**Peer Review and Communication History**

**Ms Title:** Validation of a set of stimuli to investigate the effect of attributional processes on social motivation in within-subject experiments

**Author names:** Fucci E., Baquedano C., Abdoun O., Deroche J., Lutz A.

**Submitted:** Nov 10, 2020

**Editor First Decision—Revise & Resubmit**

Jan 26, 2021

Dear Dr. Lutz,

Thank you for submitting your manuscript for consideration by Collabra: Psychology. I want to start by apologizing for how long it has taken to get you a decision – I struggled mightily to get two reviews for the paper. Fortunately, I have now received two reviews of your manuscript from researchers with considerable expertise in your field. These reviewers provided expert feedback and I thank them for their service to this journal. I also independently read your manuscript before consulting these reviews. As you will see, the reviewers were quite similar and positive in their assessments of the paper. My reading of the manuscript is also similarly positive. As a result, I am inviting you to submit a revised version of the manuscript for further consideration at Collabra (and I am hoping this decision takes a bit of the sting out of a longer than ideal review process).

Both reviews provide excellent and detailed reviews. I agree with the central treatise that both reviewers note – that there could be more clarity throughout the manuscript. For instance, one of the reviewers discusses the choice of anchors and points to Comrey (1988). I think the reviewer could have also pointed to Krosnick (1999) who also discussed (at length) the issues surrounding response scale creation. I certainly believe that you should address this issue in the paper – certainly, some researchers have argued that the typical Krosnick recommendations don’t always apply (for instance, I detail why Krosnick’s recommendations fail in the case of typically extreme emotions – Edlund & Sagarin, 2009). The other reviewer notes several methodological concerns (such as the face validity being a challenge and the meaning of the response times).

In your revision, please highlight the changes made (I would appreciate you submitting a version with change tracking enabled as well as a clean version of the manuscript). In your coverletter, please detail your responses to all of the points raised about your manuscript summarizing the changes when they are made and if changes are not made, please justify that decision.

Sincerely, John E. Edlund

**Reviewer 1**

**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 study aimed to validate video stimuli that could be used in studies on social motivations. These stimuli would be a contribution to the field because they allow for within-participants studies. To-date, most of the studies use between-participants designs. There is much to like about this study. Namely, the study seems technically sound, the figures are nice, and the authors publicly share the stimuli, data, and analysis code. However, despite these positives, I had several comments and concerns. Here are my major comments. These comments are listed roughly in the order they appear in the manuscript and not in order of importance.

1. In the Introduction, the authors home in on the fact that much of the past research has used cross-sectional data that are used to study person-level processes. An illustrative quote is when the authors say “This passing from cross-sectional data to within-person processes implicitly assumes that the structure of interindividual differences reflects intraindividual processes. However, there are many situations where this is simply not true.” I agree. However, the authors leave a lot of the mental work up to the reader. I would ask the authors to fully flesh out the idea that they imply. In this particular research domain, how might their within-participants design specifically shed new light onto attributions on social motivations?
2. The framing of the study seems off. The actual study feels like a theory-testing study, not a stimuli validation study. The authors present several reasons why the within-participants analysis might produce different results than the traditional between-participants design, which seems like the goal is to test a theoretical proposition. I know this is not a minor comment, but the authors might consider reframing the study.
3. A huge benefit of within-participants designs is that you can higher achieve statistical power. However, some research questions might be difficult to test with a within-participants design due to practice effects, carry-over effects from one trial to other trials, etc. The authors might consider discussing the possible trade-offs when deciding on a between-participants design and a within-participants design.
4. What were the hypotheses for the response time data? Additionally, the outcome is not a task where the participants were instructed to respond quickly, there is not response competition that you see in some tasks (such as a Stroop Task or the IAT), etc. What exactly did the authors believe the response times represent?
5. From what I can tell, the data are analyzed thoroughly and appropriately. In full disclosure, I am unfamiliar with the LASSO analyses, so I am not able to say with confidence that those analyses are correct.
6. How much of these results are due to demand characteristics? The contexts of the stimuli are unambiguous and the outcome measures are very face valid. Is there any way to determine whether participants merely intuited the study hypotheses and then responded accordingly?

**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.) |  |  | ✔ |  |  |

**Reviewer 2**

**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.

**2. Comments to Author:** *Summary and general impression* Authors in the current study created a set of stimuli that would allow for the study of attributional processes on prosocial motivation in within-subjects experiments. The stimuli created are videos accompanied by vignettes that influence perceptions of the main character’s responsibility. Authors then examine the effect of the responsibility manipulation on valance, arousal, empathy, responsibility, deservingness, and willingness to help. The manipulation of responsibility was successful; participants’ judgments on responsibility and deservingness of the main character changed depending on the context associated with each video. Different contexts also influenced affective responses to emotional videos, as well as participants’ willingness to help. Results were also in line with previous work as authors note a negative relationship between willingness to help and attribution of responsibility, a positive relationship between the degree of empathy for the sufferer and the willingness to help, and a strong association between attribution of responsibility and judgments on deservingness. Generally, this study provides a useful and novel tool for studying attributional processes by allowing researchers to move from between subjects designs to within subjects designs. There are several places in the manuscript where clarity and readability could be increased by the inclusion of additional detail – mainly relating to the procedures and theories underlying the authors’ work.

**Scope of the journal** The manuscript is in line with Collabra’s focus area of “Methodology and Research Practice” as it focuses on a novel paradigm for studying attributional processes on prosocial motivation in within-subjects experiments.

**Specific Comments** *Introduction* • Provides a framework for the current study and clearly identifies the need for the current work • Proofread for minor grammatical errors

*Method* • Authors explain that they will conduct exploratory analyses on response times to highlight how the stimuli can facilitate the study of attributional processes with implicit measures and provide insight on behavioral mechanisms but there is a lack of information on how response time was measured  
• In the VALIDATION – Instructions and Questions the response options seem to only include a strong positive, negative, and neutral option (e.g., 1 = very not responsible 9 = very responsible 5 = neither one nor the other). I would appreciate a more detailed explanation about why this response format was chosen as opposed to a more nuanced one that includes ‘somewhat responsible’ for example. Comrey (1988) asserts that unbalanced response distributions can lead to distorted correlational results. Comrey (1988) also argues that quantitative scales with at least five response categories are optimal because they will produce a more or less continuous distribution as opposed to a point distribution

*Analytic plan* o Authors provide a detailed analytic plan with a thorough explanation of the current study’s power o I am curious about gender differences in attribution of blame, and wonder if the authors would consider conducting analyses by gender… in the sexual violence literature there are significant differences in attribution of blame by gender (with men attributing greater blame than women for situations like domestic violence and rape; see Bryant & Spencer, 2003 and Grubb & Turner, 2012). Additionally, the literature suggests that perceptions of similarity influence attributions of blame (concept of judgmental leniency; in-group / out-group) and I’m curious as if congruency between gender of participant and main character could influence results

*Results* • I appreciate the inclusion of the supplemental files and access to data for readers who wish to examine the results further; this increases both clarity and transparency of the analyses and subsequent discussion of results

*Discussion/Limitations/Future Directions* • Subheadings assist with organization and flow • I think a future directions subsection would be worth adding just to increase the clarity and provide concrete recommendations about future work; currently it is lost in the discussion and limitations

**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.) |  |  | ✔ |  |  |

**Author Response**

March 1, 2021

Dear Editors at Collabra:Psychlogy,

We thank the reviewers and the editor for taking their time to revise our initial manuscript submission. We have made changes to the original manuscript that will hopefully clarify some of the issues raised in the review process. Together with this letter, we submit the updated manuscript in two versions (clean and with change tracking enabled).

Additionally, please find below our responses to the reviewers’ and editor’s comments, where we summarize the changes made to the original manuscript and we respond to suggestions and critical points.

Sincerely,

The authors

***The framing of the study seems off. The actual study feels like a theory-testing study, not a stimuli validation study. The authors present several reasons why the within-participants analysis might produce different results than the traditional between-participants design, which seems like the goal is to test a theoretical proposition. I know this is not a minor comment, but the authors might consider reframing the study.***

The study was born as a stimuli validation. As such, we tested main hypotheses on the effect of context-video pairs. To increase the confidence on the validity of this set of stimuli in a within-subject design, we explored associations between variables with the aim of replicating the results coming from the existing literature on attribution processes. We discovered that while most of the results corroborated established theories, some differences might be present when implementing a within-subject approach and they should be further explored. We understand that this might lead the reader to question whether this study was intended as a theory testing, but the main objective remains that of stimuli validation. Accordingly, the methods were not designed (and are not appropriate) for theory testing as the study has been powered for within-subject analyses but not for the traditional between-subject design. Therefore we chose to maintain the current framing. However, we added a “Future directions” section, as suggested, that addresses this and other comments.

***In the Introduction, the authors home in on the fact that much of the past research has used cross-sectional data that are used to study person-level processes. An illustrative quote is when the authors say “This passing from cross-sectional data to within-person processes implicitly assumes that the structure of interindividual differences reflects intraindividual processes. However, there are many situations where this is simply not true.” I agree. However, the authors leave a lot of the mental work up to the reader. I would ask the authors to fully flesh out the idea that they imply. In this particular research domain, how might their within-participants design specifically shed new light onto attributions on social motivations?***

Our intention with this sentence in the introduction was to point out a very general epistemological challenge in the field of psychology: the discrepancy between the nature of the most common objects of study in psychology, i.e. intrapersonal mental and behavioral processes, and the nature of one of the major methodologies employed in the field, namely the modelling of the structure of between-subject variability. We hope we have been reasonable in expecting readers with a background in psychology to be familiar with this epistemological issue, and because a more elaborate exposition of the topic is out of the scope of this paper, we referred to a thorough and well-cited treatment by Hamaker (2012).

Considering that our study is a stimulus validation, we have also deemed inappropriate to formulate specific hypotheses on how results from within-participant designs might differ from the ones established in the existing literature based on between-subject variability. However, and following the invitation of the reviewer, we have suggested in the Future directions section a few ways by which the within-subject methodology can provide a basis for experimental designs with the potential of refining our understanding of attributional processes.

***A huge benefit of within-participants designs is that you can higher achieve statistical power. However, some research questions might be difficult to test with a within-participants design due to practice effects, carry-over effects from one trial to other trials, etc. The authors might consider discussing the possible trade-offs when deciding on a between-participants design and a within-participants design.***

We thank the reviewer for this valuable suggestion. We have added remarks of such nature in a Future directions section.

***What were the hypotheses for the response time data? Additionally, the outcome is not a task where the participants were instructed to respond quickly, there is not response competition that you see in some tasks (such as a Stroop Task or the IAT), etc. What exactly did the authors believe the response times represent?***

As stated in the manuscript, the analysis of response times was purely exploratory. It was performed to show how this type of stimuli and design could be implemented to study third-person, implicit behavioural measures. As such, there were no specific hypotheses and the interpretation of results was done post-hoc. In the revised manuscript, we added a description of how response times were calculated and how they differed from classic measures of reaction times.

***Authors explain that they will conduct exploratory analyses on response times to highlight how the stimuli can facilitate the study of attributional processes with implicit measures and provide insight on behavioral mechanisms but there is a lack of information on how response time was measured.***

As mentioned above, this point is addressed in the manuscript.

***In the VALIDATION – Instructions and Questions the response options seem to only include a strong positive, negative, and neutral option (e.g., 1 = very not responsible 9 = very responsible 5 = neither one nor the other). I would appreciate a more detailed explanation about why this response format was chosen as opposed to a more nuanced one that includes ‘somewhat responsible’ for example. Comrey (1988) asserts that unbalanced response distributions can lead to distorted correlational results. Comrey (1988) also argues that quantitative scales with at least five response categories are optimal because they will produce a more or less continuous distribution as opposed to a point distribution.   
Here the editor adds: I think the reviewer could have also pointed to Krosnick (1999) who also discussed (at length) the issues surrounding response scale creation. I certainly believe that you should address this issue in the paper – certainly, some researchers have argued that the typical Krosnick recommendations don’t always apply (for instance, I detail why Krosnick’s recommendations fail in the case of typically extreme emotions – Edlund & Sagarin, 2009).***

As far as we have understood the reviewer’s and editor’s comments on the response scale, there might have been a misunderstanding originating in the instructions shared on OSF. Indeed our instructions file mentions only the labels of the scale. But as we described in the Methods section of the paper and as is visible in the distributions of the data in Figure 2, the scales had 9 categories signaled by ticks equally distributed on a horizontal line. Only the extreme and the middle ones were labelled. Non-labelled intermediate ticks were implicitly referring to nuanced ratings and have been fully used by the participants.

We have updated the instructions file so as to clear up the misunderstanding.

***I am curious about gender differences in attribution of blame, and wonder if the authors would consider conducting analyses by gender… in the sexual violence literature there are significant differences in attribution of blame by gender (with men attributing greater blame than women for situations like domestic violence and rape; see Bryant & Spencer, 2003 and Grubb & Turner, 2012). Additionally, the literature suggests that perceptions of similarity influence attributions of blame (concept of judgmental leniency; in-group / out-group) and I’m curious as if congruency between gender of participant and main character could influence results.***

No hypothesis was formulated in terms of how gender differences could influence the validation of stimuli. As such, an analysis of self-reports including a gender factor is not guaranteed to achieve the desired statistical power. Nonetheless, we tested a model on self-reports of Deservingness attribution where we included an interaction between context and gender as fixed effect, and as by subject and video random slope. The model estimation shows neither an interaction nor a main effect of gender, as in the table below:

Type III Analysis of Variance Table with Satterthwaite's method

Sum Sq Mean Sq NumDF DenDF F value Pr(>F)

Context 174.387 174.387 1 30.424 101.8362 3.17e-11 \*\*\*

Gender 1.774 1.774 1 57.732 1.0360 0.3130

Context:Gender 2.680 2.680 1 54.640 1.5651 0.2163

Regarding the comment on in-group/out-group differences, it is addressed in the limitation section.

***How much of these results are due to demand characteristics? The contexts of the stimuli are unambiguous and the outcome measures are very face valid. Is there any way to determine whether participants merely intuited the study hypotheses and then responded accordingly?***

While there is no explicit way to test whether results were influenced by demand characteristics, the most convincing piece of evidence is that the distributions of Responsibility and Deservingness ratings across contexts are not bimodal. In particular the variances for the responsible context are quite high, indicating that even though the story made it very clear that the sufferer was "guilty", the participants more often than not chose a moderate rating. Differences in response time also indicate some kind of personal deliberation, inconsistent with merely fulfilling presumed hypotheses.

***I think a future directions subsection would be worth adding just to increase the clarity and provide concrete recommendations about future work; currently it is lost in the discussion and limitations.***

We thank the reviewer for this helpful suggestion which helped us sort our various considerations, stress out the significance and potential of our within-subject approach, and address several other points raised by the reviewers.

**Editor Second Decision—Revise & Resubmit**

March 23, 2021

Dear Dr. Fucci, Thank you for submitting your revised manuscript for consideration by Collabra: Psychology. I was fortunate to have both reviewers agree to read the revised manuscript; additionally, I independently read your revised manuscript before consulting these reviews. In summary, the reviewers and I appreciate the effort that went into the revision and the additional clarity that created. As such, I am happy to accept the manuscript pending one minor tweak to the paper; specifically, the comment by the 2nd reviewer who asks “A reference with an example of a study where the focus is on in-group / out-group affiliation and attribution of blame, empathy, and/or responsibility may be helpful to readers here.“ I agree with that suggestion and I think it would improve the paper. As such, please track down such a citation and add it to a final revision of the paper (which I anticipate accepting upon receipt). Congratulations in advance on a fine addition to the literature. Sincerely, John E. Edlund

# Reviewer 1

##### 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 reviewed the initial submission of this manuscript. No new comments or concerns were raised in the revised manuscript. The authors were responsive to the comments either by adding material to the manuscript or adding explanation in their cover letter. I think this manuscript would make a nice contribution to the field and, in my estimation, should be accepted for publication.

##### 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.) |  |  |  | ✔ |  |

# Reviewer 2

##### 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.

Summary and general impression: Authors in the current study created a set of stimuli that would allow for the study of attributional processes on prosocial motivation in within-subjects experiments. The stimuli created are videos accompanied by vignettes that influence perceptions of the main character’s responsibility. Authors then examine the effect of the responsibility manipulation on valance, arousal, empathy, responsibility, deservingness, and willingness to help. The manipulation of responsibility was successful; participants’ judgments on responsibility and deservingness of the main character changed depending on the context associated with each video. Different contexts also influenced affective responses to emotional videos, as well as participants’ willingness to help. Results were also in line with previous work as authors note a negative relationship between willingness to help and attribution of responsibility, a positive relationship between the degree of empathy for the sufferer and the willingness to help, and a strong association between attribution of responsibility and judgments on deservingness. Generally, this study provides a useful and novel tool for studying attributional processes by allowing researchers to move from between subjects designs to within subjects designs. There are several places in the manuscript where clarity and readability could be increased by the inclusion of additional detail – mainly relating to the procedures and theories underlying the authors’ work.

Scope of the journal: The manuscript is in line with Collabra’s focus area of “Methodology and Research Practice” as it focuses on a novel paradigm for studying attributional processes on prosocial motivation in within-subjects experiments.

2.1 Specific Comments Introduction • Proofread for minor grammatical errors • The inclusion of the Hamaker, 2012 citation, and brief expansion in the footnote is helpful for contextualizing why it is erroneous to assume that the structure of interindividual differences reflects intraindividual processes. Method • There is a cleaner explanation of the rationale for exploring RT as well as how RT was assessed and what information it provides. I especially appreciate the explanation that in this case RT represents “deliberation time” for each question and experimental condition, and is one example of how this work might be used by others. • I appreciate the authors being more clear about the response scale options that were used in the VALIDATION –Instructions and Questions. I think this enhances the transparency and readability of the manuscript. Analytic plan o I appreciate the authors exploring how gender may influence attribution of blame, as well as the mention of in-group/out-group (p. 26). A reference with an example of a study where the focus is on in-group / out-group affiliation and attribution of blame, empathy, and/or responsibility may be helpful to readers here. Results • Discussion/Limitations/Future Directions • The future directions subsection increases the clarity and provides concrete recommendations about future work. Making future directions its own section has helped underscore the significance of the work as well as next steps.

##### 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.) |  |  |  | ✔ |  |

**Author revision submitted**

March 23, 2021

**Editor Final Decision—Accept**

March 23, 2021

Dear Dr Fucci,

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. Our managing editor will contact you in case there are any pre-prodution file related questions. 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 Edlund