strategy) or “hot” (directionally motivated reasoning) processes… Our qualitative results suggest that both contributed to participants’ decisions….”. *We reprise and extend this discussion in the Discussion section for Experiment 2 and the General Discussion.*

1. I appreciate analyzing also qualitative responses and I think it deserves highlighting in the Introduction as one of the contributions of the current research.

*We have added sentences to the Introduction highlighting this contribution: “*We also collected open-ended responses in which participants described the rationale for their preference (or lack thereof) to better understand the reasoning underlying their choices. Our usage of qualitative data represents a unique contribution of our paper in that we classify participants’ rationales in terms of the psychological processes (e.g., confirmatory vs. disconfirmatory and motivated by directional vs. non-directional goals) highlighted above.” (page 8)

1. In the Methods section, I would recommend explicitly stating what was your main dependant variable.

*We have edited our Methods section to make our main DV, preferences for which study to conduct in the absence of evidence and presence of conflicting evidence, more clearly stated.*

1. I did not understand why results for some of the individual measures are reported in the main text and some in supplementary materials. Could you explain or include them all in the main results?

*Due to length concerns, we had originally reported some of the individual different measures in the supplementary materials only. We have edited the paper to discuss all of the individual different measures in the main text.*

1. If I understand it correctly, three questions were used to measure participants’ beliefs (related to politicized topic of gun control): political orientation (liberal vs. conservative), belief about efficacy of gun control policy and position about gun control, but only two of them were used in analyses (political orientation and gun control belief). Why? Or if they are highly correlated, maybe you could use composite measure?

*We appreciate the reviewer bringing up this point. We had overlooked the possibility of conducting an additional robustness check for gun control position or for all three belief variables combined as a composite. We added these analyses to the paper, which show the same pattern of results for the other two belief variables. Because we preregistered the analysis using political orientation as the main test of the hypothesis and the analysis with gun control belief as the robustness check, we felt that conducting additional robustness checks was more appropriate than replacing the original analyses.*

1. Table 1 – I found the title a little confusing – what does a “forced choice” mean? In materials Authors describe that participants could expressed “no preference”.

*We agree that this is confusing, as participants reported their preference for which study to conduct on a 5-point scale with a neutral midpoint. We removed forced choice from the title of Table 1.*

1. P. 17: “responses to each question were coded…” – please, specify to what question. Similarly, on p. 18: “The condition x gun control belief interaction became stronger” – please, specify which condition. Also, p. 23: “responses were coded so that higher scores indicate…” – please, specify responses to what.

*We have edited the Results section to clarify each of these points.*

1. Moreover, I did not understand how were responses coded so they were belief-consistent to both liberals and conservatives – can you specify the procedure? I also wondered if the chi-square or t-test were not more appropriate to see the differences between liberals and conservatives and their choice and differences in scores.

*We have clarified the text to specify how responses were coded: Responses to each DV (which study participants preferred to conduct in the absence of evidence, and in the presence of conflicting evidence) were coded so that higher scores indicated a stronger preference for the gun control decreases crime study in the experimental condition, and the skin cream makes rashes better study in the control condition.*

*We further clarified the paper to state that we conducted one-sample t-tests to examine whether liberals’ and conservatives’ choices significantly differed from the midpoint of no preference, asking whether participants preferred to test hypotheses consistent (or inconsistent) with their beliefs. We would be happy to add post hoc independent sample t-tests comparing liberals’ and conservatives’ choices if the reviewer thinks that these would add to our explication of our results.*

1. Table 4 – to be consistent with your terminology, Scenario should be renamed as Study.

*In our paper, we used Scenario to refer to the absence of evidence and conflicting evidence (within-subject) conditions and Study to refer to the two studies with opposing predictions/findings in each scenario. We added a clarifying statement to the Method section on p. 16 to address this confusion.*

1. Discussion to Experiment 1 – Do you consider your results to provide support for PTS explanation rather than motivated reasoning? (Because the results were similar for politicized as well as control study?) You found different rationales for participants´ choices – some of them more in line with motivated reasoning account, some of them more in line with PTS. What are the implications of these findings? Can you elaborate a bit more?

*We thank the reviewer for bringing up this important point, and we’ve expanded the Discussion section for Experiment 1 to further discuss potential mechanisms driving our results. We have also revised the title of the paper from “Motivated Reasoning in Research Decision Making” to “Hypothesis Testing Preferences in Research Decision Making” to avoid overstating the mechanisms underlying the findings.*

1. Discussion to Experiment 2 – Authors wrote (p. 34-35): “people have a tendency to favor research with belief-consistent hypotheses and do not significantly alter their decision-making strategy from before the outcomes are known to after competing outcomes are obtained. These preferences do not appear to be driven by the assumption that the researchers are manipulating the outcome, as participants showed the same belief-consistent preference when an independent research group conducted the study.” Do these results imply that people hold their opinions regardless of the evidence? And again, because there was no difference whether researchers or independent research group conducted the study – can it be concluded that people rely on PTS?

*We thank the reviewer for bringing up these interesting points. We’ve expanded the Discussion to Experiment 2 to discuss them further. The results suggest that participants held a belief-consistent preference in the absence of evidence and presence of conflicting evidence. However, because we didn’t measure participants’ beliefs about gun control after learning about the evidence, we can’t conclude from our experiments that participants hold their opinions about gun control regardless of the evidence (though prior studies have examined belief perseverance vs. change in response to mixed evidence; see Anglin, 2019; Lord et al., 1979; Munro & Ditto, 1997; Taber & Lodge, 2006). As we state in the Discussion for Experiment 2, it’s possible that our findings are driven by a positive test strategy, but the diversity of participants’ open-ended responses suggest several potential mechanisms, including a positive test strategy and motivated reasoning.*

1. It would be interesting to see the results for ambiguous results (after the research in the scenario is conducted, but inconclusive results are found) or with different polarizing topic. (This is just an idea; you do not need to respond to this.)

*Thanks for these interesting suggestions, which we’ll keep in mind for future follow-up studies.*

1. Lastly, I found the last sentence vague and generic. The research reported in the manuscript is interesting, and I believe that the paper should end with stronger and more specific message than generic conclusion that more research is needed.

*We’ve added a stronger, more specific final sentence to the paper: “*However, the present study provides preliminary evidence that the public more strongly supports research with hypotheses consistent their beliefs, which holds important implications for garnering public support and setting funding priorities for research on politicized topics.” (page 52).

In general, I enjoyed reading the paper, I found the methodology and analysis adequate and I highly appreciate pre-registration and sharing the materials, which help to understand some details better.

*Thank you for the thoughtful and detailed review.*

**Review #2**

**Open response questions**

The current manuscript tests several hypotheses related to the way motivated reasoning impacts preference for scientific methods and findings. For example, the author(s) first research question is “Do people exhibit a preference for which of two scientifically equivalent studies to conduct based on the framing of the hypotheses?” Overall, I think that this research is interesting and could be very important to understand the conditions that lead to the funding (or not) of important scientific work. My major comments about the paper have to do with structure and clarity both in terms of generating the hypotheses based on theory and clearly and cleanly reporting the results.

Big Picture Comments

Overall, I think the authors are tackling an interesting and important set of questions. They have several research questions, and I did find the overall theoretical perspective/rationale being used to derive the predictions to be a bit difficult to ascertain. The author(s) provide a thorough review of literatures relating to confirmation bias, motivated reasoning, etc., but there does not seem to be a clear thread throughout that informs the hypotheses and contextualizes them in previous theory. Furthermore, it did not seem to me that the hypotheses are necessarily related and each one could almost be fleshed out to serve as a standalone paper. I think it would be helpful to establish a central perspective that justifies and links together the study hypotheses. One way to do this might be to lay out the broad literatures of interest, describe how they may relate to each of the three research questions, and then have subsections for each of the three hypotheses.

*Thank you for this helpful suggestion. We’ve restructured the Introduction in the way that is suggested here.*

It also did not seem to me that the hypotheses were directional in nature. For example, the hypothesis given above in this review does not seem to establish an expected direction of results but is a general question. Directional hypotheses might help to clarify what it is that the researchers expected. RQ3 is similar and does not feel strongly grounded in specific previous work. Generally, I think the authors test three interested and related questions but it takes a lot of bandwidth to derive which perspectives inform which hypotheses, how the methods test the hypotheses, and how the findings map onto the predictions.

*We thank the reviewer for this suggestion, which was also echoed by Reviewer 1. The restructured Introduction now includes a section, “The Present Research,” that draws on the literature review to state directional hypotheses. We’ve also edited the paper to more clearly tie the hypotheses to the methods and findings.*

In terms of the general discussion, there is quite a lot of speculation about why the findings differed from expectations starting on page 37 of the manuscript. I think it’s great to report unexpected findings, especially if they replicate across experiments. I also found this section a bit difficult to parse given that the predictions had seemed unclear to me from the outset. I think a grounding of all three hypotheses in the literature, directional statements, and tie-ins to which patterns map onto which expectations in the results would help with this.

*We’ve edited the Discussion to more closely map onto our rewritten Introduction and our three Hypotheses.*

Comments on the Studies

Study 1

In Study 1 skin cream being ineffective is used as a control condition (i.e. a non-polarizing topic). However, I do not see a measure of people’s agreement with skin-cream effectiveness like I do for agreement with gun control. How do we know these two conditions differ in the degree of polarization induced by the stimuli?

*Thank you for bringing up this point. Although previous research suggests that skin cream effectiveness functions as a non-polarized control condition (e.g., Kahan et al., 2017), we did not measure beliefs about skin cream effectiveness as we did for gun control. We added a footnote to highlight this limitation in the Method section of Experiment 1 on p.15: “*We note that although previous research suggests that gun control functions as a polarized stimulus topic and skin cream effectiveness as a non-polarized stimulus topic (e.g., Kahan et al., 2017), the present experiment did not measure beliefs about skin cream effectiveness to compare to the strength of participants’ beliefs about gun control. Therefore, the degree to which participants found the gun control stimulus to be more polarizing than the skin cream stimulus cannot be asserted from the present data (though participants’ open-ended explanations for their preferences suggest more polarizing responses to the gun control vs. skin cream stimuli).”

I appreciate the materials being posted online. In the Study 1 Method it might be helpful to describe the predictions made by research groups A and B. It was hard for me to map the stimuli onto the hypotheses throughout the manuscript, and it can be difficult to search through supplements to make sense of methods. I think more clarity and a more specific roadmap describing how the materials tested specific directional hypotheses would be helpful.

*Thank you for this helpful suggestion. We have edited our description of the studies to make them clearer, and make it easier to follow how they test our hypotheses.*

Similarly, I think it would help to give more of a roadmap in the results to link the findings to what was expected (or not). For example, language like: “we found that Liberals showed higher (insert outcome) in condition A than condition B, but this pattern reversed for conservatives. This provided support for the (insert name of the hypothesis) because (insert brief description).

*Thank you for this helpful suggestion. We have edited the Methods, Results, and Discussion sections to more clearly link our hypotheses with our findings and their implications.*

On page 18 it appears that the authors split the data to examine simple effects. I do not think this is an appropriate strategy, as all simple effects can be probed in the context of the original model. Comparing differences from split datafiles to an original is essentially comparing apples to oranges in terms of variance explained.

*Thank you for identifying this area of confusion. We edited the Results section to clarify which analyses we performed and why. We performed the simple effects analyses in the original model, as reported on page 21: “*Simple effects analyses revealed that participants in the experimental condition were more likely to favor the study with the belief-consistent hypothesis across both scenarios: those with a stronger liberal ideology preferred to have the gun control decreases crime study conducted (and those with a stronger conservative political ideology preferred to have the gun control increases crime study conducted), β = -0.33, B = -0.22, SE = 0.04, *t* = -5.27, *p* < 0.001. Ideology was unrelated to study preference in the control condition, β = -0.04, B = -0.03, SE = 0.04, *t* = -0.59, *p* = 0.56.”

*We also split the data to conduct an additional post hoc examination of the simple effects: specifically, we split participants into ideology groups to conduct separate analyses testing whether participants exhibited belief-consistent preferences significantly differing from the midpoint of no preference. We did this to provide a further test of Hypotheses 1 and 2, asking if people’s preferences could accurately be characterized as preferring the belief-consistent or belief-inconsistent hypothesis by comparing them to the neutral midpoint.*

*We’ve edited the paper to clarify that this is a post hoc analysis that provides an additional test of our Hypothesis 1 and 2: “*To further examine whether participants displayed a preference to conduct the study testing the belief-consistent hypothesis (Hypotheses 1 and 2), we conducted a *post hoc* analysis in which we divided participants into 3 groups based on whether they reported a liberal (somewhat to very; gun control: *n* = 141, skin cream: *n* = 140), moderate (gun control: *n* = 40, skin cream: *n* = 42), or conservative (somewhat to very; gun control: *n* = 49, skin cream: *n* = 50) political orientation. We performed one sample *t*-tests to test whether each political group’s preferences differed from the midpoint of no preference, for those in the gun control (polarized) condition only.” (page 23)

Study 2:

My comments for Study 2 are similar as for Study 1. The authors describe how the study differs (and ostensibly builds) on the findings of Study 1, but it was hard for me to tell why the specific predictions were made and how the findings relate to the research questions in the introduction in a specific fashion.

I have the same comment here about data splitting as above.

*We’ve made edits to Experiment 2 consistent with the suggested edits to Experiment 1.*

*Overall, we thank the reviewer for their helpful and detailed feedback.*

**Editor First Decision**: Revise & Resubmit

Feb 7, 2023

Dear Dr. Stephanie Anglin,

I have now read your revised manuscript. I appreciate your careful attention to the concerns the reviewers and I raised. I am happy to provisionally accept your manuscript for submission. However, I found a few small things I would like you to address.

Minor revisions were suggested by reviewer 2. Please try to catch other small issues prior to final resubmission.

I look forward to receiving your final revision and accepting it for publication in Collabra: Psychology.

Please ensure that your revised files adhere to our author guidelines, and that the files are fully copyedited/proofed prior to upload. Please also ensure that all copyright permissions have been obtained. This is the last opportunity for major editing, therefore please fully check your file prior to re-submission.

If you have any questions or difficulties during this process, please contact the editorial office at [editorialoffice@collabra.org](mailto:editorialoffice@collabra.org).

We hope you can submit your revision within the next six weeks. If you cannot make this deadline, please let us know as early as possible.

Sincerely,

John Flournoy

# Reviewer 1

##### Rating scale questions

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  |  | ✔ |

##### Open response questions

### Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

I would like to thank the authors for their responses to my comments. I am happy with the changes they made and recommend article for publication.

# Reviewer 2

##### Rating scale questions

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  |  | ✔ |

##### Open response questions

### Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

I was a reviewer on the first submission of this manuscript. Overall, I think the authors have done a nice job in re-working the structure of the introduction to cover the psychological processes of interest and to clarify the theoretical perspective(s)/hypotheses. I also think the way that the author(s) have included the qualitative response examples clarifies different styles of reasoning and how they fit within the framework of the theory. I also like the way the author(s) now provide statements throughout the results referring to support/lack of support for the different perspectives. This helps clarify which perspectives seem to most fit the data.

At this stage I don’t feel I have much to add. There were a few places with some small typos that should be fixed. On page 22 the partial eta squared values of of 0.000 (two of them) should probably be < .001. On page 37 the sentence “In the presence of conflicting evidence scenario, before asked which study they would prefer to have conducted on again on another sample of 300 cities…” seems oddly worded and should probably be re-worked.

**Author Response**  
Feb 16, 2023

**Editor Final Decision:** Accept

Feb 16, 2023

Dear Stephanie Anglin,

I have now had a chance to read over your manuscript “Hypothesis Testing Preferences in Research Decision Making”, along with the letter describing the changes you made. Thank you for your responsiveness to the concerns that the reviewers and I raised. I am happy to say that your paper is now officially accepted for publication in Collabra: Psychology. Congratulations on this excellent work, I think it will make an important contribution to the literature and I look forward to seeing it published! I hope your experiences with Collabra: Psychology have been positive and that you will continue to consider it as an outlet for your work.

As there are no further reviewer revisions to make, you do not have to complete any tasks at this point.

You will be receiving separate correspondence regarding any production and technical comments, data deposits, as well as publication charges. We work with the Copyright Clearance Center to process any applicable APC charges. Please note that your APC transaction must be completed before your article gets published.

You will have an opportunity to check the page proofs before we publish your article. Thank you again for publishing in Collabra: Psychology.

Sincerely,  
John Flournoy