Author's response to reviews

Title: After a Pair of Self-Control-Intensive Tasks, Sucrose Swishing Improves Subsequent Working Memory Performance

Authors:

Evan C Carter (evan.c.carter@gmail.com)
Michael E McCullough (mikem@miami.edu)

Version: 3 Date: 24 September 2013

Author's response to reviews:

September 24, 2013

Dear Professor Laws:

Thank you very much for your letter dated September 10, 2013, in which you invited us to further revise and resubmit our paper (MS# 1586490809937252), “After a Pair of Self-Control-Intensive Tasks, Sucrose Swishing Improves Subsequent Working Memory Performance” for BMC Psychology. We have revised the paper according to the round of comments we received, all of which came from Reviewer 2. We have copied the reviewer’s comments below (in bold) and added our responses following each of them.

Reviewer 2: Stuart Ritchie

Major point 1.

I thank the authors for conducting a power analysis, as I had suggested in my initial review. In their Discussion section, the authors note that their experiment was adequately (80%) powered if one takes an estimate from a 2010 meta-analysis as the effect size. However, they then say that their own to-be-published re-analysis of the literature has shown the effect to be smaller. If one uses this new effect size estimate, their experiment is underpowered.

First, I think this section is too vague. The authors say “...if the depletion effect is, actually, say, d = .25...” But is it? Presumably they will be able to give the actual effect size number from their review articles, and work out what the achieved power of their experiment was from there.

It seems odd to base one's power estimate on an effect size number one presumably does not agree with (that is, the one from the 2010 meta-analysis), so the authors are in the rather invidious position of having shown in previous work that their own achieved power in the present experiment is too low. They respond to this problem by noting that most of the other experiments in the ego depletion literature are underpowered, but I do not find this particularly convincing – several wrongs do not make a right. Nevertheless, the low power doesn’t make the experiment useless – it just means that it shows the effect is
smaller than previous estimates, but wasn’t able to test if a smaller effect exists.

So, given all of the above, I think the best tactic is for the authors to make it clear that what their experiment shows is that the size of ego depletion effect is either non-existent *or smaller than previously estimated*, and under some circumstances (i.e. with sucrose swishing) might be contrary to the prediction made in previous studies from both sides of the debate. I also think there’s something strange about mentioning the sample size as an advantage of the experiment in the Abstract, but then noting that the experiment may have been underpowered later in the paper.

We completely agree with Reviewer 2’s points about this section of the paper and have revised it so that it is (we think) both clearer and more in line with the arguments he has presented. The section, which starts on page 18, is copied here:

“A second limitation of the current work is the possibility that our null findings were the result of inadequate power. We did not conduct an a priori power analysis for our tests of the depletion effect (as mentioned, our data collection plan was to collect as much as possible in one semester). A priori power analyses are difficult to conduct for conceptual replications because it is not known if the parameter estimates provided by previous work generalize to the procedures that constitute the conceptual replication. Nevertheless, assuming the alternative hypothesis is true (i.e., the depletion effect is non-zero) for participants in the sucralose-sweetened and unsweetened rinse conditions, then our test of the depletion effect would have had 80% power for effect sizes of $d = 0.47$ or greater. According to Hagger et al. (2010), who provided a variety of meta-analytic estimates of the depletion effect for subsamples of experiments that were methodologically similar to ours, the depletion effect is at least this large.

However, if the depletion effect is nonzero but considerably smaller than $d = 0.47$, then the tests we conducted here are underpowered, and it is possible that our failure to find evidence for the depletion effect was due to low statistical power. According to one interpretation of our re-analyses of Hagger et al.’s (2010) meta-analytic data (Carter & McCullough, 2013; in press), it is possible that the depletion effect is indeed nonzero, but smaller than was originally estimated. Specifically, we found that based on one method of correcting for the influence of publication bias (Moreno et al., 2009), it is possible that the depletion effect is $d = 0.25$. If this estimate is correct, then any test that comprises fewer than 252 participants per group will have less than 80% power. Importantly, 188 of the 198 experiments reviewed by Hagger et al. (2010) had a total sample size of $N = 100$ or less, and the two largest experiments had total sample sizes of $N = 284$ and 501. In other words, if the depletion effect is some small, nonzero magnitude, then it would appear to be the case that the vast majority of experiments that have been conducted have been underpowered, including the one we report here.

Based on the experiment described here, as well as our re-analysis of Hagger et al.’s (2010) work, we believe that the balance of the evidence supports the
conclusion that the depletion effect is either not a robust phenomenon or that it is considerably smaller than has been previously reported. This conclusion is directly contrary to those that have been drawn by some other researchers (e.g., Vohs et al., 2012; Hagger et al., 2010). Thus, as we have recommended elsewhere (Carter & McCullough, 2013; in press), we believe that it is critical that researchers conduct large-scale direct replications of the classic tests of the depletion effect (e.g., replications of the experiments reported by Baumeister et al. [1998], but with total samples of at least N = 504).

Major point 2.
I still don’t think the Abstract is adequate. In the Background subsection, it reads as if the authors are making a distinction between “published evidence” and “other evidence”, which I know is not intentional. Also, could they clarify that the drink-swishing increases motivation by acting as a reward? In the Conclusion subsection, the authors say “given our sample size”, but they haven’t previously said whether the sample size was particularly large or particularly small (though see above on this). As I said in my previous review, the writing in the Abstract contrasts with the remainder of the manuscript, which is very clearly-written.

We have rewritten the abstract based on these reviews. We believe that the sample size/statistical power issue requires more explanation than is appropriate for the abstract, so we deleted the reference to the sample size that Reviewer 2 took issue with.

Minor point 1.
There is either a typo or a tense error in footnote b (p. 21): I think "we choose to use PCA" should be "we chose to use PCA".

This was a typo and has been fixed.

Minor point 2.
A thought occurs to me about Kool-Aid: for readers outside the US, might it be worth clarifying that usually sugar is added to Kool-Aid, and it is not sweetened at all without this? What the authors have written is fine, but I think some readers unfamiliar with Kool-Aid might assume it is sweetened, as most similar drinks are, and therefore think that the "sweetened" conditions had *extra* sucrose/sucralose added. This is a very minor comment, though: the authors may wish to leave things as they are.

We have clarified this point on page 8 with the addition of the following sentence: “Note that the Kool-aid flavoring mix we used was not sweet by itself.”

SUMMARY
We believe we have addressed all of the issues raised during this round of reviews. We have reworked the conclusion section so that it is clearer what conclusions can be reasonably drawn from this experiment. We have also modified the abstract so that it is (we think) less vague. We look forward to the possibility that this version of the manuscript might be acceptable for publication
in BMC Psychology. If you have additional concerns that you would like for us to address, feel free to contact me directly and we can deal with them immediately.

Thank you very much for your work on our behalf thus far.

Sincerely,

Michael E. McCullough, PhD
Professor