XX XXX 2025

Dear Professor Conway,

We are submitting a revised version of the manuscript “Space to Think: Testing the Effect of Distancing on Moral Dumbfounding - A Registered Report” for consideration in *Journal of Experimental Social Psychology*. We would like to thank you for this opportunity to submit a revised manuscript. We would also like to thank you and the reviewers for the very helpful and insightful feedback that has contributed to, in our view, a much-improved manuscript. We thoroughly engaged with the feedback provided and have made substantial changes to the manuscript.

Below, we address each comment from the review team, point by point. All changes in the manuscript are highlighted in yellow.

We hope that the changes described below sufficiently address the concerns raised and that the revised manuscript can be granted in-principle acceptance as a Stage 1 Registered Report in *Journal of Experimental Social Psychology*.

Kindest regards

The authors.

|  |  |  |
| --- | --- | --- |
| **Editor** |  |  |
| *Ed Feedback* |  | *Dear Dr McHugh*  *Thank you for submitting JESP-D-24-00351, Space to Think: Testing the Effect of Distancing on Moral Dumbfounding - A Registered Report, to JESP. Thanks especially for your patience in waiting for feedback; as you likely know, it is increasingly difficult to secure quality reviews in a timely fashion. Fortunately, I was able to secure reviews from three experts in the field whose opinions I value very much, and I read your paper myself with interest. I salute the reviewers for their time and effort.*  *Overall, the reviewers and I certainly saw merit in this work. For example, reviewers wrote that this is a “promising approach,” “the general aim is worthwhile,” and “conceptually clean and sensible.” That said, the reviewers and I noted a number of issues with the paper that should be addressed before this project moves forward. Therefore, I have decided to award a decision of major revision. If you choose to submit a new version of this registered report plan, please address all reviewer comments. I will note those I deem most important to focus on below.* |
| Ed Response |  | |
|  |  |  |
|  |  |  |
| *Ed Comment 1* |  | ***Theory*** |
| *Ed Comment 1.1.1* |  | *First, it strikes me that the literature review and theorising is undercooked in this paper, with a number of relevant findings not mentioned despite important implications for the current paradigm, especially CLT papers that show findings inconsistent with the proposed theory. How do the authors account for those findings? I have become increasingly skeptical of CLT as the years pass after many failures getting CLT predictions to consistently work, and like the reviewer I would frankly be surprised if you obtain clear replicable patterns from these manipulations. Naturally, this is not necessarily a reason to avoid the study—after all, I could be proven wrong, and things may well work as expected—but I worry about how clearly readers can interpret findings using these manipulations, especially in case of null effects—what would we learn then exactly? Perhaps you can say that about many studies, but this design seems fairly likely to result in null effects, and it is unclear to me the evidentiary value that would be obtained if they arose. What then—the paper should be clearer about what we stand to learn depending on the patterns of data.* |
| Ed Response 1.1.1 | We are grateful to the editor and to the reviewers for highlighting the rich and complex work examining CLT and morality. We appreciate that there are inconsistencies in the patters of results observed in the literature. We have expanded our introduction to incorporate and discuss these inconsistent findings. We have also expanded the scope of the current project to examine participants’ judgments (in addition to reason-giving), and derived additional hypotheses in relation to participants’ judgments. Our expanded hypotheses are informed by and address the inconsistency in the literature (e.g., H2 has two competing predictions). We have also attempted to reconcile the competing predictions by examining one possible moderator of the effect of construal level on moral judgments and on reason-giving (ambiguity). Finally, at the end of the introduction, we address the limitations and inconsistencies in the literature and specifically address the possibility of null findings through ineffective manipulation(s).  The relevant passage reads:  We note two potential limitations associated with the hypotheses as presented above. First, it assumes that our manipulations will be successful in influencing construal level, and that these changes in construal level will have downstream effects on rates of reason-giving, and on moral judgments. However, the effectiveness and reliability of construal level manipulations has come under question in recent years, with multiple failed replications (for discussion, see Maier et al., 2022). A recent meta-analysis failed to find strong evidence for construal level theory effects when controlling for publication bias (Maier et al., 2022). Second, our hypotheses assume that the effects of two different manipulations will be comparable. While there is some evidence to support this assumption (e.g., Amaral & Jiao, 2023; Eyal et al., 2009; Lammers, 2012; Ledgerwood et al., 2010), it is also the case that different manipulations can lead to different effects (Alper, 2020; Žeželj & Jokić, 2014). The studies presented by Žeželj and Jokić particularly notable because they show inconsistencies in the direction of effects across manipulations, as well as showing null effects in come cases.  To mitigate these limitations, and ensure that meaningful conclusions can be taken from our findings (even in the case of null results), we have included manipulation checks in both studies. These means we can determine if any null findings are a consequence of the manipulation failing or if they provide evidence that the manipulation did not affect the measure(s) of interest. We will also investigate the effect of construal level on judgments made, in addition to our primary measure of interest (reason-giving), further helping us to make meaningful conclusions from any null results that we may observe. | |
|  |  |  |
| *Ed Comment 1.1.2* |  | *Relatedly, reviewers cited several papers that should give pause when using CLT as a theory for this work. Indeed, the current lit review is far too thin and too biased to give the reader an accurate impression of the state of construal theory with regard to morality. In fact, one can find multiple different CLT papers showing the exact opposite impacts on dilemma decision-making and moral decisions in general—most decent papers in this area identify at least one or two key moderators that completely flip effects, whereas the current paper seems to assume there will be a simple main effect across all cases, measures, and so on. Further care should be taken in discussing this literature, considering that CLT research has been plagued with accusations of publication bias, inconsistent meta-analytic results, and much of the strongest evidence (including evidence cited here) comes from papers that were, um, ‘pre-replication crisis’ and should be taken with a grain of salt, not to mention some authored by people later accused of doctoring data. Although some theorists have attempted to reconcile these inconsistencies, I suggest that current state of the art thinking here is that CLT by itself may be inappropriately generalizing across a series of many different types of thinking. Rather than ‘abstract construal’ generally, there is the fact that outcomes can seem abstract (unless they are concrete!), moral rules can seem abstract, harm can seem concrete, and so on. I suggest it may be important to drill down and really conceptually clarify ‘what exactly’ is being processes in an abstract or concrete way according to the thinking here?*  *Maier, M., Bartoš, F., Oh, M., Wagenmakers, E. J., Shanks, D., & Harris, A. (2022). Adjusting for publication bias reveals that evidence for and size of construal level theory effects is substantially overestimated. PsyArXiv Preprints.*  *Gamliel, E., Kreiner, H., & McElroy, T. (2017). The effect of construal level on unethical behavior. The Journal of Social Psychology, 157(2), 211-222.*  *Agerstrom, J., & Bjorklund, F. (2009). Moral concerns are greater for temporally distant events and are moderated by value strength. Social Cognition, 27, 261282.*  *Amaral, N. B., & Jiao, J. (2023). Responses to ethical scenarios: The impact of trade-off salience on competing construal level effects. Journal of Business Ethics, 1-18.*  *Xiao, W., Wu, Q., Yang, Q., Zhou, L., Jiang, Y., Zhang, J., ... & Peng, J. (2015). Moral hypocrisy on the basis of construal level: to be a utilitarian personal decision maker or to be a moral advisor?. PloS one, 10(2), e0117540.*  *Hofer, M., Tamborini, R., & Ryffel, F. A. (2021). Between a rock and a hard place: The role of moral intuitions and social distance in determining moral judgments of an agent in a moral dilemma. Journal of Media Psychology: Theories, Methods, and Applications, 33(3), 103–112. https://doi.org/10.1027/1864-1105/a000284*  *Gamliel, E., Kreiner, H., & McElroy, T. (2017). The effect of construal level on unethical behavior. The Journal of Social Psychology, 157(2), 211-222.*  *Körner, A., & Volk, S. (2014). Concrete and abstract ways to deontology: Cognitive capacity moderates construal level effects on moral judgments. Journal of Experimental Social Psychology, 55, 139-145.*  *Alper, S. (2020). Explaining the complex effect of construal level on moral and political attitudes. Current Directions in Psychological Science, 29(2), 115-120.* |
| Ed Response 1.1.2 | We are grateful to the reviewers and the editor for highlighting the need to expand on our literature review and for providing helpful pointers regarding key considerations. We also thank you and the reviewers for suggesting specific papers to be considered as we approached our revision.  We found the varying and inconsistent findings intriguing and attempting to make sense of them for the current project required considerable reflection. Based on this reflection, our revised manuscript now includes a direct examination moral judgments in addition to reason-giving/dumbfounding. We now articulate specific hypotheses regarding participants’ moral judgments. We have also developed a more detailed set of hypotheses for reason-giving/dumbfounding (which remains our primary outcome of interest). Some of these additional hypotheses are exploratory, they are informed by our predictions regarding moral judgment, and they attempt to address/incorporate the varied findings in the literature highlighted by the reviewers and editor, e.g., competing hypotheses in H2, while H3 attempts to reconcile these competing hypotheses. We applied similar logic when developing H4 and H5. | |
|  |  |  |
| *Ed Comment 1.2* |  | *Second, the paper says the current model, “assumes that making a moral judgment is an intuitive/habitual response,” without much support for this claim. This seems to oversimplify things quite a bit. Beyond the fact that most philosopher-types talk about moral reasoning, even papers like Cushman (2013) cited here do not seem to go so far as to make this claim—rather, there is generally discussion of qualitatively or quantitatively different kinds of moral judgments being made under intuitive or deliberative processing. Furthermore, this phrasing implies that any judgment derived from deliberation is therefore not moral? I do not think most theorists would agree. This section of the paper needs rephrasing for enhanced clarity and better grounding in the literature. Consider:*  *Monin, B., Pizarro, D. A., & Beer, J. S. (2007). Deciding versus reacting: Conceptions of moral judgment and the reason-affect debate. Review of general psychology, 11(2), 99-111.* |
| Ed Response 1.2 | We thank the editor for highlighting this imprecision and oversimplification in our original submission. We also thank the editor for pointing us towards this paper to assist us in improving this section. We agree that the wording in our original submission appeared to overgeneralize the nature of all moral judgments.  In our revised submission we have qualified that we are referring specifically to moral judgments within the moral dumbfounding paradigm. We specify that we regard these responses to be in line with model-free responses described by Cushman (2013), and we support this characterization by drawing on the distinction between first person recommendations vs third person judging of others described by Monin et al., (2007). The revised passage now reads:  This approach assumes that making a moral judgment in the dumbfounding paradigm involves an intuitive/habitual response (in line with model-free responses described by Cushman, 2013; or habitualized categorizations described by McHugh et al., 2022). In support of this view, participants in studies of moral dumbfounding are typically asked to judge the (often shocking) behavior of others, rather than being asked to compare possible outcomes and make recommend a decision from a first person perspective (see Monin et al., 2007). | |
|  |  |  |
| *Ed Comment 1.3* |  | *Third, interpretation of results: As one reviewer noted, “I found the authors explanation of why "nothing wrong" should count as reflecting a moderate degree of deliberation quite uncompelling. In fact, I wonder whether these answers are less likely to track level of deliberation.” Indeed, I am unclear how to appropriately interpret this response with regard to reasoning or dumbfounding.* |
| Ed Response 1.3 | We thank the editor and the reviewer for highlighting this concern. We agree that attempting to characterize “nothing-wrong” is not straightforward. In our revised manuscript we are more careful in how we walk the reader through the logic of expecting “nothing-wrong” to involve more deliberation than dumbfounding, but less deliberation than reason-giving. We also explicitly acknowledge that the complexity of attempting to characterize this response. Furthermore (and in line with the suggestions from the reviewer and the editor) we note that any effects on participants judgments would likely confound our interpretation of effects on reason-giving/dumbfounding (because of the effect on nothing-wrong). We now include the following regarding mitigating this:  Hypothesis 1 is based on a straightforward mapping of the responses in the dumbfounding paradigm both to relative amount of deliberation and to construal level. However, this does not account for the possible relationship between construal level and moral judgments more generally. This poses a particular challenge for attempting to incorporate nothing-wrong responses, that is, if construal level influences people’s judgments, our results in relation to nothing-wrong responses would be confounded by this. To account for this we will take the following steps. First we will explicitly examine participants’ judgments using a measure that is separate from the reason-giving/dumbfounding/nothing-wrong measure. Second, we will conduct all analyses twice, where our first analyses includes all participants, and the second analyses excludes participants who selected nothing-wrong to directly examine reason-giving against dumbfounding. | |
|  |  |  |
| *Ed Comment 1.4* |  | *I should add that the original dumbfounding studies used a rather different paradigm—it is not that participants voiced no reasons for their judgments—but rather that experimenters noted the (many) reasons they provided were not valid within the scenario. It is unclear to me how the current paradigm deals with such cases—can't participants indicate they ‘have reasons’ but they are thinking of invalid reasons, like ‘It’s wrong because the government should step in instead’? See Bennis et al., (2010) for a discussion of the ‘closed world assumptions’ involved in the kinds of scenarios used here and the challenge of getting participants to adhere to the exact standards stipulated in the scenarios. I would like to see some consideration of this possibility and perhaps a method of dending against it.*  *Bennis, W. M., Medin, D. L., & Bartels, D. M. (2010). The costs and benefits of calculation and moral rules. Perspectives on Psychological Science, 5(2), 187–202.* [*https://doi.org/10.1177/1745691610362354*](https://doi.org/10.1177/1745691610362354) |
| Ed Response 1.4 | We are grateful to the editor for highlighting this limitation with our approach. In our original submission we took a more conservative measure of dumbfounding (participants are only identified as dumbfounded if they admit to not having reasons). This conservative measure has very high face validity and is therefore relatively robust to criticisms of the paradigm. We agree that this conservative measure does not account for participants who claim to have reasons but who fail to provide valid reasons (in the spirit of the original studies).  To account for these participants, in our revised submission, we now include an additional measure of reason-giving/dumbfounding where we code the responses provided by participants who indicated that they can provide reasons.  We note that we remain relatively conservative in our coding. For the majority of reasons provided we will not adjudicate on their validity – if participants provide a reason that they believe is valid there is no reason to judge them as dumbfounded. In line with McHugh et al., (2017) we will identify two classes of response as dumbfounded responding: : unsupported declarations (“*it’s just wrong*”) and tautological responses (“*because it’s incest*”).  McHugh, C., McGann, M., Igou, E. R., & Kinsella, E. L. (2017). Searching for Moral Dumbfounding: Identifying Measurable Indicators of Moral Dumbfounding. *Collabra: Psychology*, *3*(1), 1–24. <https://doi.org/10.1525/collabra.79> | |
|  |  |  |
| *Ed Comment 2* |  | ***Methods*** |
| *Ed Comment 2.1* |  | *A reviewer argued, “I am not sure why a 2 x 3 design is needed to cross temporal and social distance.” I tend to agree—the need for this cross remains mysterious to me, and especially the possibility of an interaction, which seems not fully considered in the current design. This relates back to the CLT point earlier—what exactly do we learn from temporal versus social distance, are they expected to work the same (even Trope and Liberman 2010) admit they may operate differently at times, and what happens when we combine them? Overall, more precision in theorizing please or else maybe a different design.* |
| Ed Response 2.1 | We thank the editor and the reviewer for highlighting the concerns regarding the 2 × 3 design. In line with the reviewer suggestion, our revised submission proposes two separate studies, testing two different manipulations of construal level independently. Study 1 employs a revised temporal distance manipulation, informed by previous research manipulating temporal in morality. Study 2 employs the “how-and-why” task (suggested by Reviewer 1). (in response to other concerns raised, we have removed the social distance manipulation). | |
|  |  |  |
| *Ed Comment 2.2* |  | *Re: “Social distance will be manipulated by asking participants to think about the moral dilemma for themselves (control) or from the perspective of another person (increased social distance).” This manipulation confounds many other factors beyond mere social distance. Moral decisions by myself, as opposed to someone else, may well have important consequences for me, may reflect upon my character, may involve more social desirability, etc. I am not convinced this is a clean manipulation of the construct in play. As one reviewer put it, “The social distance manipulation is fatally confounded with the identity of the target.” They offer a suggested alternative that may side-step much of the issue.* |
| Ed Response 2.2 | We are grateful to the editor and to the Reviewer for highlighting, and clearly articulating the potential issues with employing a social distance manipulation in a moral judgment/dumbfounding task. We agree that the confounding influences and possible sources of error mean the social distance manipulation may not be appropriate for the proposed studies. We thank the Reviewer for their suggested alternative. We have removed the social distance manipulation and instead employ the “how-and-why” task suggested by Reviewer 1. | |
|  |  |  |
| *Ed Comment 2.3* |  | *Reviewers also argued for shifting temporal elements. One argued that “the assignment, not only the deadline should be further away in time to avoid the possibility of imagining finishing early” and another, “imagining that the assignment is due in several hours seems not very time-pressured to me. I wonder if the effect would be cleaner by shrinking that time window.” I won’t give any particular direction here as I trust you see the general issues raised and can shift things accordingly.* |
| Ed Response 2.3 | We appreciate the concerns raised regarding the temporal distance manipulation proposed in the original submission, and we thank the editor and the reviewers for highlighting these. To address these concerns we have adopted a temporal distance manipulation that has previously been used in studies involving moral judgments (Žeželj, & Jokić, 2014). The “deadline” aspect (and its associated confounds) has been removed, and participants are now asked to imagine the *event* occurring either tomorrow or in a year.  Žeželj, I. L., & Jokić, B. R. (2014). Replication of Experiments Evaluating Impact of Psychological Distance on Moral Judgment. *Social Psychology*. <https://econtent.hogrefe.com/doi/10.1027/1864-9335/a000188> | |
|  |  |  |
| *Ed Comment 2.4* |  | *A reviewer noted that when answering the dumbfounding questions in social distance, it is not clear whether I am answering from the perspective of "Alex"--this may elicit a whole set of responses you were not envisioning. I recommend clarifying this to participants and to readers.* |
| Ed Response 2.4 | We are grateful to the reviewer and to the editor for highlighting this as a source of confusion in the original submission. Our revised submission no longer includes a social distance manipulation so all responses are from the perspective of the participant. | |
|  |  |  |
| *Ed Comment 2.5* |  | *As reviewers noted, the paper would really benefit from manipulation checks.* |
| Ed Response 2.5 | We thank the reviewers and the editor for highlighting the need for manipulation checks. In our revised manuscript we propose 2 studies, with manipulation checks in each.  In Study 1 (temporal distance manipulation), in line with previous research using a similar manipulation, our manipulation check will ask ‘participants to rate on a seven-point Likert scale “how far into the future did you imagine this event?”1 = *very close/very soon*, 7 = *very far/very distant*’ p. XX  In Study 2 (how-and-why task) we include a manipulation check similar to previous studies using a version of this task. The revised manuscript now includes the following description of the manipulation check (p. XX):  As a manipulation check, two judges who are unaware of the condition will rate the abstractness of responses to the second part of the how-and-why task. Any response that is identified as a subordinate means to the original statement will be assigned a score of -1, while any response that is a superordinate end of the original statement will be assigned a score of +1 (responses not meeting either criterion will be assigned a score of 0). Ratings for the four responses will be summed to create construal level index with a potential range of -4 to +4 where higher scores represent higher levels of construal (see Fujita et al., 2006).  Fujita, K., Trope, Y., Liberman, N., & Levin-Sagi, M. (2006). Construal Levels and Self-Control. *Journal of Personality and Social Psychology*, *90*(3), 351–367. <https://doi.org/10.1037/0022-3514.90.3.351> | |
|  |  |  |
| *Ed Comment 2.6* |  | *Reviewers also noted that “(Alex) is not really described at all, but subjects will be told to think about the scenario from Alex's perspective. By contrast, in the pliot studies we got a description of a student who was described as a good student who usually understands the material. The description prompts taking a cognitive approach to the scenario to be described. I didn't see anything in the materials explaining why the switch in approach was made.” I agree—consider revising this element of the procedure or better defend and support this decision.* |
| Ed Response 2.6 | We thank the editor and the reviewers for this question. As noted above, in response to this, and other concerns raised, our revised manuscript no longer includes a social distance manipulation. The description of Alex has been removed, so these concerns regarding potential confounds in the description no longer apply in our revised submission. | |
|  |  |  |
| *Ed Comment 2.7* |  | *Further, reviews mentioned “it would have been nice to see more of a defence of the convenience sampling plan” and “Is there a plan to equalize samples per condition by site? To examine site effects and interactions with the factors of interest?” Indeed, there is not much theoretical hay made out of the fact that you apparently plan to recruit hundreds of people in different languages from different places. I gather this is done in the name of power? But there does not seem to be consideration for how that might affect things, especially in light of aspects like the foreign language effect, which not only influences moral judgment, but also may impact things like cognitive effort and construal. This paper needs more thought and care put into sampling and decisions surrounding it to persuade the reader that results will be interpretable in light of these plans.* |
| Ed Response 2.7 | We thank the editor and the reviewers for highlighting the need for more detail and a more robust sampling plan. We agree that the multi-site collaborative aspect of the project was not well presented in the previous version of the manuscript.  Our revised manuscript provides a more careful and systematic approach. We propose 2 studies. The Irish team will conduct both studies in Ireland, and this Irish data will be sufficient to test the hypotheses.  The Irish data will be supplemented with data from collaborating teams in Poland, UK, Chile, USA, and Belgium. Collaborating teams will conduct either Study 1 or Study 2 (or both studies where resources allow) in full in their respective countries; e.g., Study 1a Irish data (target *N* = 685), Study 1b Polish data (target *N* = 685), Study 1c (target *N* = 685) etc.  Our revised analysis plan will treat the separate sites as separate studies and aggregated results will be collated by conducting internal meta-analyses (with any pooled analyses being used for exploratory purposes only) | |
|  |  |  |
| *Ed Comment 2.8* |  | *Reviewers also raised questions about the line mentioning, “Participants will then be given an opportunity to provide reasons for their judgment” following the initial DVs—this would seem to come before the ‘key slide’ DV that analyses focus on? This would seem to pre-emt the point of the dumbfounding DV—most/all participants will have already been prompted to think of reasons, which may reduce or eliminate dumbfounding, the point of the study.* |
| Ed Response 2.8 | We thank the editor and the reviewer for noting this and for highlighting it as a potential source of error for the current design. Our methods were informed by previous research which has asked participants to provide reasons immediately after their initial judgments (e.g., McHugh et al. 2020). We note that in this previous work these initial responses were coded for specific content. This coding of initial responses is not relevant to the research question we are examining here, and we agree that this prompt to provide reasons could be a potential source of error for the current studies and as such we have removed this from the proposed studies in the revised manuscript.  McHugh, C., McGann, M., Igou, E. R., & Kinsella, E. L. (2020). Reasons or rationalizations: The role of principles in the moral dumbfounding paradigm. *Journal of Behavioral Decision Making*, *33*(3), 376–392. <https://doi.org/10.1002/bdm.2167> | |
|  |  |  |
| *Ed Comment 2.9* |  | *Furthermore, how will you analyse the DVs that are not the focal DV? The paper mentioned multiple DVs that do not seem to appear in the analysis plan. Moreover, the analysis is sparse in detail—it should be clear exactly what patterns you expect to find for each key DV and how the reader should interpret these patterns. For example, what pattern is expected for ratings of how moral actions are?* |
| Ed Response 2.9 | We thank the editor for highlighting the need for more specificity in our predictions and in our analyses. Our revised manuscript now includes moral judgment as an additional variable of interest, and we have now derived specific hypotheses in relation to this measure – and these hypotheses informed the development of more specific hypotheses in relation to our primary measure of interest (reason-giving).  We also now include a much more detailed analysis plan, specifying the models that we plan to run to test our hypotheses (and we have tested these analyses on simulated data). We clearly label the analysis of the additional measures (confidence, confusion, gut feeling etc.) as exploratory, and we do not provide specific predictions in relation to these measures. | |
|  |  |  |
| *Ed Comment 3* |  | ***Minor Points*** |
| *Ed Comment 3.1* |  | *A reviewer mentioned, “The only example in the paper given of what would count as a qualifying reason for reason-giving was a clearly utilitarian”--indeed, it would be useful to clarify what kinds of reasons you anticipate, including a variety of types of reasons, and also examples of something that may not qualify as a moral reason per se.* |
| Ed Response 3.1 | We thank the editor and reviewer for raising this concern. In our revised manuscript we now include a plan to code the open-ended responses provided by participants who indicate that they can provide a reason. In describing the coding we clearly articulate that we will take a very strict approach to coding responses as dumbfounded, specifically only unsupported declarations (“*it’s just wrong*”) and tautological responses (“*because it’s incest*”) will be taken as dumbfounded responses. We also acknowledge that there is some disagreement in the literature regarding whether such responses count as dumbfounding, and as in light of this we treat the coded response measure as a secondary measure, for follow-up analyses. In describing the we provide some examples of possible responses that we will accept as reasons:  These responses (unsupported declarations/tautological responses) are the only responses that we will classify as dumbfounding, and as such, we anticipate that the majority of responses will be acceptable as reasons, e.g., participants may invoke religion (“*because God doesn’t permit it*”), the law (“*because it’s illegal*”), or social norms (“*because it is unacceptable in our society*”). | |
|  |  |  |
| *Ed Comment 3.2* |  | *Your opening referring to politicians describing the right thing to do sensible cites Skitka 2010, but reminded me very much of a paper I was missing from your references:*  *Kreps, T. A., & Monin, B. (2014). Core values versus common sense: Consequentialist views appear less rooted in morality. Personality and Social Psychology Bulletin, 40(11), 1529-1542.*  *You may want to incorporate this into your opening argument, but I leave this up to you.* |
| Ed Response 3.2 | We thank the editor for highlighting this interesting paper and its relevance for our work. We have incorporated it into our opening paragraph, with a sentence that reads:  Interestingly, it is not even clear if a position would be strengthened by the inclusion of reasons; research has shown that including consequentialist justifications can reduce the perceived moral relevance of an argument, potentially reducing its moral appeal (Kreps & Monin, 2014). | |
|  |  |  |
| *Ed Feedback* |  | *In sum, the reviewers and I see potential in this project, but are requesting richer theory and method modifications to ensure that you have a clear plan to test exactly what you aim to test and the results will be clear to readers.*  *I hope you find this feedback fair and useful in revising this registered report plan for the next phase. Please let me know if you need clarification on any points.* |
| Ed Response |  | |
|  |  |  |
| **Reviewer 1** |  |  |
| *R1 Feedback* |  | *Thank you for the opportunity to review the registration report for Stage 1, 'Space to Think: Testing the Effect of Distancing on Moral Dumbfounding.' I believe that the core idea of examining the moral dumbfounding phenomenon through the lens of Construal Level Theory is a promising approach.*  *However, I have some methodological concerns that temper my enthusiasm for the paper. Below, I include comments and questions.* |
| R1 Response | We thank you for your encouraging assessment of the overall aim of this work. Below we address the methodological concerns point-by-point. | |
|  |  |  |
| *R1 Comment 1* |  | *(1) The authors should check the methodology of these two replication studies falsifying CLT:*  *Gong, H., & Medin, D. L. (2012). Construal levels and moral judgment: Some complications. Judgment and Decision making, 7(5), 628-638.*  *Žeželj, I. L., & Jokić, B. R. (2014). Replication of experiments evaluating impact of psychological distance on moral judgment. Social Psychology.* |
| R1 Response 1 | We thank the reviewer for highlighting these two papers. We agree that they are very relevant for the current project, and that there is a clear need to incorporate and accommodate these findings within our work. | |
|  |  |  |
| *R1 Comment 2* |  | *(2) For instance, Zezelj and Jokic (2014) asked participants to evaluate the moral wrongness of the classic three scenarios (eating one’s dead pet, sexual intercourse with sibling, dusting with national flag). The researchers manipulated temporal distance by asking participants to imagine that the event will happen “the next day” or “the next year”. Temporal distance had no effect on moral wrongness judgments; in Study 3, they tested again and could not verify the effect of temporal distance. However, social distance manipulations were found to be effective.*  *The authors aimed to manipulate temporal distance on pg. 8 “ In the control condition, participants will not be provided with any instructions about a time frame; in the increased temporal distance condition, participants will be asked to imagine responding in the distant future (five weeks from now); in the reduced distance condition, participants will be asked to imagine responding in the near future (two to three hours from now).”*  *Even though DVs are different, Zezelj and Jokic examined moral wrongness judgments; however, the authors want to examine reason-giving, and I still doubt the effectiveness of this manipulation.* |
| R1 Response 2 | We agree that this limited effect of temporal distance in Žeželj and Jokić (2014) poses a challenge to the design as originally proposed. We have modified our temporal distance manipulation in two ways to strengthen it: first we have removed the “assignment” aspect, participants are asked to imagine the actual behaviour describe occurring at different times (rather than imagining completing an assignment with different deadlines); second, there is a larger contrast between near (previously 2-3 hours; now tomorrow) and the far (previously 5 weeks; now in a year) conditions.  While these changes make the temporal distance manipulation stronger, there is still a possibility that we observe null findings in line with Žeželj and Jokić (2014). We now present more detailed hypotheses that accounts for opposing effects.. We also include an additional passage where we discuss the limitations and the possibility of null findings. The passage reads:  We note two potential limitations associated with the hypotheses as presented above. First, it assumes that our manipulations will be successful in influencing construal level, and that these changes in construal level will have downstream effects on rates of reason-giving, and on moral judgments. However, the effectiveness and reliability of construal level manipulations has come under question in recent years, with multiple failed replications (for discussion, see Maier et al., 2022). A recent meta-analysis failed to find strong evidence for construal level theory effects when controlling for publication bias (Maier et al., 2022). Second, our hypotheses assume that the effects of two different manipulations will be comparable. While there is some evidence to support this assumption (e.g., Amaral & Jiao, 2023; Eyal et al., 2009; Lammers, 2012; Ledgerwood et al., 2010), it is also the case that different manipulations can lead to different effects (Alper, 2020; Žeželj & Jokić, 2014). The studies presented by Žeželj and Jokić particularly notable because they show inconsistencies in the direction of effects across manipulations, as well as showing null effects in come cases.  To mitigate these limitations, and ensure that meaningful conclusions can be taken from our findings (even in the case of null results), we have included manipulation checks in both studies. These means we can determine if any null findings are a consequence of the manipulation failing or if they provide evidence that the manipulation did not affect the measure(s) of interest. We will also investigate the effect of construal level on judgments made, in addition to our primary measure of interest (reason-giving), further helping us to make meaningful conclusions from any null results that we may observe. | |
|  |  |  |
| *R1 Comment 3* |  | *(3) A manipulation check is necessary to ensure that temporal distance manipulation increases reason-giving. Then, testing this idea in cases that trigger moral dumbfounding would be better. Another option would be testing the manipulations of social distance, how-why technique, and temporal distance on reason-giving. Then, the most effective ones might be used in the main study.* |
| R1 Response 3 | We thank the reviewer for these helpful suggestions.  Regarding the manipulation check suggestion, whether or not distancing increases reason-giving is an open question that this work aims to investigate. We use four different scenarios, all of which have been used in previous studies of moral dumbfounding, but that have shown differences in the rates of dumbfounding they elicit – *Julie and Mark* (*incest*) and *Jennifer* (*cannibal*) elicit dumbfounding more reliably than *Trolley* and *Heinz*. In this way, the inclusion of *Trolley* and *Heinz* serve a similar function to the reviewer’s suggestion of including a manipulation check. In line with this, the studies proposed in our revised manuscript now present all four scenarios to all participants (in the previous version, participants were randomly presented with only one of the four scenarios).  We are also grateful to the reviewer for suggesting pretesting social distance, temporal distance, and the how-why technique separately, with a follow-up main study examining the most effective ones. In response to other reviewer comments highlighting issues with manipulating social distance within a moral judgment task, we opted to drop the social distance manipulation in our revised manuscript. We have replaced it with the “how-and-why” task (suggested by the reviewer – thank you). Our revised manuscript now includes two separate studies, Study 1 tests temporal distance (3 levels: increased, decreased, control) while study 2 tests a direct manipulation of construal level using the “how-and-why” task. We had previously piloted versions of the temporal distance (and social distance) manipulations (reported in the supplementary materials) however these pilots were underpowered and we could not draw meaningful conclusions. To avoid this we propose two “main” studies in our revised manuscript, both studies will be sufficiently powered to detect a medium effect size.  We also thank the reviewer for highlighting the need to include a manipulation check. Both studies now include manipulation checks.  In Study 1 (temporal distance manipulation), in line with previous research using a similar manipulation, our manipulation check will ask ‘participants to rate on a seven-point Likert scale “how far into the future did you imagine this event?”1 = *very close/very soon*, 7 = *very far/very distant*’ p. XX  In Study 2 (how-and-why task) we include a manipulation check similar to previous studies using a version of this task. The revised manuscript now includes the following description of the manipulation check (p. XX):  As a manipulation check, two judges who are unaware of the condition will rate the abstractness of responses to the second part of the how-and-why task. Any response that is identified as a subordinate means to the original statement will be assigned a score of -1, while any response that is a superordinate end of the original statement will be assigned a score of +1 (responses not meeting either criterion will be assigned a score of 0). Ratings for the four responses will be summed to create construal level index with a potential range of -4 to +4 where higher scores represent higher levels of construal (see Fujita et al., 2006).  Fujita, K., Trope, Y., Liberman, N., & Levin-Sagi, M. (2006). Construal Levels and Self-Control. *Journal of Personality and Social Psychology*, *90*(3), 351–367. <https://doi.org/10.1037/0022-3514.90.3.351>  Gamliel, E., Kreiner, H., & McElroy, T. (2017). The effect of construal level on unethical behavior. *The Journal of Social Psychology*, *157*(2), 211–222. <https://doi.org/10.1080/00224545.2016.1208139> | |
|  |  |  |
| *R1 Comment 4* |  | *(4) On pg. 11-12 “Participants will be asked to rate on a 7-point Likert scale how right or wrong the behavior described is (where 1 = morally wrong; 4 = neutral; 7 = morally right), as well as their confidence in this judgment (where 1 = not at all confident; 7 = extremely confident). Participants will then be given an opportunity to provide reasons for their judgment” I do not understand here. Following the evaluation of the scenario, will participants be allowed to provide reasons for their judgment? Why?*   1. *If yes, this might interfere with the manipulation effect as people are equipped with reason-based thinking. So, the question is whether the high-reason-giving results are obtained from CLT manipulation or by asking participants’ reasons before they provide reasons.* 2. *Will participants provide their answers in response to this question -julie&mark scenario-, e.g. “And do you accept that they are both consenting adults, and that they both consented and enjoyed it?” Is this question asked to test whether participants understood/read the scenario?* |
| R1 Response 4 | We thank the reviewer for noting this and for highlighting the prompt to provide reasons as a potential source of error for the current design.  The reason for including this was based on previous research where participants were to provide reasons immediately after their initial judgments and these responses were coded for specific content (e.g., McHugh et al. 2020). However this coding of initial responses is not relevant to the research question we are examining here and therefore this prompt to provide reasons is not needed in the current study. Furthermore, we agree that this prompt to provide reasons could be a potential source of error for the current studies and as such we have removed this from the proposed studies in the revised manuscript.  Regarding a. We thank the reviewer for highlighting this potential source of error, and we agree this may interfere with the manipulation, and in response to this we have removed this question.  Regarding b. Participants will provide yes/no responses to the counter arguments (e.g., “And do you accept that they are both consenting adults, and that they both consented and enjoyed it?”). The purpose of these questions is not to collect participants’ responses, rather the questions are intended to encourage participants to engage more fully with the counter arguments being presented such that participants may begin to question the validity of reasons they may hold for judging the behaviour as wrong. By requiring a response (rather than just having participants read the counter-arguments) participants will take a position on the validity of the counter-arguments (whether they agree or disagree), and potentially re-evaluate their own positions based on these counter-arguments.  Participants’ responses will be recorded and this may provide the basis of future exploratory secondary analysis, however these responses will not be included in the primary analyses for the current research. This approach is consistent with previous studies using this method for studying dumbfounding.  McHugh, C., McGann, M., Igou, E. R., & Kinsella, E. L. (2020). Reasons or rationalizations: The role of principles in the moral dumbfounding paradigm. *Journal of Behavioral Decision Making*, *33*(3), 376–392. <https://doi.org/10.1002/bdm.2167> | |
|  |  |  |
| *R1 Comment 5* |  | *(5) I also do not understand why participants will be asked whether they change their minds. Changing minds is a measure that can be tested with a within-subjects-design. However, if I do not miss anything here, participants will be randomly assigned to manipulation conditions. If the authors aim to compare participants’ responses under “counter-arguments” and “critical slide,” this also seems problematic to me. As I do not grasp what is measured with these responses. Let’s say I am one of the participants, and I respond, “No, I do not accept what they did, but yes, they consented and enjoyed the act as it is written in the scenario.” to your question, “do you accept that they are both consenting adults, and that they both consented and enjoyed it?” Does it assess whether I understand the acts in the scenario or my point of view?* |
| R1 Response 5 | We thank the reviewer for these queries relating to (a) asking participants if they changed their mind, and (b) responses to the counter arguments.   1. Asking participants if they changed their minds   The procedure for eliciting dumbfounding in a computerized task adopted here is based on previous studies of the phenomenon. In these previous studies a range of additional measures were included for exploratory purposes – e.g., is there any relationship between incidences of dumbfounding and whether or not a participant self-reports changing their mind. In adapting the procedure for the current study we retained all questions/measures in our original submission. Based on Reviewer feedback we have reflected on the relevance of these measures and agree that asking participants if they changed their minds is unnecessary for the current study. As such we have removed this question in our revised submission.   1. Responses to the counter arguments   Regarding the responses to the counter-arguments, previous studies using this procedure have typically not examined the responses provided to the counter-arguments. These responses are also not typically included any analyses.  As highlighted in our response to the previous comment (R1 Response 4),  The purpose of including a response option is to ensure that participants read, engage with, and think about, the content of the counter-arguments. We will record the responses and these may form the basis of future exploratory secondary analysis, but they are not relevant for the current studies. | |
|  |  |  |
| *R1 Comment 6* |  | *(6) Theory-driven question: On pg. 6, it is written that “Applied to the available responses in the dumbfounding paradigm, providing a reason for a moral judgment (reason giving) involves the most deliberation; providing a dumbfounded response (dumbfounding) involves the least amount of deliberation; and selecting “There is nothing wrong” (nothing wrong) involves less deliberation than reason-giving, but more deliberation than dumbfounding (deliberation may lead participants to accept the counter-arguments, and revise their judgment rather than deliberate further to identify alternative reasons).”*  *For instance, will participants who are dumbfounded say “there is nothing wrong” with the incest scenario? Haidt explained that if the intuitions are highly strong, people are dumbfounded and find the act unacceptable. If there are possible alternatives to making the act acceptable, then people will not be dumbfounded as they can think reason-based. Dumbfounded people can also find reasons to support their intuitions, so the problem is not finding a reason but a counter-reason. So, the main question is, which reasons will the authors look at? Is High-reason-giving automatically considered that people think reason-based and less dumbfounded?*  *One possible solution would be to look at “reasons supporting the intuition (with intuition, I mean disapproving or approving the act in a given scenario)” and “reasons falsifying the intuition” across manipulation conditions.* |
| R1 Response 6 | We thank the reviewer for this question and for highlighting one of the challenges inherent in trying to study moral dumbfounding. We agree that a person experiencing dumbfounding may indeed provide “other” responses, including saying there is nothing wrong, or searching for (possibly invalid) reasons to support their judgment. This kind of measurement error is unavoidable (we have even seen it in interview studies, where, only once the study is over some participants admit to having been dumbfounded), and we regard the reported rates of dumbfounded responding in published literature on the conservative side.  We also agree that there is an asymmetry regarding the kinds of reasons participants are asked to provide, i.e., participants are typically asked to defend (provide reasons for) their disapproval of an act (and participants are not asked to defend/provide reasons for approving an act).  In our original submission we did not intend to look at the reasons provided by participants, if participants claimed to have reasons for their judgment we would take this claim at face-value and assume that the participants were not experiencing dumbfounding. Based on the feedback received as part of this review we have revised this plan. | |
|  |  |  |
| *R1 Feedback* |  | *Overall, I hope my questions and comments do not overwhelm the authors. While I am impressed with the core idea, I am confused by certain methodological and theoretical aspects.* |
| R1 Response |  | |
|  |  |  |
| **Reviewer 2** |  |  |
| *R2 Feedback* |  | *I should say up front that I am a philosopher, not a psychologist, even if one who has done a lot of interdisciplinary work. That said, I thought that setup of the study is very conceptually clean and sensible. I have a few questions to flag but nothing that makes me question the value of doing the study.* |
| R2 Response |  | |
|  |  |  |
| *R2 Comment 1* |  | *So, here are a few mild questions or concerns:*  *- I notice that in the study the team wants to run the hypothetical character in the experimental conditions (Alex) is not really described at all, but subjects will be told to think about the scenario from Alex's perspective. By contrast, in the pliot studies we got a description of a student who was described as a good student who usually understands the material. The description prompts taking a cognitive approach to the scenario to be described. I didn't see anything in the materials explaining why the switch in approach was made between the pilot studies and the study to be proposed. I could imagine a few possible reasons, but it would be good to see that defended. As it stands, I do wonder whether one of two things will happen when subjects are asked to reflect as if they are Alex. Possibility #1 is that because we are not told anything about Alex that subjects will give the same answers as they gave in the control condition. The other possibility I'm wondering about is whether the lack of information will lead to uncontrolled mental modeling of Alex (what would an average male college student think?) as opposed to just spurring reflection. Regardless, the pivot to underdescribed Alex from the model philosophy student seemed in need of explanation if not defense.* |
| R2 Response 1 | We thank the reviewer for raising this concern. Our rationale for including less detail in the description of Alex was to reduce the potential confounds within the description. Along with the concerns noted here, several other issues with the social distance manipulation were raised and in response we have removed this manipulation entirely. Our revised submission now includes two studies, Study 1 uses a temporal distance manipulation, and Study 2 uses the “how-and-why” task. The removal of the social distance manipulation eliminates the concerns regarding mental modelling of Alex. | |
|  |  |  |
| *R2 Comment 2* |  | *- Similarly, it would have been nice to see more of a defense of the convenience sampling plan. One would imagine that convenience sampling is going to get one mostly college students albeit perhaps an international set of the same. If there is any realistic hope of getting a less homogenous sample through convenience sampling, this was not explained. One would imagine that the ease and facility of pivoting from an intuitive to a deliberative mode of cognition would interact relevantly with age.* |
| R2 Response 2 | We thank the reviewer for highlighting the need for more detail and a more robust sampling plan. We agree that the multi-site collaborative aspect of the project was not well presented in the previous version of the manuscript.  Our revised manuscript provides a more careful and systematic approach. We propose 2 studies. The Irish team will conduct both studies in Ireland, and this Irish data will be sufficient to test the hypotheses. This Irish data will involve a combination of convenience sampling and online research participation systems. The revised sampling includes the following:  Participants in Study 1a will be recruited through a combination of convenience and snowball sampling in Ireland. Some participants may be recruited using the SONA research participation system where college students will receive course credit in return for their participation. Where resources allow, additional participants may be recruited through online research participation systems (e.g., Prolific / Lucid / CloudResearch / MTurk) if necessary, specifically targeting older demographics to balance out the student samples. This combination of convenience sampling and online recruitment will help us to achieve a better balance of generalizability, and statistical power within resource constraints.  The Irish data will be supplemented with data from collaborating teams in Poland, UK, Chile, USA, and Belgium. Collaborating teams will conduct either Study 1 or Study 2 (or both studies where resources allow) in full in their respective countries; e.g., Study 1a Irish data (target *N* = 685), Study 1b Polish data (target *N* = 685), Study 1c (target *N* = 685) etc.  Our revised analysis plan will treat the separate sites as separate studies and aggregated results will be collated by conducting internal meta-analyses (with any pooled analyses being used for exploratory purposes only) | |
|  |  |  |
| *R2 Comment 3* |  | *-The only example in the paper given of what would count as a qualifying reason for reason-giving was a clearly utilitarian one related to the trolley problem. That's a fine example to use, but it would be a bit more reassuring as a philosopher to see explicitly in the paper examples of, for instance, deontological reasons as well or an explicit statement that reasons from different ethical perspectives would count as instances of deliberative moral reasoning.* |
| R2 Response 3 | We thank the reviewer for raising this concern and apologize for any lack of clarity in our previous submission. In most recent studies of moral dumbfounding the researchers do not adjudicate the validity of reasons provided by participants, if participants declare that they can provide reasons, this response is typically taken at face value and participants are only identified as dumbfounded if they admit to not having a reason. This was the proposed approach in the original submission.  That said, in our revised submission we have now adopted another approach. Where participants indicate that they can provide a reason for their judgments we will code these reasons for responses that may be indicative of dumbfounding. Based on McHugh et al., (2017) we have selected 2 such responses: unsupported declarations (“*it’s just wrong*”) and tautological responses (“*because it’s incest*”). We acknowledge that there is some disagreement in the literature regarding whether such responses count as dumbfounding, and as in light of this we treat the coded response measure as a secondary measure, for follow-up analyses, with the primary analyses focusing on whether participants declared they had reasons or not. In describing the coding of reasons we make it clear that these unsupported declarations/tautological responses are the only kinds of responses that will count as dumbfounding, and we provide examples of possible responses that we will accept as reasons:  These responses (unsupported declarations/tautological responses) are the only responses that we will classify as dumbfounding, and as such, we anticipate that the majority of responses will be acceptable as reasons, e.g., participants may invoke religion (“*because God doesn’t permit it*”), the law (“*because it’s illegal*”), or social norms (“*because it is unacceptable in our society*”).  McHugh, C., McGann, M., Igou, E. R., & Kinsella, E. L. (2017). Searching for Moral Dumbfounding: Identifying Measurable Indicators of Moral Dumbfounding. *Collabra: Psychology*, *3*(1), 1–24. <https://doi.org/10.1525/collabra.79> | |
|  |  |  |
| *R2 Comment 4* |  | *- I imagine that the analyses will be run and reported in a way that makes this point moot, but I found the authors explanation of why "nothing wrong" should count as reflecting a moderate degree of deliberation quite uncompelling. In fact, I wonder whether these answers are less likely to track level of deliberation and more likely to track some other factor. Someone who reverses course after reporting that something is wrong under the pressure of counter-argument might, for instance, simply be high on agreeableness or high on conflict aversion.* |
| R2 Response 4 | We thank the reviewer for highlighting this additional potential source of error regarding how nothing-wrong responses should be understood. In response to this comment we have made the following changes:  First, we acknowledge the complexity of attempting to characterize this response, and we are more careful in how we walk the reader through the logic of expecting “nothing-wrong” to involve more deliberation than dumbfounding, but less deliberation than reason-giving.  Second, we note that any effects on participants’ judgments are likely to confound our interpretation of effects on reason-giving/dumbfounding (because of the effect on nothing-wrong). We now include the following regarding mitigating this:  Hypothesis 1 is based on a straightforward mapping of the responses in the dumbfounding paradigm both to relative amount of deliberation and to construal level. However, this does not account for the possible relationship between construal level and moral judgments more generally. This poses a particular challenge for attempting to incorporate nothing-wrong responses, that is, if construal level influences people’s judgments, our results in relation to nothing-wrong responses would be confounded by this. To account for this we will take the following steps. First we will explicitly examine participants’ judgments using a measure that is separate from the reason-giving/dumbfounding/nothing-wrong measure. Second, we will conduct all analyses twice, where our first analyses includes all participants, and the second analyses excludes participants who selected nothing-wrong to directly examine reason-giving against dumbfounding. | |
|  |  |  |
| *R2 Comment 5* |  | *- I wondered whether the near term temporal condition was near enough. Especially if one just has to have an opinion about a scenario and perhaps provide a reason for one's opinion, imagining that the assignment is due in several hours seems not very time-pressured to me. I wonder if the effect would be cleaner by shrinking that time window.* |
| *R2 Response 5* | We agree that the previous temporal distance manipulation may not have been adequate to manipulate construal level in the manner intended. We previously framed the manipulation as part of an assignment and varied the deadline (this was intended to make the combined social/temporal distance manipulations more plausible). Removing the social distance manipulation removed some of the plausibility constraints associated with the temporal distance manipulation and as such we have removed the “assignment” aspect of the manipulation entirely. The temporal distance manipulation now asks participants to imagine the *events* as happening either tomorrow or a year from now.  We acknowledge that this remains a larger window (tomorrow is further away than a few hours), however this is in line with previous research using similar manipulations (e.g., Žeželj & Jokić). We believe this is mitigated in the new design in two ways: (i) the increased distance condition has been moved further into the future (now one year, previously five weeks), and (ii) the target of the distance manipulation is now the actions/events the participants are imagining (rather than a deadline on an assignment), that is, participants are now thinking about *events* in the near/distant future.  Žeželj, I. L., & Jokić, B. R. (2014). Replication of Experiments Evaluating Impact of Psychological Distance on Moral Judgment. *Social Psychology*. <https://econtent.hogrefe.com/doi/10.1027/1864-9335/a000188> | |
|  |  |  |
| **Reviewer 3** |  |  |
| *R3 Feedback* |  | *On the face, this Registered Report proposal seems to be a perfectly reasonable A + B type of research, where manipulation A (construal level) is applied to outcome B (moral dumbfounding) to generate a research question. While I think the general aim is worthwhile, adding knowledge about the cognitive conditions that are conducive to the dumbfounding phenomenon, there are some details that do require comment and possibly course correction before the research is approved.* |
| R3 Response | We thank the reviewer for their positive evaluation of the worth of the proposed project. We are also very grateful for their considered and insightful suggestions for improvement. We have indeed taken the suggestion of a “course correction” and believe the revised manuscript is much stronger as a result. | |
|  |  |  |
| *R3 Comment 1* |  | *1. This is actually a positive point the authors missed. If I had reviewed the previous paper where cognitive load was applied to dumbfounding, I would worry that the role of the load would be ambiguous. Does it influence the ability to express reasons, or actually the ability/motivation to process reasons? For example, some examples of dumbfounding might be due to overthinking (processing) which load might reduce, even as it also increases dumbfounding due to judgment that there isn't enough time to work on really good emotions. These manipulations are more contextual/cognitive in nature so do not have this fault.* |
| R3 Response 1 | We thank the reviewer for highlighting this strength of the current work. We now include a short passage describing this strength of the proposed studies as addressing the ambiguity limitation of the cited cognitive load study. The addes passage reads:  The proposed studies not only extend and build on McHugh McGann, et al., (2023), but also address a limitation associated with this work. Specifically, it is unclear if manipulating cognitive load inhibited participants’ ability to provide reasons, or if participants under cognitive load were simply less motivated to provide reasons. The manipulations proposed here do not have the same limitation. | |
|  |  |  |
| *R3 Comment 2* |  | *2. I am not sure why a 2 x 3 design is needed to cross temporal and social distance. If there are no well-founded predictions for the interactions, one could just have a 2 + 3 design (i.e. 2 different studies side by side) and save 1/6 of the participants. There don't seem to be good predictions for the interaction, and the influence of one factor could just end up making the main effect less interpretable as compared to the simple effect in the other factor's control group.* |
| R3 Response 2 | We thank the reviewer for highlighting this. We have taken on board the suggestion of different studies and in our revised manuscript we now propose two separate studies.  Study 1 manipulates temporal distance (with three levels, increased, decreased, control; and a modified distance manipulation in response to other comments).  Study 2 manipulates construal level using a mindset manipulation (how-and-why) task (with two levels abstract and concrete). We replaced the social distance manipulation with this task in response to other comments. | |
|  |  |  |
| *R3 Comment 3* |  | *3. The revelation that participants are to be sourced from 3 different sites in as many languages took me by surprise. Is there a plan to equalize samples per condition by site? To examine site effects and interactions with the factors of interest? More detail is needed, and if these three samples are to be mashed together, a defense of this approach.* |
| R3 Response 3 | We thank the reviewer for highlighting the need for more detail and a more robust sampling plan. We agree that the multi-site collaborative aspect of the project was not well presented in the previous version of the manuscript.  Our revised manuscript provides a more careful and systematic approach. We propose 2 studies. The Irish team will conduct both studies in Ireland, and this Irish data will be sufficient to test the hypotheses.  The Irish data will be supplemented with data from collaborating teams in Poland, UK, Chile, USA, and Belgium. Collaborating teams will conduct either Study 1 or Study 2 (or both studies where resources allow) in full in their respective countries; e.g., Study 1a Irish data (target *N* = 685), Study 1b Polish data (target *N* = 685), Study 1c (target *N* = 685) etc.  Our revised analysis plan will treat the separate sites as separate studies and aggregated results will be collated by conducting internal meta-analyses (with any pooled analyses being used for exploratory purposes only) | |
|  |  |  |
| *R3 Comment 4* |  | *4. The social distance manipulation is fatally confounded with the identity of the target. Presumably any effects could be attributed to role-taking of an ethicist - either as someone who should have an answer, or someone who would overthink the solution and so become flummoxed. If you imagined social distance to a caveman you might get a different effect. A cleaner manipulation would be to imagine someone like you but in another (culturally similar) country or a different part of the same country. Even so, the temptation might be to double-guess what kind of mentality that person would have and give a different answer from your own, which is not exactly the same as a manipulation of distance per se. I can't see a way around the interpretive problems social distance and for this reason I would simply drop this condition or choose something else like spatial distance.* |
| R3 Response 4 | We appreciate this critical analysis of our originally proposed social distance manipulation, and agree with the concerns raised. We thank the reviewer for the suggestion of dropping this manipulation in favour of something else. We have adopted this suggestion and our revised manuscript now includes a second study where we manipulate construal level directly using a mindset manipulation (how-and-why task). The decision to use this task was based on its use (and the use of similar tasks) in previous studies of construal level and moral judgment (e.g., Žeželj & Jokić, 2014; Gong & Medin, 2012). We note that this task was also suggested by Reviewer 1.  Gong, H., & Medin, D. L. (2012). Construal levels and moral judgment: Some complications. *Judgment and Decision Making*, *7*(5), 628–638. <https://doi.org/10.1017/S1930297500006343>  Žeželj, I. L., & Jokić, B. R. (2014). Replication of Experiments Evaluating Impact of Psychological Distance on Moral Judgment. *Social Psychology*. <https://econtent.hogrefe.com/doi/10.1027/1864-9335/a000188> | |
|  |  |  |
| *R3 Comment 5* |  | *5. A related problem, more concrete: when answering the dumbfounding questions in social distance, it is not clear whether I am answering from the perspective of "Alex" (in which case the response should be worded respecting Alex) or my own (in which case keep the "I" wording).* |
| R3 Response 5 | We thank the reviewer for highlighting this as a potential source of confusion. Our revised manuscript no longer included a social distance manipulation, and as such all responses are from the perspective of the participant, and this is reflected in the updated materials. | |
|  |  |  |
| *R3 Comment 6* |  | *6. In the temporal manipulation, the assignment, not only the deadline should be further away in time. It is easy and socially desirable to imagine myself virtuously completing the assignment sooner than it is due.* |
| R3 Response 6 | We thank the reviewer for highlighting this, and agree that this does present a potential confound in our original temporal distance manipulation. As part of our revisions we changed the temporal distance manipulation such that it no longer involves an assignment, instead participants are asked to imagine the events they are judging as happening either tomorrow (near) or in a year (distant). | |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Once again we thank the reviewers for their insightful and helpful feedback, and we thank the editor for the opportunity to revise this manuscript. We believe the revised manuscript is much stronger, and we hope our revisions have addressed all your concerns. Thanks again. The authors.