Review | Stage 1 submission

Revisiting celebrity contagion and the value of objects: Replication and extensions Registered Report of Newman et al. (2011)

Mannix Chan, Yaqi Jin, Eva Yiyu Chen, Shanshan Peng, Aaron Charlton, Gilad Feldman

https://osf.io/esg2k | *version 1*

Dear all,

Thanks a lot for the opportunity to review this stage 1 submission. Overall, I think that the article is well-chosen and that the theoretical background as well as the justification for choosing this article is reasonable. I also consider the methodology of the replication to be thorough, well-rounded, and close to the original studies. I have however noticed some aspects that may help advance the stage 1 manuscript.

First, I have a few comments concerning the background sections, specifically regarding the alignment of your manuscript vis-à-vis the original paper.

The beginning of the paper and the theoretical setup are a bit hard to follow. Specifically, it was not always clear which elements in the original study you focused on in your manuscript since the "Main hypotheses and key findings in the target article" section is rather short and mostly lists the constructs. Providing more information on the original study vis-à-vis the objectives of the replication could help clarify this section. For example, you mention some analyses (e.g., mediations, moderations) that you do not plan to test. For the contagion sensitivity moderation specifically (H2a and H2b in the original manuscript), you only later explain why you will not test it. The hierarchical regression that is mentioned in the original paper, for instance, is not mentioned at all. Maybe you can specifically expand this section to be clearer about which constructs, relationships, hypotheses, and tests you consider for our replication, which you consider for exploratory purposes, and which you discard from your manuscript.

As a related, but minor point: In Table 1 you specify the main hypotheses by Newman et al. (2011). Experiment 1 however does not directly specify hypotheses in the original article while Experiment 2 overall lists four hypotheses.

There are also instances in which the terminology is not consistent. For example, on p. 13 for Experiment 2, you write about "item valuation" which is "purchase intention" in the original manuscript; or in Table 2, you write market value instead of demand. Can you clarify the construct names in your manuscript vis-à-vis the original paper?

Second, concerning the extensions: I think that the hypotheses overall make sense since contagion should depend on psychological distance but I noticed a few things about their operationalization.

Concerning the *physical contact extension*: The item "How much would you like to meet this person?" does not specify whether there will be physical contact or not. Since many people do shake hands or hug whenever they meet in person, the item wording could be ambiguous since it does not directly exclude physical contact and therefore might miss its objective. Instead, maybe specify that the meeting would be a video meeting or a phone call to ensure that the respondents would be unambiguously aware of the missing physical contact.

Concerning the *temporal proximity extension*: Much of the theoretical background for this section rather refers to physical proximity and not just to temporal proximity. There is also a small mismatch in your argument. On p. 16 you write that "implying a weaker impact of time on negative contagion" which suggests that temporal proximity may not be as effective for negatively (vs. positively) perceived individuals. This is however not reflected in the hypothesis in Table 3.

The manipulation of temporal proximity is also quite obvious. Since the assessment of both close and distance temporal contact will be on the same page, this position could trigger carry-over effects from one measure to the other or at least prompt participants to compare their ratings. I know that you may not like to introduce another between-subject factor here but it might already help to separate the measure on two different pages.

Third, concerning the analyses and reporting:

You chose to apply two approaches to determine the required sample size but discarded the analysis that resulted in n = 736 participants. Is there a specific reason to leave the discarded analysis in the manuscript? If you decide to keep the power analysis in the manuscript, can you elaborate a bit more on the methods? Currently, you only mention the R packages (which is good to ensure that the developers get the appropriate credit), but the rationale for the analysis as well as why you decided on a target sample of n = 1200 remains implicit.

In the manuscript, you specify that you do not plan to exclude any participants which strikes me as odd since excluding inattentive participants may increase data quality. Further, you are collecting multiple variables that could serve as a quality check. Do you, for example, plan to check variables such as the time each participant took to complete the questionnaire? Do you plan to do consistency checks (e.g., if participants indicate that a person is "Not at all famous" in the celebrity condition)? How do you proceed if participants indicate that they were not filling out this questionnaire seriously?

I further really like that you provide the R markdown files with all the code and results. Going through the files, however, I think that there should be more comments or explanations about the analyses. For example, concerning the reliability analyses in the Extensions, it is unclear for which construct the reliability is calculated unless one checks the code directly.

I have some minor points which I noticed when reading the manuscript:

- In the hypotheses, you write about "willingness to contact celebrities and their possessions." "Willingness to contact" sounds a bit like initiating communication via mail, Messenger apps, etc., and not like contagion or actual physical contact. Can you rephrase the wording here?
- On the bottom of p. 11, you forgot to mention "liking" as a DV of the original study. You also excluded liking from Table 2. Is there a specific reason for it?
- Is there a reason not to include the interaction *fame x valence* from Newman's Experiment 1 in Table 2?
- Table 5 (Experiment 2) could be a bit misleading on the experimental design. Since IV2 (physical contact) and IV3 (demand) are placed below each other it seems as if each participant receives information for both IVs.
- For Table 2, could you please add a brief note that explains the multiplier for the required sample size?

- On p. 31, you indicate different effect size estimates for the original effect compared to Table 1.
- Figures 1 (and similar Figures): Could you add to the Figure note what the error bars represent (a 95% CI, I assumed)? It also seems as if the colors for the error bars are not easily distinguishable from the points.
- For Figures 6 and 7, can you please clarify that the y-axis represents the difference in purchase intentions (pleasure)?
- The manuscript includes two Tables 5 (p. 22 and 23)

All the best,

Susanne Adler