**PCI Registered Report Peer Review: Revisiting stigma attributions and reactions to stigma: Replication and extensions of Weiner et al. (1988)**

***Signed, Dr Charlotte Pennington, Aston University, Birmingham, UK.***

**Summary:**

This is a fantastic Stage 1 Registered Report which aims to (conceptually)\* replicate Weiner et al. (1988) to assess whether physically-based stigmas are perceived as less controllable, more stable (irreversible), and associated with more pity, less anger, and more willingness to help compared to mental-behavioral stigmas. It also aims to extend this original study by assessing four additional stigmas that have become prevalent in the last decade and re-assess the original categorizations of stigma sources (origins) by asking participants whether they perceive these as mental-behavioral or physically-based. Overall, I think this is a well-planned study with important implications, and I fully endorse the recognition of replications, and their importance, within psychological science. The proposal meets the PCI RR guidelines for the ***scientific validity of the research questions, the logic, rationale and plausibility of the proposed hypotheses, the soundness of the proposed methodology and analysis pipeline, the clarity of study procedures and analysis pipeline to reduce undisclosed flexibility, and the consideration of sufficient outcome-neutral conditions (e.g., attention checks, MTurk data quality checks).*** Although my suggestions for revision below are extensive, they all relate to minor improvements to clarity and detail rather than anything substantial/major. I must say, I am impressed by the quality of this work and am inspired by the implementation of replications within teaching/pedagogy (this represents a students’ proposed project and it is of extremely high quality).

Before I outline my minor comments below, I just want to provide some general points to reflect upon:

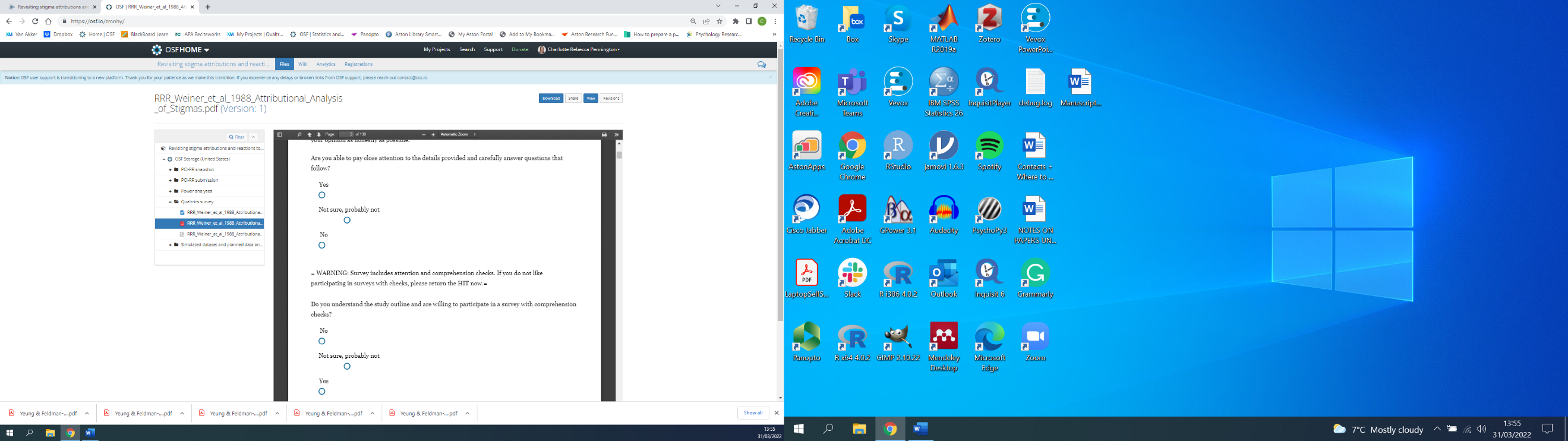
* The authors suggest this is a direct replication, according to LeBel’s guidelines, but in my opinion, it’s a conceptual replication due to the addition of four new stigma labels, a change in the Vietnam War label (which I agree with, but my point remains), slight changes in the procedure and measuring participant’s own cateogorisations of the stigma origins, and the analytical strategy (e.g., removing “order” because it showed no significant effect in the original, but the inclusion of an additional factor in the original will have changed its power and potentially the subsequence interpretations). This might be worth thinking about with regards to the original authors replying to this replication.
* The literature around public stigma towards mental and physical health conditions is vast, and there is a lot of work consistent with the theory and hypotheses proposed in the original work by Weiner et al. (1988). For example, reviews suggest that mental health disorders are more heavily stigmatised than physical health conditions, at both the explicit and implicit level. Do you think it’s important to include some of this more general work to show how vast and, seemingly, consistent this work is? Or do you want to focus only on the replication? Perhaps this detail could be included in the part where you signify the importance and impact of this original study.
* As an exploratory analysis, will you assess which conditions may be most-least stigmatised? For example, a new review by Kilian et al. (2021) suggests that alcohol use disorders are one of the most heavily stigmatised health conditions:
* Kilian, C., Manthey, J., Carr, S., Hanschmidt, F., Rehm, J., Speerforck, S., & Schomerus, G. (2021). Stigmatization of people with alcohol use disorders: An updated systematic review of population studies. *Alcoholism, Clinical & Experimental Research, 45,* 899-991. <https://doi.org/10.1111/acer.14598>

*\*\*To note, unfortunately, I could not initially access the original article by Weiner et al. (1988) because it was behind a paywall that my institution did not have access to. The Recommender – Chris Chambers – kindly sent this to me, but I think it’s worth considering that many other readers might also not be able to access this, so it’s important that the information presented in the current manuscript does not rely on having access to the Weiner et al. article. I note ambiguities in phrasing within my recommended revisions below\*\*.*

**Open materials, code, and data:**

The OSF Page includes simulated power analyses, the Qualtrics survey, and the simulated dataset and planned data analysis; everything is very clearly labelled and in relevant folders. I looked through the documents and they are clear and would aid an independent replication and reproducibility. I document any minor issues found below.

* In the screenshot below from the Qualtrics materials, should the second question now follow the same format as the first, with YES, NOT SURE, and NO, or is this an attention check in itself? (i.e., a participant might quickly click ‘NO’ in the second question thinking it is ‘YES’ because of the structure of the first question.



* In the power analysis, I noticed that alpha was set at .10. Is this correct? Should it be a = .05 or a = .01?

**Title:**

* Would this be clearer as “A replication and extension of Weiner et al. (1988)” (compared to “Replication and extensions of Weiner et al. (1988)”.

**Abstract:**

* The findings sub-section mention “stigma source” and “emotional reactions” but these are fairly ambiguous given the prior information given – could the same terms be used as is currently presented in the first introductory sentences? I am confused by what “stigma source” is when reading the Abstract alone; could this be ‘origins’ instead (and you could provide the e.g., of mental-behavioural and physical-based again here).

**Introduction:**

* Please check APA or BPS guidelines for terminology, but I think the term “obese people” can be seen as stigmatising, and it is best to use the term “individuals with obesity”.
* “Stigmas are negative social reactions to perceived deviations from the norm (Jones et al., 1984).”. There is a distinction between public and self-stigma; could you clarify here that this is a definition of Public stigma rather than capturing self-stigma.
* More information could be given about the Weiner et al. (1988) study on Page 8 of the Introduction, such as examples of “stigmas with somatic origins” and “mental-behavioral origins”. I was surprised to not see more detail here given that this is the study under replication.
* Page 9, remove the term “suggested” from “followed by our suggested extensions” if this article receives IPA.
* Page 9, what is “attributional search”?
* Please correct the references on Page 12: “importance of reproducibility and replicability in psychological science (Open Science Framework & Lakens, 2012)”. Is this Open Science Collaboration, 2012, 2015? And another reference to Lakens?

**Table 1**

* Can you further explain the issue of ‘randomization’– as it currently reads, I cannot see what the specific issue is (and the table is meant to provide an overview of these issues). Can you do the same for “Comparison between the control and experimental condition in Experiment 2”. In simple terms, it would be good to see the issues briefly described in this table so it is clear (just like with the “scale of meansurement” and “nature of stigma” inclusions).

**Replication target’s experimental design, hypotheses and findings**

* Page 13, first paragraph: What do you mean by “adding one generalized hypothesis for causality” – can you expand on this? Did you also test this generalized hypoithesis? Can you also expand on what is meant by “instrumentality” or “intervention techniques” in this paragraph; this assumes prior knowledge of the reader/the terms are quite ambiguous.

**Experiment 1: Baseline:**

* Does Weiner et al. explain how the different stigmas were cateogorised? For example, I would say that obesity is a physical health stigma, but Weiner et al. describe this as a mental behavioral stigma. Further information to clarify how these stigmas were cateogorised into the two groups would be helpful. Reading on, I see that your extension covers this by assessing “Participants’ categorization of stigmas”. Perhaps you can use my example here within this paragraph on Page 17 to show how their categorizations might not be have been fully accurate (and there is literature that supports obesity as a physical health condition).

**Four new current stigmas:**

* “Since the original article’s publication, the coverage and prevalence of different physiological and psychological conditions have changed considerably.” - can you support this statement with a reference?

**Rethinking the “child abuse” stigma item (Page 19):**

* “The described stigma regarding “child abuse” differed from other stigmas in that individuals suffering from the other stigmas were commonly described as the victims, yet the stigma “child abuse” was about the person who perpetrated the abuse rather than its victim”. Can you provide the specific statement from Weiner et al. which shows that this was about the perpetrator and not the victim; this will solidify your point.

**Table 5:**

In my opinion, Table 5 needs the most work to aid reader understanding.

* Why are there no relevant hypotheses proposed in Table 1 for the first row of Alzheimer’s disease, blindness, cancer, paraplegia, child abuse, and drug abuse. It would be best to specifically state these here for completeness.
* The sentences following the first of the “stigmas in the original study that will not replicate as well in the current replication” are confusing and could be reworded to aid clarity; here you state that H1 was “stigmas having a mental-behavioral origin are perceived as more onset-controllable than stigmas having a somatic genesis”, and then go onto state “heart disease, which is categorized as a physically-based stigma, will be perceived as more onset-controllable in our current replication. US Americans may perceive heart disease to be a mental-behavioral stigma nowadays (Waters et al., 2014). Such current perceptions on heart disease will contradict with original findings, which indicated that heart disease received low ratings on perceived controllability”. Paradoxically, then, such a proposed finding (heart disease will have *higher* ratings on perceived controllability) would actually support the original hypothesis that stigmas having a mental-behavioral origin are perceived as more onset-controllable because perceptions of it have changed from being physically-based to mental-behavioral based. In simple terms, the contradiction is with the categorisation of this condition (from physical to mental-behavioral), rather than contradicting the prediction (“stigmas having a mental-behavioral original are perceived as more onset controllable”).
* Condition 2 is presented under “Stigmas in the original study that will not replicate as well in the current replication”. However, it seems that, actually, you will support the original findings (“Original hypotheses will be disconfirmed again in the current replication.”).

**Method:**

**Pre-registration and open-science**:

* “We pre-registered the experiment on the Open Science Framework (OSF) and data collection was launched later that week. Pre-registrations, power analyses, and all materials used in these experiments are available in the supplementary materials. We provided all materials, data, code, and pre-registration on the OSF: <https://osf.io/gwcbt/>. We provided additional openscience details and disclosures in the supplementary materials under “Open Science disclosures” sub-section”. Can I clarify that this is written in past tense to aid the completion of the Stage 2 manuscript, but that data collection *has not yet begun?* Also, note that a space is needed between open and science.

**Power analysis/sample size:**

* “To demonstrate what the results would look like after data collection, we simulated a dataset of 1000 participants using Qualtrics, which we will later update with the real data and our sample of ~800”. Why did you base this on a sample of 1000 participants not your plan of 800 participants?
* “A sensitivity analysis indicated that a sample of 800 would allow the detection of f = 0.14 (groups = 3, df = 1) and d = 0.29 (independent samples with 266 participants in each condition; both 95% power, alpha = 5%, one-tail)”. Sorry to be pedantic, but 266 \* 3 = 798 and not 800.
* Do you plan for equal amounts of participants in each of the three conditions? Some text is confusing because it implies two conditions, when there are three, e.g. “In the simulation the no-information condition and the experimental condition each comprised 500 participants as samples, but in the actual data collection the three conditions will have a fairly even split”.

**Table 7:**

* Table 7 suggests that the gender was not reported in Weismann et al. (1988), but perhaps you could state what they did report: “149 male and female UCLA students and 171 male and female University of Manitoba students”. I understand that they did not split them, but it would be good to have this information.

**Design and procedure:**

* You ask for specific feedback on the following: “*We followed the target’s decision to try and balance the controllability/uncontrollability conditions by having the two information conditions evenly mixing controllability and uncontrollability stigma information. [Sidenote: After contemplating this design for long, we admit to struggling to understand this decision and having strong reservations regarding their design, we believe this should have been randomized, yet decided to try and replicate their design as is. More information is provided in the “deviations” section below. It is possible that we misunderstood or overlooked something important here, and so we would appreciate reviewers feedback on this point].*” ***Feedback:*** I think they evenly mixed controllability and uncontrollability so that more of one (e.g., controllability) didn’t bias the responses (i.e., participants would have viewed a greater proportion of one condition, which might have biased their responses/confounded stigma ratings). I have spent the past hour reading through, reflecting, and evaluating their Methods on page 727 and I believe you have got this replication method/procedure correct (however, I note that their Method section of Experiment 2 is very complex and contradictory in parts).
* Sorry if I’ve missed it, but whereabouts in the Procedure do you measure participant’s own categorizations of mental-behavioral or physical? I can see that this is asked at the start of the Qualtrics questionnaire (by looking at the questionnaire), but I don’t seem to be able to find it within the written procedure? Moreover, do you think that by presenting this first, this might weaken the manipulation of mental-behavioral vs. physical stigmas? (i.e. people’s pre-existing beliefs influence the stigmas, and this being salient nullifies the experimental manipulation?).

**Table 8:**

* This is excellent and really clarifies the design/procedure.

**Data Analysis Strategy:**

* This is clear and I agree with all analytical decisions here – the addition of a 3-way ANOVA is a good extension.

**Results:**

* Where are the simulated results for H1 and H2? Am I missing something here? For example, H1: Stigmas having a mental-behavioral origin are perceived as more onset-controllable than stigmas having a somatic genesis.
* Excellent simulation of results to test analytic pipelines.

I hope this peer review proves valuable in the revision of your manuscript,

Charlotte