Still No Evidence That Exergames Improve Cognitive Ability: A Commentary on Stanmore et al. (2017)

Giovanni Sala[1*], K. Semir Tatlidil[2], and Fernand Gobet[3]

[1]Osaka University

[2]Brown University

[3]University of Liverpool

Author's Note

*Corresponding author: Giovanni Sala, Graduate School of Human Sciences, Yamada Oka 1-2 Suita, Osaka University 565-0871, Japan. E-mail: sala@hus.osaka-u.ac.jp

Sala, G., Tatlidil, K. S., & Gobet (2019). Still no evidence that exergames improve cognitive ability: A commentary on Stanmore et al. (2017). *Neuroscience & Biobehavioral Reviews.*

Abstract

A recent meta-analysis (Stanmore et al. Neurosci. Biobehav. Rev. 78:34–43, 2017) claimed that exergames exert medium-size positive effects on people's overall cognitive function. The present article critically tests this claim. We argue that the meta-analysis reported inflated effect sizes mainly for three reasons: (a) some effect sizes were miscalculated; (b) there was an excessive amount of true heterogeneity; and (c) no publication-bias-corrected estimates were provided. We have thus recalculated the effect sizes and reanalyzed the data using a more robust approach and more sophisticated techniques. Compared to Stanmore's et al., our models show that: (a) the overall effect sizes are substantially smaller; (b) the amount of true heterogeneity, when any, is much lower; and (c) the publication-bias analyses suggest that the actual effect of exergames on overall cognitive function is slim to null. Therefore, the cognitive benefits of exergames are far from being established.

Keywords: aging; cognitive training; exergames; meta-analysis; transfer.

1. Introduction

A recent meta-analysis (Stanmore et al., 2017) has investigated the impact of exergames on overall cognitive ability. The meta-analysis included 17 Randomized Control Trials (RCTs) and a total of 926 participants. In most of the studies (n = 15), the participants were older adults (mean age > 55) with either no or some clinical condition (e.g., Parkinson's disease). The cognitive performance of the exergames-treated participants was compared to the performance of participants involved in several activities (e.g., stretching and cycling) or no activity at all. The meta-analysis reported a medium overall effect size ($\bar{g}$ = 0.436), indicating that exergames may be an effective tool to improve general cognition.

However, due to methodological issues, we think that the results of this meta-analysis are substantially unreliable. First, due to mistakes in the effect-size calculation, some effect sizes are inflated. Also, it is not clear what formula was used to calculate sampling error variances. Second, the amount of true heterogeneity is quite high (τ^2 = 0.170). Beyond making the results hard to interpret, such large τ^2 values inflate the overall effect size when the distribution of the effects is asymmetrical as in Stanmore et al. (2017). Third, even though Stanmore et al. (2017) includes two publication-bias analyses – the rank-correlation test and fail-safe N – neither of these methods provides an adjusted estimate of the overall effect size. In addition, the fail-safe N has been found to provide uninterpretable results (Schmidt and Hunter, 2015; pp. 531-534). Based on these issues, we present a re-analysis of Stanmore et al.'s data (2017).

2. Method

2.1. Effect Size Extraction

We included all the studies (RCTs) included in Stanmore et al.'s (2017) meta-analysis except one, Ackerman et al. (2010). This study investigated the effects of the Wii *Big Brain Academy* program that consists of a set of brain-training – rather than exergaming – activities. The number of included studies and independent samples was 16 (N = 883). We recalculated all the effect sizes and sampling error variances using the formulas provided by Schmidt and Hunter (2015).

2.2. Modeling Approach

We implemented *robust variance estimation* (RVE) with hierarchical weights (Hedges, Tipton, and Johnson, 2010). RVE allows one to model statistically dependent effect sizes and adjusts (i.e., increases) overall standard errors. Furthermore, RVE provides estimates of within-study and between-study true (i.e., not due to random error) heterogeneity components (ω^2 and τ^2, respectively). The effect sizes extracted from one study were thus grouped into the same cluster.

We then ran publication-bias analyses. First, the statistically dependent effects were merged using Cheung and Chan's (2014) method, and a random-effect model was run. Second, we used the trim-and-fill analysis with the *L0* and *R0* estimators (Schmidt and Hunter, 2015; pp. 538-540). Finally, since trim-and-fill analysis sometimes fails to fully correct for publication bias when the null is true, we employed the PET-PEESE method as an additional technique to assess publication bias.

2.3. Sensitivity Analysis

We controlled for potential outliers with influential-case analysis. The analysis individuated those studies that exerted a particularly strong influence on the model's estimates (e.g., overall effect size or true heterogeneity). We then removed the influential studies and reran the same set of analyses as described above.

3. Results

3.1. Main Model

The overall effect size of the RVE model was $\bar{g}$ = 0.212, 95% CI [-0.010; 0.434], m = 16, k = 75, p = .058, ω^2 = 0.000, τ^2 = 0.039. The model thus yielded a substantially smaller effect size and between-study true heterogeneity than Stanmore et al. (2017; $\bar{g}$ = 0.212 vs $\bar{g}$ = 0.436; τ^2 = 0.039 vs τ^2 = 0.170). After merging the effects, the overall effect size of the random-effect model was $\bar{g}$ = 0.246, p = .006, k = 16, τ^2 = 0.044. The trim-and-fill estimates were $\bar{g}$ = 0.076, p = .445 and $\bar{g}$ = 0.053, p = .586 with the *L0* and *R0* estimators, respectively. The PET and PEESE estimators were, $\bar{g}$ = 0.002, p = .986 and $\bar{g}$ = 0.079, p = .242, respectively.

3.2. Sensitivity Analysis

One influential study was detected. The overall effect size of the RVE model without this study was $\bar{g}$ = 0.113, 95% CI [-0.023; 0.248], m = 15, k = 63, p = .084, ω^2 = 0.000, τ^2 = 0.000. Excluding the influential study thus explained all the observed true heterogeneity (from ω^2 = 0.000, τ^2 = 0.039 to ω^2 = 0.000, τ^2 = 0.000) and reduced the overall effect of approximatively by a half (from $\bar{g}$ = 0.212 to $\bar{g}$ = 0.113). The overall effect size of the random-effect model was $\bar{g}$ = 0.109, p = .028, k = 15, τ^2 = 0.000. The trim-and-fill estimates were $\bar{g}$ = 0.066, p = .168 and $\bar{g}$ = 0.058, p = .219 with the *L0* and *R0* estimators, respectively. The PET and PEESE estimators were, $\bar{g}$ = -0.009, p = .889 and $\bar{g}$ = 0.048, p = .331, respectively.

4. Discussion

The aim of the present paper was to test the reliability of the findings of Stanmore et al.'s meta-analysis about the effects of exergame intervention on overall cognitive ability. Contrary to the findings of that meta-analysis, our reanalysis of the data has shown that the impact of exergaming on one's cognitive ability is very small at best and null at worst. Corrected overall effect sizes ranged from zero (PET estimates) to about 0.050-0.100 (all the other publication-bias estimates). Also, our reanalysis has yielded much more homogeneous and, hence, interpretable results. Finally, the methods used to model statistically dependent effect sizes (RVE and Cheung and Chan, 2014) do not seem to substantially affect the results (see the additional analyses). Based on the relatively small number of studies conducted at this point, our findings provide limited or even no evidence of the effectiveness of exergames on cognition. Future studies will contribute to updating the present findings.

Acknowledgements

We gratefully thank all the authors who provided additional data for the calculation of the effect sizes. The support of the Japan Society for the Promotion of Science [17F17313] is gratefully acknowledged.

References

Ackerman, P.L., Kanfer, R., Calderwood, C., 2010. Use it or lose it? Wii brain exercise practice and reading for domain knowledge. Psychol. Aging 25, 753–766. doi:10.1037/a0019277

Cheung, S.F., Chan, D.K., 2014. Meta-analyzing dependent correlations: an SPSS macro and an R script. Behav. Res. 46, 331–345. doi:10.3758/s13428-013-0386-2

Hedges, L.V., Tipton, E., Johnson, M.C., 2010. Robust variance estimation in meta-regression with dependent effect size estimates. Res. Synth. Methods. 1, 39–65. https://doi.org/10.1002/jrsm.5

Schmidt, F.L., Hunter, J.E., 2015. Methods of meta-analysis: correcting error and bias in research findings (3rd ed.). Newbury Park, CA: Sage.

Stanmore, E., Stubbs, B., Vancampfort, D., de Bruin, E.D., Firth, J., 2017. The effect of active video games on cognitive functioning in clinical and non-clinical populations: a meta-analysis of randomized controlled trials. Neurosci. Biobehav. Rev. 78, 34–43. http://dx.doi.org/10.1016/j.neubiorev.2017.04.011