Author’s response to reviews

Title: The Transportability of Memory Specificity Training (MeST): Adapting an Intervention Derived from Experimental Psychology to Routine Clinical Practices

Authors:

Kris Martens (kris.martens@kuleuven.be)

Tom Barry (tom.j.barry@icloud.com)

Keisuke Takano (Keisuke.Takano@psy.lmu.de)

Filip Raes (filip.raes@kuleuven.be)

Version: 2 Date: 14 Jan 2019

Author’s response to reviews:

Dear Dr Baguley,

Thank you for your feedback, we’re happy that this manuscript is near publication. Here is our answer to your final remark.

With kindest regards,

The authors.

“I've gone through your careful and considered responses to the reviewers and I'm very happy with the final version except for one technical point. I missed this in the earlier version because you didn't report a CI d for the pre-post measures. This is in relation to the definition of d for a within-subjects design. There is no unique definition of d in a repeated measures design and so it is important to explain how you calculated d and how the CI is obtained. One option is Morris and DeShon (2008)'s suggestion of using the SD of the pre-test group (e.g., see https://www.psychometrica.de/effect_size.html#cohc and http://jakewestfall.org/blog/index.php/2016/03/25/five-different-cohens-d-statistics-for-within-subject-designs/ ). My own preference is to use unstandardised effect sizes to compare between designs as it easier to separate the estimate of the effect from its variability. Some options (such
as using the t to d conversion formula) are known to be problematic. In any case d needs to be calculated appropriately to compare between studies - which means using a common standardizer (the denominator for the calculation). I'm also not convinced that the CI for d needs to be provided for all settings separately given the small n.”

A: For clarification: the previous version included a Cohen’s d with a 95% confidence interval as calculated by JASP software, which uses the effsize package in R.

Following your recommendation, we now use simple means of pre-post differences per setting. The lower and upper limits of a 95% CI were calculated in which a SD of the pre-intervention measures (pooled across settings) was used.

We now mention this in the manuscript:

- Result section, line 382 – 383, p16
  “Using a mean of pre-post difference scores and a 95% CI using a pooled SD of pre-intervention scores this results in a mean difference of 2.70, 95% CI [1.90 – 3.50].”

- Both table 5 and 6 now no longer mention a Cohen’s d per setting but a mean of the pre-post difference scores. Both only mention a 95% CI the total sample using a pooled SD of pre-intervention scores. This calculation method is mentioned as a Note in Table 5 and 6:
  “A mean of pre-post difference scores is calculated per setting. A 95% CI is calculated for the total sample by using a pooled SD of the pre-intervention scores.”

- Discussion, line 433 to 439:
  “Comparing the current effect (a mean difference of 2.70 on the AMT, 95% CI [1.90 – 3.50]) with a previous study with high internal validity which used the same inclusion criterion [20] in which participants increased from a mean of 5.2 (SE = 0.4) to a mean of 8.0 (SE = 0.4), shows that the adaptations made for RCPs here did not decrease the efficacy of MeST in a significant way. Also, a translation to an open version (with a mean difference ranging from 2.33 to 3.42 in Settings 1, 2 and 3; see Table 5) showed comparably strong effects.”