**Editor comments**

Summarizing remarks: In summary, I think this is a very interesting paper and I imagine it will have an important impact on the field if you can collect additional data to address the issues raised above.  Thus, I hope you will collect new data and revise this for further consideration at SPPS.  I realize this may be more substantial revision than you are willing to conduct, but I hope you will consider it.   I look forward to receiving your revision.

*Response: We are very grateful for all the helpful feedback we received and for your encouraging remarks (as well as that of the reviewers) about this work. We were also encouraged to read that both reviewers shared your assessment that the topic was important and contributing to the scientific literature on charitable giving.*

*We have indeed collected additional data in a new, pre-registered experiment. We have also added a number of new analyses to further examine the robustness of the different effect (e.g. using Bayesian ANOVAs; implemented in JASP in line with Love et al., 2015;, and Rouder et al., 2012). Finally, we have substantially rewritten the manuscript to address the comments from you and the reviewers. We hope you find the revision satisfactory.*

Comment 1: My reading of the paper is that one of the key results is the null effect of information about charity efficiency and effectiveness. I think this is an interesting and important finding that is worth reporting. However, if this is in fact the focal effect, it is important that uninteresting alternative explanations (e.g., ineffective manipulations) be ruled out. Reviewer 1 raised some important concerns about the efficacy of the manipulations/measures used in the studies. First, they noted that the manipulation may not have worked because participants didn’t pay enough attention. Although participants who failed attention checks were excluded from analyses, I am not completely convinced that correctly answering the attention check items (e.g., “to monitor quality, please respond with a seven for this item”) demonstrates that the participant also carefully read the information about the charity. For this reason, I think the present manuscript could be strengthened by an additional study (preferably pre-registered) that includes a manipulation check that clearly distinguishes between participants who have, and have not, absorbed the information about efficiency and effectiveness.

*Response: This is a point well taken. We will discuss your concerns separately from those from reviewer 1, as they have slightly different foci (the reviewer’s remarks are mainly focused on a lack of attentiveness due to “long and tiring” text materials; see response to reviewer 1, comment 2). As for your more specific comments, we agree that our attention checks might not capture how well participants attend to the manipulated materials in particular. We also agree that an additional, pre-registered study would be beneficial to better evaluate the impact of effectiveness.*

*We ran a pre-registered experiment with the aim of providing a clearer, albeit brief, effectiveness information manipulation. In this experiment we also asked participants about their recall from this text (from a segment appearing in all conditions, to also test whether participants paid equal attention across conditions).*

*This study focused on donations for blind people and participants could choose from two charities, one focusing on treating river blindness and the other focusing on providing guide dogs. In the effectiveness condition we provided information about the costs and outcomes of each program. The manipulation check asked about where the efforts to combat river blindness are focused (Africa and Latin America; again, this information was mentioned in both conditions). 84% of the participants of the participants reported one of the two regions, of which 44% recalled both regions correctly. There was no significant difference between the effectiveness and control condition in this recall (coded as 0 = no region correct, 1 = one region correct, 2 0 both regions correct); X2 = 0.55 (df = 2), p = .76. More important, the impact of effectiveness information on donations did not differ significantly by degree of attentiveness. In the revised manuscript, we report these results in the supplementary materials (along with additional results for attentiveness and the image manipulation).*

*We also conducted additional Bayesian analysis to provide a more formal evaluation of the probability that the null hypothesis is true given the observed data, as compared to a model positing an effect of efficiency/effectiveness information. We did this for each study for the whole distribution of donation responses. The findings are reported along with the ANOVA results in Table 1. In every single study, the null hypothesis was more likely, often considerably so, compared to the alternative (impact of effectiveness info). In fact, the null hypothesis was at least 6 times as likely (strong evidence for the null hypothesis according to rules of thumb for interpreting Bayes factors) and those “best odds” represent a case in which the average donation was lower with effectiveness information (study 6).*

*The Bayesian analyses do not speak to alternative explanations why we observed the null so consistently (e.g. lack of attentiveness). However, they do corroborate the intuitive conclusion that across a variety of manipulations of efficiency and effectiveness there is no evidence that effectiveness information leads to greater donations. In fact, we tested 5 different manipulations (positive efficiency [study 1], negative effectiveness [study 2&3], conjoint efficiency comparison [study 4], outcome effectiveness [study 5], and conjoint effectiveness comparison [study 6]) in reasonably large samples, and two study populations (Mturk & Prolific). Furthermore, in our meta-analysis (Table 1, final column) the estimated correlation coefficient is small, with fairly-tight confidence intervals.*

Comment 2:Reviewer 1 also points out that the image used in Studies 1-4 to manipulate identifiability (of Omran Dagneesh) may also have been familiar to many participants, as it was widely publicized. This could present a confound, as you may have been manipulating familiarity as well as identifiability. Again, a new study could use an alternative image, and help to rule out this potential confound.

*Response: We appreciate the comment that familiarity could contribute to donations when the manipulations involve images like the one in Study 1-4. We therefore used an unfamiliar image in the new experiment. It is also important to point that we also used an unfamiliar image in study 5. That study indicated an image effect that was very similar to those observed in studies 1-4. See response to reviewer 1, comment 3, for more details.*

Comment 3:  
Reviewer 1 also makes what I think is a valuable suggestion regarding the operationalization of “empathic concern.” They note that the specific photo chosen may be more likely to evoke personal distress than empathic concern (an important distinction in the empathy literature), and that this possibility seems to be supported by participants indicating that they felt “sad,” “devastated,” and “horrified,” in response to the images. Thus, it may make sense to conceptualize some of your results as effects on personal distress rather than empathic concern.

*Response: We agree that some of the reported emotions fit with descriptions of personal distress and we appreciate that the reviewer brought this up. At the same time, we believe it is somewhat oversimplified to argue that evidence of personal distress constitutes evidence against empathy. We also believe this could be more of an issue with how we phrased our question about emotions, rather than the experimental effects. For a more detailed response, see comment 4 by reviewer 1.*

Comment 4: I, like Reviewer 2, found the evidence for the interaction to be relatively weak. This interaction was found to be significant in Study 2, but not in the other four studies. Although the meta-analysis found a significant interaction, I share Reviewer 2’s concerns that internal meta-analyses can lead to misleading results (and in particular a greater risk of false-positives), especially when the set of studies is not pre-registered. To me, it seemed that this interaction was not the main focus of the paper, and for this reason I’m not particularly concerned that it might not be a robust effect. I think, however, that the interpretation of this effect could be even more tentative.

*Response:* *We agree that the interaction should be interpreted with more care. The Bayesian analysis along with the new data further suggests that our findings are quite plausible under the null hypothesis of no interaction (and consistently more plausible than the alternative hypothesis that there is an interactive effect). We have revised the discussion accordingly (see pp. 21).*

Comment 4:Finally, I thought that Reviewer 2 had some very useful suggestions regarding the dual process framing throughout the paper. They point out that this framing may currently be overly simplistic, and may place these two “systems” on a false continuum.   
  
*Response: We appreciate this comment as well, and it is true that could be more specific with our terminology. We have updated the manuscript throughout to avoid discussions about two broad systems. Even if we do not fully concur with the reviewer’s criticism (see responses to reviewer 2, comment 2 and 3), we believe a broader discussion on the validity of dual-process models is beyond the scope of the paper. Practically speaking, we agree with the reviewer that it is easier to just be more specific and not to get dragged into the broader discussion in the paper.*

**Reviewer 1 comments**

Summarizing remarks:In this paper the authors present five studies examining the combined effect of victims' identifiability and information about the efficiency and effectiveness of charities on people's donation decisions. Overall, results demonstrate a strong identifiability effect on willingness to donate as well as on donation amounts. Efficiency and effectiveness did not have a significant effect on donations. There is some evidence, however, that information on efficiency and effectiveness may attenuate the identifiability effect. The paper addresses an important, timely topic, while employing different designs to examine it (in five different studies). Therefore, it has the potential to contribute to the field of charitable giving and to be of interest for readers of the journal.   
  
*Response: We are happy that the reviewer found our paper interesting and timely. We are also glad to read that s/he thought it would be of interest for readers of this journal and the literature on charitable giving.*

Comment 1: I appreciate the authors efforts to be succinct in the description of the methods of the five studies, to fit with the journal's short format. However, I believe more information is needed to clarify important nuances that may help interpret the results. For example, what was the exact framing of the donation request? Was it identical in all conditions (e.g. would you be willing to donate to a charity operating in Syria?) or was it different in the identified condition? Which image was used in study 5? I think that a short description of each of the studies in terms of the experimental design, procedure and material would make it easier to follow the different methods.

*Response: We appreciate the general comment that more information would be desirable and the specific questions are indeed important ones. Since it is difficult to fit this in the manuscript we have added a longer description of all the studies in the supplementary materials, along with a table that provides an overview of important methodological choices in each study (e.g. image used etc.). We have also provided an example of the donation request and in the method descriptions for each study and we describe the minor changes from study to study (it was always the same across conditions within each study). These materials should answer the specific questions and we hope they will help other readers as well. We appreciate the request as it will hopefully make for a more compelling article.*

Comment 2*:* Regarding the efficiency/effectiveness manipulations, it seems that participants were asked to read a long description of the cause. This may decrease the participant's attention to the specific information that distinguishes between the conditions (the information regarding the efficiency/effectiveness). I am afraid that participants did not read all the detailed information and therefore missed the manipulation. This may provide an alternative explanation to the lack of an overall efficiency/effectiveness effect. It may also increase the effect of the image on donations, since when the text is long and tiring, participants may rely more heavily on the image.

*Response:* *We understand this concern with regards to study 5 (effectiveness information for Polio vaccines), but it should be emphasized that studies 1-4 did not include long descriptions of this kind. In all of those studies, the efficiency information was presented succinctly in a graphic format (using Charity Navigator ratings). As an illustration, the efficiency manipulation from study 4 was provided in the supplementary materials. Experiments 1-3 involved similar manipulations, except they only saw the ratings for one charity. When we manipulated negative information (study 2 and 3) it was done with a news headline and an image, again containing minimal text. As such, we do not believe these manipulations contained too much information, or were long and tiring. If anything, one might argue that the manipulations in study 1-4 were too minimal to have an effect (hence the more elaborate text in study 5). Nevertheless, a new study (see study 6 in the paper) was designed to provide clearer, albeit brief effectiveness information. The results from the last study also failed to provide evidence for effectiveness information influencing donation decisions. Overall then, the results were similar regardless if the efficiency/effectiveness information was presented succinctly or not.*

Comment 3: If I understand correctly, the identifiability manipulation in studies 1-4 included the photo and the name of the injured Syrian boy in the ambulance (Omran Daqneesh), an image that, as the authors noted, was highly publicized. Moreover, this photo along with the photo of Aylan Qurdi (lying dead near a Turkish beach) received much attention by the media, and became symbols of the humanitarian crisis in Syria. I was wondering whether in addition to the identifiability manipulation per-se, this image also manipulated familiarity with the victim, as well as the memory of the general public reaction to it at the time it was published (for example, see the general reaction to the photo of Aylen Kurdi, Slovic et al 2017).

*Response: It is a good point that the image effects in study 1-4 could be driven by familiarity, in addition to identifiability. On the other hand, study 5 indicated similar results, despite the use of an unfamiliar image. However, study 6 did not reveal an image effect, so this study does not bolster our conclusions about identifiable victim images. Still, the identified victim effects (in the absence of effectiveness information) seem fairly robust in studies by other researchers (see Lee & Feeley, 2016), so this is perhaps not the most critical point of this paper.*

Comment 4: Importantly, this specific photo may evoke feeling of personal distress, rather than empathy, in the perceiver - as the boy is looking shocked and is covered with blood. This type of shocking stimuli is not typical to studies on the identifiable victim effect (which usually use more neutral photos). The authors write that they were using an empathy-inducing image; however, the dominant emotions mentioned by the participants in their study were “sad”, “devastated” and “horrified”, which seem to be more self-oriented emotions, representing distress from seeing the victim, rather than empathy. Empathy is a more other oriented emotion, which may include compassion, sympathy and caring. Since the authors report measuring empathic concern, it would be good to show the difference between the identified and the unidentified conditions in this measure. I would encourage the authors to address the limitations of their stimuli and how it may affect their results.  

*Response: This is another good point and we agree that the distinction is relevant to make. At the same time, the seminal work that distinguishes empathic and distressed reactions and behaviors suggest that people will mainly look for a way out of the situation if they primarily experience distress (Batson et al., 1981). If it is easy to relieve the distress in other ways than helping, then distressed individuals are unlikely to help (e.g., Toi & Batson, 1982; see also Dovidio et al., 2012 for a review). In our experiments there are number of easier alternatives. In particular, as the helping situation was anonymous and took place online, a distressed person could easily click past the donation question or quit the study to escape the situation. In other words, there is little reason to expect images to generate donations in this context, unless they also triggered emotions associated with a desire to help (i.e. compassion/sympathy). In these experiments it seems clear that people had a reaction to the images, and this reaction motivated them to help – even though simpler options were available to alleviate personal discomfort.*

*The research on empathy and personal distress also suggests that these emotions are substantially correlated (e.g., Dovidio et al., 1990) and it has been proposed that most people start with a distressed reaction and then that can (sometimes) transform into compassion etc. (e.g., Dickert et al., 2011). From such a perspective, the real question is whether the initial negative state transforms into an other-oriented focus or not. When we asked participants about their emotional reaction, most of them might still have been at that first stage, or had this in mind when answering the question (after all, we asked them to report their first reaction). In retrospect, we realize that our measure might not have been a convincing way to assess if they reacted in the “right” emotional way – it was mainly aimed to test if they reacted emotionally at all (and not completely misplaced – e.g., expressing fear). We now briefly discuss this issue in the manuscript (see Procedures and Materials).*

*We understand why the reviewer would consider looking at the measure of empathic concern to evaluate if the manipulation worked as intended or not. The problem is that the instrument we used is meant to assess trait empathy (i.e. stable individual differences) rather than state empathy (intra-individual, changeable emotional reactions). A measure of state empathy (e.g., Cikara et al., 2014) might have worked as a manipulation check, but a null-effect on a trait measure could just reflect that the measure focuses on aspects of empathy that are fairly stable (and thereby non-susceptible to experimental manipulation).*

*That said, we tried to address the concern by the reviewer in our new experiment. Study 6 included a question directly aimed at distinguishing the signature of empathic (I want to help) and distressed (I want out) reactions. Again, the image manipulation did not seem to work as intended in this study, so it is difficult for us to use these data to draw conclusions about participants’ experiences of empathy and personal distress.*

Comment 5: On the same note, it could also be interesting to code the emotions participants mentioned as self vs. other oriented, to learn more about the type of emotions that were raised by the specific victim (content analysis).

*Response: We agree that could be interesting to do content analysis to get a more fine-grained understanding of those responses, but it would be beyond the scope (and word limit) of this paper to do so. For reasons discussed in our previous response (how we asked the question, and how distress can turn to empathy), it is also uncertain how informative such an analysis would be for distinguishing empathy and personal distress.*  
  
Comment 6 (under other comments): There is research suggesting that people do care about the impact or effectiveness of their donation decisions that not mentioned in the paper. For example, prospective donors prefer to help more (rather than less) people with the same amount of donations (in a comparative context, Duncan, 2004); people are less motivated to help when overhead costs are high (Sargeant & Woodliffe, 2007); campaigns that are close to reaching fundraising goals receive more donations (Cryder, Loewenstein, & Seltman, 2013); providing tangible details about a charity’s interventions increases donations, by increasing perceptions of their impact (Cryder, Loewenstein, & Scheines, 2013).

*Response: We focused on research examining how people react to factual information about the efficiency or effectiveness of a charity, rather than subjectively wanting help to be meaningful (as Duncan, 2004 put it: “someone who wants to personally ‘make a difference’”). That is, there is a difference between a subjective sense of effectiveness (a property of the donor) and objective effectiveness (a property of the charity). The references cited above are dealing with the subjective aspect. For example, donating when a charity is close to reaching its goal (Cryder, Loewenstein, & Seltman, 2013) is informative about the fact that people want to feel like they personally make a difference, but it is uninformative about the effectiveness of the charity (in fact, in many cases late donations could be argued to typically be less impactful - they have lower marginal utility [Steinberg & Morris, 2010] ).*

Comment 7 (under other comments): On the top of page 3, the authors write: "Many studies have shown that people give more to help one featured victim than many unknown ones. This has been called the identified victim effect (e.g., Kogut & Ritov, 2005; Lee & Feeley, 2016; Small, Loewenstein, & Slovic, 2007)". I believe this is inaccurate. The identified victim effect refers to people's greater willing to help an identified victim about whom they have some information (like a name or a photo), compared with an unidentified (anonymous) victim. The preference to help an identified single victim over a group of victims (whether victims are identified or not) is called "the singularity effect (of identified victims)", see Kogut & Ritov, 2005; Small, Loewenstein, & Slovic, 2007, Slovic, 2007).   
  
*Response: We agree that it would be helpful to distinguish the numerical aspect from the question of identifiability, but we do not find this to be clearly expressed in previous definitions of the identifiable victim effect (e.g., in the sources cited by the reviewer). Kogut and Ritov (2005) make such a distinction, but none of the other sources do. The other sources tend to compare identified victims with “statistical victims” (e.g., Lee & Feeley, 2016; Small, Loewenstein, & Slovic, 2007). Statistical victims refers to multiple people, both in the writing and empirical operationalizations by these scholars.*

*That said, to better match how scholars have defined the identified victim effect in the past, we have now changed the writing to directly match the definition given by Lee and Feeley (2016) in their meta-analysis for this phenomenon (“The identifiable victim effect (IVE) refers to individuals’ tendency to offer greater aid to specific, identifiable victims than to anonymous, statistical victims”, p. 199). More important, in our empirical studies we do not compare the image (identified) condition with a condition with many victims, but a rather a “blank” control that entails no specific victims. We believe that primarily speaks to the role of identifiability, rather than numerical variation. In any case, we are grateful that the reviewer brought this up, because we believe the literature could benefit from a discussion about what the identifiable victim effect really entails.*

**Reviewer 2**   
  
This paper, Empathic and Numerate Giving: The Joint Effects of Victim Images and Charity Evaluations, presents five studies examining whether manipulations of victim identifiability interact with statistical information regarding charity effectiveness to influence helping, with an eye towards implications for effective altruism. The paper has several strengths. The topic of the paper is important and relevant to both scientific discourse and everyday life. The five studies and meta-analysis presented represent a lot of work, and I applaud the authors’ use of real money and real charities, which enhances the importance and value of these studies. I also appreciate the large sample sizes and the use of meta-analysis to examine results.

*Response: We appreciate the overall positive impression described by the reviewer here.*

Comment 1: Overall, the introduction does a decent job of setting up the main questions and reviewing relevant literature, but one glaring omission would be the various papers by David Rand and colleagues that frequently demonstrate intuitive/emotional/empathic influences on donations decisions, as well as some cases where deliberative process contribute. The paper would be more informative if some discussion of these findings and how they relate to the current work were included.

*Response: We agree there is an interesting parallel between these studies and those by Rand and colleagues. In fact, an earlier version of this paper described that parallel, but we received feedback from peers that it was inappropriate to compare real donation decisions to economic game behaviors mostly focused on mutual cooperation and trust (e.g., public goods game; see Rand, 2016; Rand et al., 2012). We left that out because we had been explicitly advised to do so.*

*As we are more inclined with the current reviewer, we have reintroduced references to Rand’s work. At the same time, we outline the differences between their work and ours (e.g., between reciprocal and non-reciprocal prosociality).*

Comment 2: The authors approach this topic from a dual process perspective, but their discussion of this is overly simplistic. Modern approaches to dual process theory clarify that there is no such thing as ‘system 1’ and ‘system 2’ processing, as there is little to no coherence among the 4-5 features of these systems (e.g., those described by Bargh, 1994). For example, responses may be rapid and uncontrollable, yet available to conscious awareness (e.g., Hahn & Gawronski, 2018). Hence, authors should be clearer about which features of system 1/system 2 processing they are describing rather than treating these as monolithic constructs. At times the authors dive down into specifics, talking about empathy as rapid and spontaneous; they are on much better theoretical ground to focus on such more specifics rather than dual process models generally.

*Response: We agree with the general sentiment of the reviewer’s comments, as well as the advice to be more specific in the paper. On the other hand, we are not entirely convinced that discussions about explicit versus implicit attitudes and stereotypes (the topic of both the Bargh and Hahn/Gawronski reviews) can speak to the validity of the whole dual-process paradigm in cognitive psychology, which also deals with other mental concepts and behaviors. For instance, Evans (2008; psych review) discusses similarities and differences between dual-process models in (a) reasoning, (b) judgement and decision-making, and (c) social cognition. We could provide arguments in our paper for defining system 1 and 2 based on criteria set in the first two domains and especially work on affective heuristics (e.g., Slovic et al., 2007). We could further discuss more specific criteria for defining something as a heuristic, which in turn defines system I in many dual-process models (and this discussion is not altogether reducible to debates about how to define implicit cognition, as described by e.g., Bargh, De Houwer, Gawronski). However, we do not have the space to go into those discussions in detail and we believe it would be better anyway to just be more specific, as the reviewer suggests.*

Comment 3: On a related note, the authors speculate that “Most any deliberation may disrupt empathic giving because it shifts people’s thinking from system I to system II” (p. 6). This claim is odd because it places these systems on a false continuum—it need not be the case that increases in deliberation reduce other forms of thinking; they are conceptually independent and methods that investigate whether they operate independently often confirm that they do operate independently and demonstrate that analyses predicated on treating them as opposites often distort results or conflate processes (e.g., see Payne & Bishara, 2009). I suggest moving away from this language and toward the possibility that empathic and deliberate processes may have independent influences on judgment, possibly interacting influences (as they generally do elsewhere in the paper), but that they need not be conceptualized as competing. Hence, I am a little skeptical of the claim later in the paper that introducing statistical information may ‘disrupt’ the use of empathic considerations (p. 17). I would be open to this possibility were the data on this point more persuasive. However, as my next comment indicates, I am not yet there.

*Response: We agree that we should have been more specific about what we mean by those statements. As for the broader point about the validity of two mental “systems,” we understand where the reviewer is coming from. However, it is also important not to over-generalize findings from one dual-process domain (social cognition) to broad statements about any-and-all research on “system I” and “system II” (see also previous response). Certainly, there are parallels to be made between different lines of dual-process research, but sorting out the similarities and differences is beyond the scope of this paper. In the end, we agree with the reviewer that a broad discussion about dual-process systems is unlikely to benefit the paper anyway.*

Comment 4: The studies are fairly well powered, although the trichotomous nature of the DVs inhibits this somewhat. More concerning, effects are quite small. I remain confident that the main effect of showing an identifiable victim is robust as this pattern emereged for nearly all studies as well as in the meta-analysis. No problem. However, that effect is less interesting than the interaction effect, which emerged only in Study 2, and slightly trended in a couple other studies. I am rather nervous about interpreting such an effect. I appreciate the authors’ decision to conduct a meta-analysis to try to find the best evidence, but recent papers suggest that in cases like this meta-analyses can lead people to end up reporting type I errors as significant and meaningful findings (e.g., Ueno, Fastrich, & Murayama, 2016, JEP:Gen).

*Response: In the original submission we deliberately used words such as “may” and “unclear” to signify the uncertainty about the robustness of this effect. It is a point well taken that we could have been even more careful though. The new data and especially the addition of Bayesian analysis clearly suggest that the evidence is weak. We have updated the discussion accordingly.*

*In terms of the broader point about type I errors and mini-meta analyses, we agree with the argument that p-values should be interpreted cautiously. At the same time, Ueno et al. (2016) also* *emphasize that the core problem has to do with stopping rules and not with the practice of synthesizing effect size estimates.* *It is problematic to run another study just to get a significant aggregate effect, just like it is problematic to collect more data points in a single study to reach a significant result. Ueno et al. commented that “without using the flexible stopping rule, one can reliably keep false positive error rates to the conventional nominal level (e.g., 5%) regardless of how many studies are integrated within an article. It is important to note that Cumming’s (2008, 2012, 2014) ‘new’ statistics and its accompanying recommendation for internal meta-analysis actually emphasized the value of estimation (i.e., CI), and he never recommended making a dichotomous decision based on NHST with internal meta-analyses” (p. 651).*

*We strongly concur with this position, and it should be emphasized that we did not decide on a number of studies after seeing the outcome of the meta-analysis (i.e. the aggregate effect size and p-value did not serve as a stopping rule). We only use the meta-analysis to synthesize the data and to illustrate the aggregate effect and its uncertainty (i.e. effect size and CI) – precisely in line with the recommendations by Cummings and Ueno et al. Nevertheless, we recognize that the aggregate estimate of the interaction might still have been a false positive in classic NHST, as indicated by the new study and the Bayesian analyses. We appreciate that the reviewer raised concerns about the interaction effect and we are happy to report a credible null-finding, rather than a potentially intriguing, but speculative effect.*

Comment 5: Hence, I hate to say it after so much work, but the most interesting claim in this paper regards the interaction effect, and I am personally not yet persuaded by the data that it should be interpreted. Perhaps were the authors to conduct a large scale preregistered study where they make this interaction prediction and then included these data that would be enough for me. I would not even require that real money be used (as such studies can be expensive); I would be content with a similar pattern emerging for something participants view as psychologically valuable. But due to my hesitation to interpret such a weak and inconsistent interaction as psychologically meaningful, and this interaction forming the main contribution of the paper, I am afraid I cannot recommend that SPPS accept this version of the manuscript at this time.

*Response: We are sympathetic to the reviewer’s position on this issue and we did run an additional pre-registered experiment. However, the results were not highly conclusive and we did not quite reach our target N for practical reasons (a lower-than-expected response rate to the Omnibus survey that this was connected to, after several weeks). In any case, we believe the study adds value to the paper, especially for the interpretation of the more questionable effects (e.g. the interaction). In other words, even if the experiment “failed” it still says something about the prospect of finding these effects in an experiment that was designed to satisfy known concerns with the previous studies (e.g., about the image and effectiveness manipulations).*

*At the same time, we would not consider the interaction to be the sole contribution of this paper and we hope that it will not solely be evaluated on the basis of what the findings show in that regard. After all, we submitted this paper to SPPS because we hoped that its editors, reviewers, and readers would see the value in an inquiry that examines relevant theoretical and practical questions – even if the data did not come out “sensational.”*

Comment 6 (minor points): The paper is not always structured in the way I typically expect of an SPPS paper; it seems as though it may have been designed for a business or economics journal and then resubmitted here. For example, hypotheses in APA format are often simply stated as part of the text rather than given specific reference numbers like Q1, and also tend to be phrased in terms of statements rather than questions. I will leave it to the editor to decide to what degree the authors might want to align their structure more closely with typical SPPS structure.   
  
*Response:* *We have rewritten the aim section and hope the reviewer will find it clearer and more aligned with conventions in psychological journals (although we are not aware of any formal APA requirements/suggestions for the exact phrasing of research questions, aims or hypotheses; see e.g. APA manual, 6th ed., p. 28).*

Comment 7 (minor points): There were a fair number of typos and minor grammatical issues that should get resolved before publication.

*Response: We have made a more careful proof-read of the revised manuscript.*

Comment 8 (minor points): The authors should italicize the phrase ‘identified victim effect’ when defining it to make it clear this is a technical term.

*Response: We have now italicized the phrase “identified victim effect.”*

Comment 9 (minor points): It also might be nice to clean up the graphs a little (e.g., delete background lines as per by APA format)

*Response: We have now removed the lines from the graphs.*