Mozart effect: A meta-research study on statistical power, effect size, and false discovery rate - Preliminary results

Marie Meunier[1], Ezio Tirelli[1], Nancy Durieux[1], François Léonard[2]

[1] Department of Psychology; [2] Department of Educational Sciences

Introduction

The Mozart effect, first reported by Rauscher et al. (1993), refers to the claim that listening to Mozart's Sonata K.448 temporarily enhances spatial or spatial-temporal abilities. Although the effect has been widely studied, replication attempts have yielded inconsistent findings. Meta-analyses report small to moderate effects, with evidence of publication bias (Pietschnig et al., 2010), casting doubt on the robustness of the effect.

The aim of this preliminary meta-research study is to assess statistical power and the false discovery rate (FDR) in the Mozart effect literature. This preliminary study is part of a larger project registered on the Open Science Framework: https://osf.io/3854h.

Methods

We systematically searched PsycINFO, Medline, Embase, and ERIC for studies comparing Mozart's Sonata K.448 to control conditions (silence, white noise, or relaxation). Screening and data extraction will be conducted by three reviewers. For these preliminary results, data extraction was performed by a single trained reviewer.

Effect sizes (Hedges' g) were computed from reported statistics. We conducted random-effects meta-analyses (REMA) using the restricted maximum likelihood estimator (Viechtbauer, 2005) to estimate heterogeneity (τ^2), and applied the Knapp-Hartung adjustment (Knapp & Hartung, 2003) to estimate confidence intervals. To adjust for small-study effects, we used the Trim and Fill method (Duval & Tweedie, 2000).

To address dependent effect sizes and improve estimates' reliability, we applied multilevel meta-analysis (MLMA) and a robust two steps approach (Yang et al., 2024) with clustered variance–covariance estimation.

Statistical power was estimated per study based on meta-analytic effects, and FDRs were computed under various plausibilities of H_1 (Ioannidis, 2005).

Results

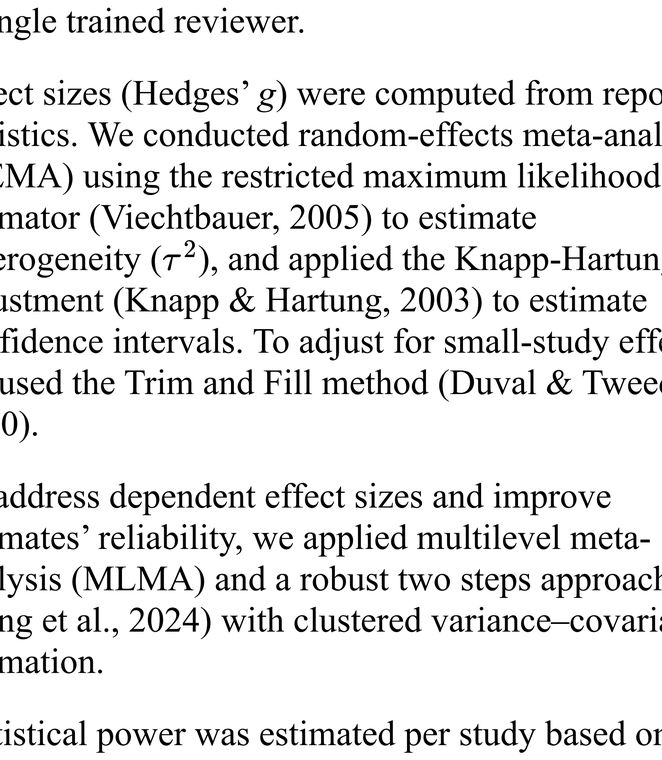

Figure 1. Flowchart of study selection process (identification, screening, inclusion)

Table 1. Estimated effect sizes (Hedges' g) and 95% confidence intervals for each meta-analytic model.

Method	Effect size	95% CI
REMA	0.236	[0.111; 0.36]
Trim & Fill	0.099	[-0.061; 0.259]
MLMA	0.231	[0.123; 0.339]
Two-Step RVE	0.167	[0.07; 0.264]

Overall, all meta-analytic models converged on small effect sizes (Hedges' $g \approx 0.099$ - 0.236). Notably, the Trim & Fill correction reduced the effect size estimate and widened the confidence interval, suggesting potential small-study effects.

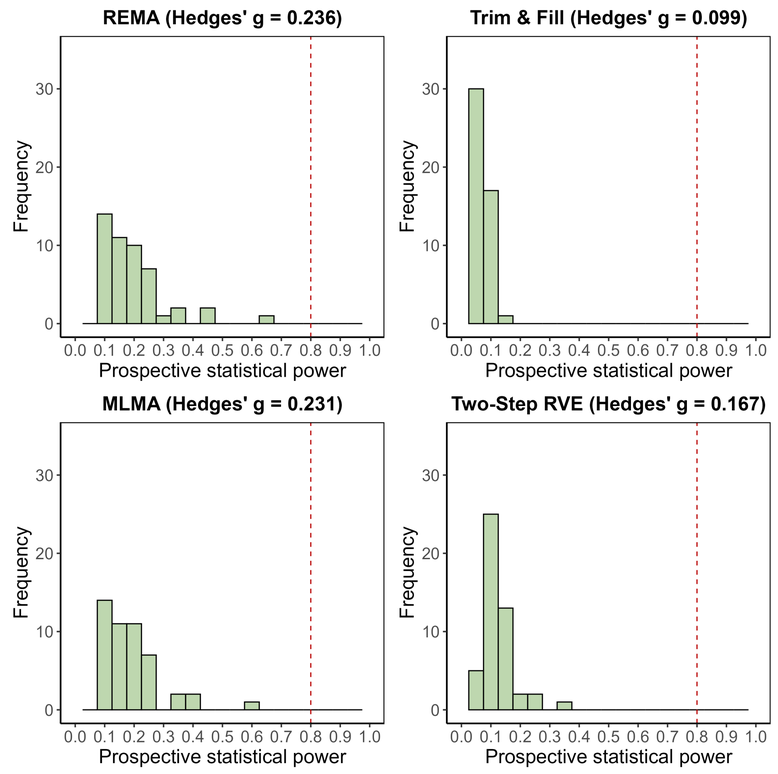

Figure 2. Histograms of estimated prospective power across four effect size estimates. Each panel represents a meta-analytic model (REMA, Trim & Fill, MLMA, Two-Step RVE) with the associated effect size (Hedges' g). The dashed vertical line marks the 80% power threshold.

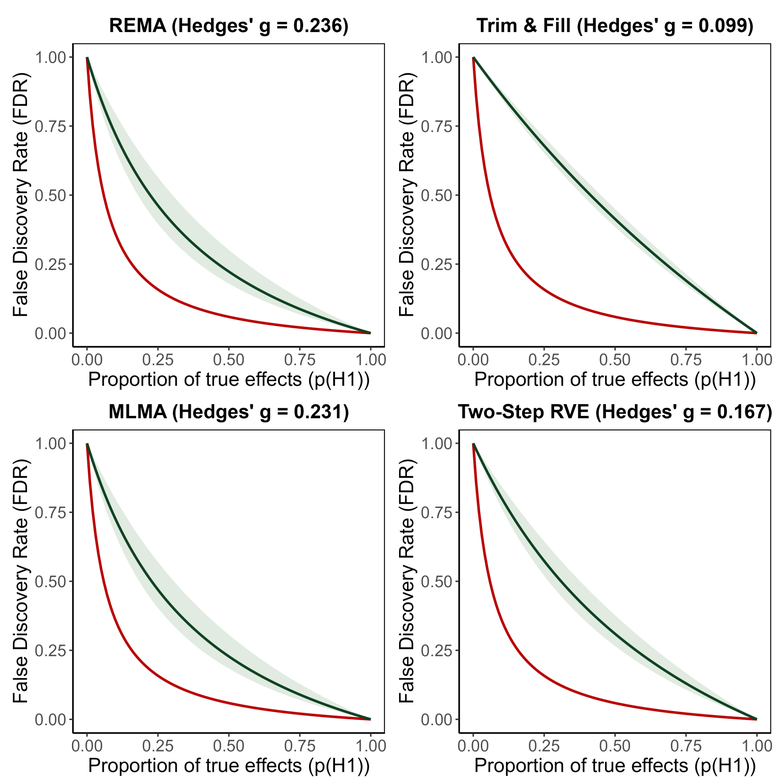

Figure 3. False Discovery Rates (FDRs) across prior probabilities of a true effect (p(H_1)), plotted separately for each meta-analytic model. Shaded ribbons represent 95% uncertainty bounds based on the confidence intervals of the estimated effect sizes. The dashed red line indicates the FDRs under a conventional 80% power assumption.

Across models, all studies were underpowered to detect the estimated effect sizes at the 80% threshold. This low power likely increases the probability of both false negatives and overestimated effects. Due to consistently low statistical power across studies, observed FDR curves exceed commonly recommended FDR curves for a power of 80%, particularly when prior probability is low.

Discussion

Estimated effect sizes across models were consistently small (Hedges' $g \approx 0.099$ - 0.236), suggesting limited practical significance of the Mozart effect.

The persistently low statistical power observed in the primary studies raises concerns about the stability of the reported findings and the risk of inflated estimates. Under plausible prior assumptions, the associated FDRs were high, further challenging the credibility of the literature.

In addition, methodological limitations were common: many studies lacked essential design features such as blinding, randomisation or sample size justification (results not shown), which may distord the estimated effects and contribute to inconsistent results.

Together, these findings suggest that the evidence for the Mozart effect is weakened by suboptimal methodological practices. More robust designs and adequate sample sizes are needed to clarify whether any reliable effect exists.

Beyond the Mozart effect, these preliminary results illustrate the broader need for stronger methodological standards in this literature.

References

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, *22*(17), 2693–2710. https://doi.org/10.1002/sim.1482

Pietschnig, J., Voracek, M., & Formann, A. K. (2010). Mozart effect–Shmozart effect: A meta-analysis. *Intelligence*, *38*(3), 314–323. https://doi.org/10.1016/j.intell.2010.03.001

Rauscher, F. H., Shaw, G. L., & Ky, C. N. (1993). Music and spatial task performance. *Nature*, *365*(6447), 611–611. https://doi.org/10.1038/365611a0

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293. https://doi.org/10.3102/10769986030003261

Yang, Y., Lagisz, M., Williams, C., Noble, D. W. A., Pan, J., & Nakagawa, S. (2024). Robust point and variance estimation for meta-analyses with selective reporting and dependent effect sizes. *Methods in Ecology and Evolution*, *15*(9), 1593–1610. https://doi.org/10.1111/2041-210X.14377