Note: The following manuscript is a draft. It has not gone through peer-review yet.

This is the first version of the manuscript. The supplemental materials including all the effects sizes and other technical details will be provided soon.

Feedback from fellow researchers is welcome. Please direct your comments to Giovanni Sala, email:sala@hus.osaka-u.ac.jp

Near and Far Transfer in Cognitive Training: A Second-Order Meta-Analysis

Giovanni Sala[1], N. Deniz Aksayli[2], K. Semir Tatlidil[3], Tomoko Tatsumi[4], Yasuyuki Gondo[1], and Fernand Gobet[4]

[1]Osaka University

[2]Istanbul, Turkey

[3]Brown University

[4]University of Liverpool

Authors' Note

GS and FG conceptualized this paper. GS wrote the first draft of the paper and performed the statistical analyses. NDA and KST extracted the effect sizes of the meta-analysis of working-memory training in older adults. GS and KST extracted the effect sizes of the meta-analysis of exergame training. GS and TT extracted the effect sizes of the meta-analysis of chess training. All the authors contributed to drafting the current version of this preprint.

Corresponding author: Giovanni Sala, Department of Clinical Thanatology and Geriatric Behavioral Science, Yamada Oka 1-2 Suita, Osaka University 565-0871, Japan. E-mail: sala@hus.osaka-u.ac.jp

Giovanni Sala is a JSPS International Research Fellow.

Abstract

Recent empirical and meta-analytic evidence has shown that the benefits of cognitive-training programs hardly go beyond the trained task and similar tasks. However, it is yet to be established whether the effects differ across cognitive-training programs and populations (children, adults, and older adults). We addressed this issue by using second-order meta-analysis (i.e., a meta-analysis of meta-analyses).

In Models 1 ($k = 99$) and 2 ($k = 119$), we investigated the impact of working-memory training on near-transfer (i.e., memory) and far-transfer (e.g., reasoning, speed, and language) measures, respectively. Model 3 ($k = 207$) extended Model 2 by adding four meta-analyses assessing the far-transfer effects of other cognitive-training programs (action video-games, music, chess, and exergames). Model 1 showed that working-memory training does induce near transfer, and that the size of this effect is moderated by the type of population. On the contrary, Models 2 and 3 highlighted that far-transfer effects are small or null. Crucially, when placebo effects and publication bias were controlled for, the overall effect size and true variance equaled zero. That is, no impact on far-transfer measures was observed regardless of the type of population and cognitive-training program. The lack of generalization of skills acquired by training is thus an invariant of human cognition.

Keywords: cognitive training; meta-analysis; second-order meta-analysis; transfer.

Near and Far Transfer in Cognitive Training: A Second-Order Meta-Analysis

Transfer of skills is the generalization of skills acquired by training across different domains. Transfer is a central phenomenon in cognitive psychology because it is a manifestation of how humans acquire and process information. It is customary to distinguish between near and far transfer (Barnett & Ceci, 2002): while the former refers to the generalization of skills across similar domains, the latter indicates the transfer of skills across domains that are not, or very weakly, related to each other. More precisely, the distinction between near and far transfer relies on the overlap between the source and target domains. In other words, the definition of the type of transfer is directly related to the extent to which the domains share common features. The more the shared features, the nearer the transfer. Importantly, such features include both perceptual and conceptual information (Singley & Anderson, 1989).

According to the common elements theory (Thorndike & Woodworth, 1901), the likelihood of transfer to take place is directly related to the degree to which the source domain and the target domain share common features. Substantial research into learning, skill acquisition, and expertise has corroborated the theory: while near transfer often occurs, far transfer is rare (Detterman, 1993; Donovan, Bransford, & Pellegrino, 1999; Ritchie, Bates, & Deary, 2015; Sala & Gobet, 2017a). Relying on common sense leads to the same conclusion. For instance, it is reasonable that learning analytic geometry facilitates the acquisition of knowledge in calculus because there is some overlap between the two fields. Conversely, there is no clear reason why learning Latin sentence structures should be of any use for learning calculus (or vice versa).

Obtaining Far Transfer Through Cognitive Training

As seen, it is unanimously acknowledged that near transfer is much more common than far transfer. Nonetheless, far transfer is undoubtedly a much more interesting phenomenon to researchers, policymakers, and practitioners. Most (if not all) the theories and cognitive architectures of memory and skill acquisition, implicitly or explicitly, make predictions about the possible occurrence of far transfer (e.g., Gobet, 2016; Gobet & Simon, 1996; Singley & Anderson, 1989; Taatgen, 2013). The presence/absence of far transfer is thus a valuable litmus test for theories of human cognition. Furthermore, knowing whether and under what conditions far transfer occurs would represent a breakthrough in education and training in general. Skill acquisition is a costly endeavor and acquiring expertise in more than one specific field is a rare achievement (Ericsson & Charness, 1994). Knowing how to generalize skills acquired in a particular domain to many different domains would help trainees to develop a broad set of skills in many areas more efficiently. Thus, understanding the mechanism of transfer is a major challenge in cognitive science with profound theoretical and societal implications.

Researchers are yet to reach an agreement about the actual possibility to obtain far transfer of skills. Some authors have, implicitly or explicitly, suggested that the lack of far transfer is a fundamental characteristic in human cognition (e.g., Chase & Ericsson, 1982; Detterman, 1993; Sala & Gobet, 2017a; Simons et al., 2016). Domain-specific skills acquired by training exert an impact on the relevant domain but hardly generalize to other domains. Moreover, even transfer of skills from one particular field of expertise to one of its sub-domains appears to discount significant decrease in performance (e.g., Bilalić, McLeod, & Gobet, 2009; Rikers, Schmidt, & Boshuizen, 2002). This line of research does not deny that some people manage to excel in more than one domain. However, the explanation of this phenomenon does not rely on transfer. People with superior cognitive ability are more likely to excel in several domains because they acquire knowledge better and faster than the general population (e.g., Burgoyne et al., 2016; Detterman, 2014; Schmidt, 2017).

Other scholars are more optimistic and have suggested that it is possible to elicit far transfer. To date, the most influential and systematic attempt to obtain far transfer of skills is represented by cognitive training (for a review, see Strobach & Karbach, 2016). The cognitive-training program of research assumes that general cognitive ability or, at least, some core cognitive mechanisms (e.g., working memory, inhibition, and processing speed) can be enhanced by engaging in cognitively demanding exercises. Some of these activities, such as working-memory-training and brain-training programs, have been purposely designed to boost cognitive function. Other training programs implement mentally challenging activities such as music instruction, video games, and chess (for reviews, see Sala & Gobet, 2017a; Simons et al., 2016; Strobach & Karbach, 2016).

The basic idea underlying cognitive-training programs is that the enhancement of domain-general cognitive mechanisms is a by-product of training in domain-specific activities (Taatgen, 2016). Consistent with the research on skill acquisition and expertise, engaging in cognitive-training programs has been found to improve participants' performance on the trained task and related tasks (e.g., Simons et al., 2016). However, these activities are also believed to foster overall cognitive function or, at least, some domain-general cognitive skills (e.g., memory and processing speed). Once improved, enhanced domain-general cognitive skills are supposed to boost professional and academic domain-specific capabilities that depend on them. Neural plasticity is believed to be the mediator of this process (Karbach & Schubert, 2013).

The Present Study

Hundreds of experimental studies have examined the impact of cognitive-training programs on people's ability to perform cognitive and academic tasks. In order to make some sense of the sometime mixed results, the evidence provided by these investigations has often been evaluated via meta-analysis. The meta-analytic evidence collected so far suggests that, overall, cognitive-training programs exert small to medium near-transfer effects and small to null far-transfer effects (e.g., Melby-Lervåg, Redick, & Hulme, 2016; Sala & Gobet, 2017a). However, it is yet to be understood whether some specific cognitive-training regimens can induce transfer effects better than others. Also, it is unclear whether the type of population (e.g., young adults, older adults) that undergoes a particular cognitive-training regimen moderates the degree to which transfer occurs. In other words, it is yet to be clarified whether there is genuine between-regimen and between-population variability regarding both near- and far-transfer effects. The present investigation employs second-order meta-analysis to address these questions.

Second-order meta-analysis is defined as “a meta-analysis of a number of statistically independent and methodologically comparable first order meta-analyses examining ostensibly the same relationship in different contexts” (Schmidt & Oh, 2013, p. 204). One of the major advantages of conventional (i.e., first-order) meta-analysis is reducing sampling error by merging effect sizes from different sources. That allows researchers to produce more precise measures of an effect than the single primary study. However, sampling error can never be ruled out entirely because the number of included samples is always less than infinite. Hunter and Schmidt (2004) define this residual sampling error as *second-order sampling error*. Second-order meta-analysis aims to estimate to what extent second-order sampling error accounts for the difference across overall meta-analytic means in a set of first-order meta-analyses regarding a particular topic. First, first-order meta-analytic means are used to calculate a grand mean ($\bar{\bar{g}}$). Then, the proportion of the between-meta-analysis variance explained by second-order sampling error is calculated and used to produce more accurate estimates in first-order meta-analyses. If second-order sampling error explains all the observed variance, all first-order meta-analytic means are corrected into their grand mean because no true variance is observed across first-order meta-analytic means (i.e., $\sigma^2 = 0$). Conversely, if second-order sampling error accounts only for a portion of the variance (i.e., $\sigma^2 > 0$), corrected first-order meta-analytic means are closer, but not necessarily identical, to the grand mean than the uncorrected means. To date, second-order meta-analysis thus represents the highest level of cumulative quantitative knowledge.

We here run three main models. In Models 1 and 2, we investigate whether the effects of working-memory (hereafter WM) training on performance in cognitive tasks are mediated by the type of population. Model 1 analyzes the effects of WM training on memory tasks (i.e., near-transfer effects). Model 2 focuses on far-transfer tasks (e.g., fluid reasoning, language, and cognitive control). To date, WM training is the most studied and probably most influential cognitive-training program. Moreover, the average quality of WM-training studies is excellent. In fact, the primary studies often include pre-post-test assessments, active control groups, and measures of both near and far transfer. For these reasons, WM training is the most suitable cognitive-training program for testing the extent to which trained skills transfer across different cognitive tasks. Finally, Model 3 is an extension of Model 2. In Model 3, we included other four meta-analyses of other cognitive-training programs: action video-game training, music instruction, chess instruction, and exergames (i.e., cognitive-training games combined with physical activities).

General Method

Inclusion Criteria

We established four inclusion criteria to guarantee a minimum standard of design quality in the primary studies:

(1) The primary study included at least one control group;

(2) The primary study included a pre-test to assess baseline effects;

(3) The experimental samples were not self-selected;

(4) The transfer effects were measured by a cognitive/academic task. Self-reported measures were excluded.

Some studies and effect sizes from the published first-order meta-analyses were excluded according to these criteria.

Effect Sizes

The effect size used in all the meta-analyses was the corrected standardized mean difference, that is, Hedges' *g* (Hedges & Olkin, 1985). The effect size represented the amelioration of the experimental groups over the controls immediately after the end of the training. Due to their dearth, no follow-up effects were included. For most of the meta-analyses, we used the original effect sizes. In two cases (see Model 3), we recalculated all the effect sizes and variances to uniform the standard across the meta-analyses. The formula for the effect size was:

$$g = \frac{\left(M_{1post} - M_{1pre}\right) - \left(M_{2post} - M_{2pre}\right)}{SD_{pooled_{pre}}} \times \left(1 - \frac{3}{(4 \times N) - 9}\right)$$

that is, the post-pre between-group mean difference standardized by the pooled pre-test standard deviations and corrected for upward bias. The formula for effect size variance was:

$$Var_g = \frac{N}{N_e \times N_c} + \frac{g^2}{N \times 2}$$

where N, N_e, and N_c are the total sample size of the study, experimental group, and control group, respectively (Hedges and Olkin, 1985; Schmidt & Hunter, 2015; pp. 292-293).

Active Controls

Active control groups are necessary to control for possible placebo effects. For this reason, we also ran models including only experimental groups matched with active controls. According to commonly accepted guidelines (e.g., Boot, Simons, Stothart, & Stutts, 2013; Simons et al., 2016), we considered a control group as "active" only if it consisted of an engaging and cognitively demanding activity (e.g., non-adaptive training, visual-search training, etc.).[1] Alternative tasks with negligible cognitive demand (e.g., watching videos and filling in questionnaires) were labeled as "non-active." Two coders independently judged whether the primary studies implemented such an active control group.

Correction for Statistical Dependence

Primary studies often report more than one measure of cognitive ability. Measures from the same samples are, by definition, statistically dependent. Modeling these effect sizes as statistically independent does not introduce any systematic bias in the estimation of meta-analytic means (Schmidt & Hunter, 2015) or even confidence intervals (Tracz, Elmore, & Pohlmann, 1992). Nevertheless, not correcting for statistical dependency leads to an underestimation of sampling error variances. Conversely, just merging the effects without applying any additional correction overestimates sampling error variances (Schmidt & Hunter, 2015). Given that a major goal of second-order meta-analysis is to estimate the amount of between-meta-analysis variability explained by second-order sampling error, this bias must be corrected or at least reduced. To address the problem, we use Cheung and Chan's (2014) samplewise-adjusted-individual correction. This technique has been designed to estimate an adjusted variance based on (a) the number of the

---

[1] The only exception is the meta-analysis of exergame training. In this field, the active controls usually consist of participants involved in physical activities. For more details, see Model 3.

dependent effect sizes and (b) inter-effect-size correlation. The procedure thus calculates a more precise sampling error variance.

The method requires transforming effect sizes into Pearson correlations (*r*s). We thus use the formulas reported in Borenstein, Hedges, Higgins, and Rothstein (2009) to convert standardized mean differences into correlations and vice versa:

$$r_i = \frac{g_i}{\sqrt{4 + g_i^2}}$$

$$g_i = \frac{2 \times r_i}{\sqrt{1 - r_i^2}}$$

The corrected variances (V_{r_i}) calculated by the R code provided in Cheung and Chan (2014) are then converted into variances of standardized mean differences with the following formula (Borenstein et al., 2009):

$$V_{g_i} = \frac{4 \times V_{r_i}}{(1 - r_i^2)^3}$$

Finally, the adjusted variances and the reconverted *g*s are employed to run a random-effect model.

It is important to note that Cheung and Chan’s (2014) correction does not distinguish between totally or partially dependent effects. In almost all the cases, the primary studies report either totally dependent effect sizes—that is, extracted from the same sample (i.e., experimental group vs. control group)—or effect sizes extracted from partially overlapping samples comparing the experimental group with both an active and passive control groups. The models including only comparisons between the experimental groups and active controls do not suffer from this minor technical issue and thus serve as a further control for the reliability of the whole procedure. Finally, a few primary studies report effect sizes extracted from two independent samples (i.e., experimental 1 vs control 1 and experimental 2 vs control 2). We model these effect sizes as statistically dependent to reduce the weight of these studies in the analyses. The number of effect sizes in the first-order meta-analyses is thus equal to the number of studies. This more conservative approach is commonly employed in other meta-analytic techniques for modeling dependent effect sizes (e.g., Tanner-Smith & Tipton, 2014). (It is worth noting that other methods to model nested effect sizes [e.g., Hedges, Tipton, & Johnson, 2010; Konstantopoulos, 2011] are not ideal for running a second-order meta-analysis because they do not allow us to use multiple techniques for publication-bias correction.)

Publication Bias Analysis

Naïve (i.e., uncorrected) meta-analytic means are often less reliable than the publication-bias corrected estimates (Schmidt & Oh, 2013; Stanley, 2017). We therefore ran a set of publication bias analyses for all the first-order meta-analyses and built a parallel set of second-order meta-analyses using publication-bias corrected estimates. It is usually recommendable to employ multiple publication-bias detection techniques to triangulate the most likely true (i.e., unbiased) overall effect size (e.g., Kepes & McDaniel, 2015). First, we reported funnel plots for the visual inspection of the distribution of the effect sizes. Second, we used the "precision effect test" (PET) and the "precision-effect estimate with standard error" (PEESE; Stanley & Doucouliagos, 2014). The PET estimator is the intercept of a weighted (by precision) linear regression where the dependent variable is the effect size and the independent variable is its standard error. The PEESE estimator is obtained by replacing the standard error with the standard error squared (i.e., variance) as the independent variable. If PET suggests the presence of a non-zero effect (at the 10% of significant level; Stanley, 2017), the PEESE estimator is employed. Third, when both the PET and PEESE produced inaccurate (i.e., very high standard error) or excessively negative estimates[2] we used the trim-and-fill analysis with all the three estimators (L0, R0, and Q0) described in Duval and Tweedie (2000). Since the trim-and-fill has been found to underestimate the amount of publication bias

[2] We assume that the true transfer effect of a training program cannot be significantly negative.

(Carter, Schönbrodt, Gervais, & Hilgard, 2017; Moreno et al., 2009; Simonsohn, Nelson, & Simmons, 2014), the PET-PEESE estimates were preferred when they did not suffer from the abovementioned flaws (i.e., large negative values and high standard errors). As a general rule, when the PET test did not show evidence of non-zero effect, the estimate that was the closest to zero was picked up for the second-order meta-analysis. (It is worth noting that selection methods such as selection models (e.g., McShane, Böckenholt, & Hansen, 2016; Vevea & Woods, 2005) and *p*-curve (Simonsohn et al., 2014) are not suitable when statistically dependent effects have been merged because the *p*-value distribution significantly differs from the original one.)

Heterogeneity

Between-study within-meta-analysis degree of true heterogeneity was assessed by the I^2 statistic (Higgins, Thompson, Deeks, & Altman, 2003). High degrees of true heterogeneity artificially increase the relative weight assigned to small studies, which in presence of publication bias can result in an overestimation of the overall effect size (Stanley & Doucouliagos, 2017). Moreover, high degrees of true heterogeneity (e.g., $I^2 > 75\%$) can bias publication-bias corrected estimates (Schmidt & Hunter, 2015; Stanley, 2017). Therefore, when the degree of true heterogeneity was statistically significant ($p < .100$), we ran Viechtbauer and Cheung's (2010) influential case analysis. The detected influential cases were excluded to reduce the degree of heterogeneity and enhance the reliability of the corrected effect sizes.

Second-Order Meta-Analytic Procedure and Omnibus Meta-Analysis

Second-order meta-analysis requires the studies in the first-order meta-analysis being statistically independent (Schmidt & Oh, 2013). In other words, no study (or sample) in must be included in more than one first-order meta-analysis. Since we divided the first-order meta-analysis by type of population and type of cognitive-training program, this assumption was met in all the second-order meta-analytic models.

None of the included first-order meta-analyses corrected the effect sizes for measurement error. Thus, we implemented the equations for second-order meta-analysis of bare bones meta-analyses (Schmidt & Oh, 2013; pp. 207-209). The data related to first-order meta-analysis (overall effect size, standard error, degree of true heterogeneity) were calculated by the metafor R package (Viechtbauer, 2010). The amount of total between-study heterogeneity (τ^2) was calculated with the REML estimator. We used random-effect models for all the first-order meta-analyses. We ran two different sets of second-order meta-analytic models. The first one included all the uncorrected (naïve) meta-analytic means. The second one included a corrected effect size from each first-order meta-analysis estimated by the publication bias analysis. This whole procedure was carried out twice: we ran a second-order meta-analysis with first-order meta-analysis including (a) all the primary studies and (b) only the comparisons between experimental and active control groups to check for placebo effects. Therefore, four second-order meta-analyses were carried out for each of the three models.

Finally, we also ran an omnibus meta-analysis—that is, a first-order meta-analysis including all the primary studies of all the first-order meta-analysis (Borenstein et al., 2009)—for each of the three models. While estimating an overall effect size similar to the grand mean in second-order meta-analysis, omnibus meta-analysis does not provide an estimation of second-order sampling error. This additional analysis was performed as a sensitivity analysis to control for the overall effect sizes, publication-bias estimates, and true between-study heterogeneity calculated in the first-order meta-analyses.

Model 1: Working Memory Training (Near Transfer)

Description of The First-Order Meta-Analyses

We selected four meta-analyses including typically developing (TD) children (Sala & Gobet, 2017b), children with learning disabilities (LD; a subsample of Melby-Lervåg et al., 2016), healthy adults (a subsample of Melby-Lervåg et al., 2016), and healthy older adults (unpublished). The three published meta-analyses were selected because (a) they were recent and (b) the authors employed comparable methods. In fact, they calculated the standardized mean difference with the same formula (i.e., the difference between mean pre-post gains of experimental and control groups divided by the pooled pre-test standard deviation), reported only effects from cognitive/academic tests (no self-reported questionnaires) and included both near-transfer and far-transfer measures. The fourth meta-analysis (healthy older adults) was carried out accordingly to the criteria of the other three meta-analyses. No study or effect size reported in the original meta-analyses was excluded. This section examines the impact of WM training on the participants' performance in memory tasks (i.e., near transfer).

Results

TD children.

The overall meta-analytic mean was $\bar{g}$ = 0.45, *SE* = 0.07, 95% CI [0.31; 0.58], k = 16, p < .001. The test of heterogeneity was not significant, Q = 15.63, I^2 = 5.74%, p = .407. The distribution of the effect sizes in the funnel plot looked substantially symmetrical (Figure 1).

Insert Figure 1 around here

Figure 1. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the TD children (near transfer).

PET and PEESE estimators were $\bar{g} = 0.59$, $SE = 0.20$, $p < .001$ and $\bar{g} = 0.50$, $SE = 0.11$, $p < .001$, respectively. Since the PET test showed evidence of a real effect ($p < .100$), the PEESE estimator was used in the relevant second-order meta-analysis.

With regard to the type of control group (active vs. non-active), there was 100% inter-rater agreement. The classification of the type of control group was the same as in Sala and Gobet (2017b). When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g} = 0.43$, $SE = 0.09$, 95% CI [0.26; 0.61], $k = 11$, $p < .001$. The test of heterogeneity was not significant, $Q = 6.59$, $I^2 = 0.00\%$, $p = .764$. The distribution of the effect sizes in the funnel plot looked symmetrical (Figure 2).

Insert Figure 2 around here

Figure 2. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the TD children (near transfer, active controls).

PET and PEESE estimators were $\bar{g} = 0.54$, $SE = 0.26$, $p = .073$ and $\bar{g} = 0.49$, $SE = 0.13$, $p < .001$, respectively. Since the PET test showed evidence of a real effect ($p < .100$), the PEESE estimator was used in the relevant second-order meta-analysis.

LD children.

The overall meta-analytic mean was $\bar{g} = 0.55$, $SE = 0.10$, 95% CI [0.35; 0.76], $k = 17$, $p < .001$. The test of heterogeneity was significant, $Q = 54.65$, $I^2 = 74.12\%$, $p < .001$. The distribution of the effect sizes in the funnel plot looked highly asymmetrical (Figure 3).

Insert Figure 3 around here

Figure 3. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the LD children (near transfer).

PET and PEESE estimators were $\bar{g}$ = 0.25, *SE* = 0.15, *p* = .110 and $\bar{g}$ = 0.40, *SE* = 0.09, *p* < .001, respectively. Since the model showed a high degree of true heterogeneity, we ran an influential case analysis. The influential case analysis found one influential study (*g* = 1.40, *SE* = 0.21). After removing this study, the overall meta-analytic mean was $\bar{g}$ = 0.46, *SE* = 0.08, 95% CI [0.30; 0.62], *k* = 16, *p* < .001. The test of heterogeneity was still significant, *Q* = 34.15, I^2 = 54.70%, *p* = .003. PET and PEESE estimators were $\bar{g}$ = 0.25, *SE* = 0.12, *p* = .057, and $\bar{g}$ = 0.37, *SE* = 0.07, *p* < .001, respectively. This PEESE estimator was used in the relevant second-order meta-analysis.

Regarding the type of control group (active vs. non-active), there was 100% inter-rater agreement. The classification of the type of control group was the same as in Melby-Lervåg et al. (2016). When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g}$ = 0.48, *SE* = 0.11, 95% CI [0.27; 0.69], *k* = 12, *p* < .001. The test of heterogeneity was significant, *Q* = 25.52, I^2 = 62.36%, *p* = .008. The distribution of the effect sizes in the funnel plot still looked highly asymmetrical (Figure 4).

Insert Figure 4 around here

Figure 4. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the LD children (near transfer, active controls).

PET and PEESE estimators were $\bar{g}$ = -0.19, *SE* = 0.20, *p* = .374 and $\bar{g}$ = 0.08, *SE* = 0.11, *p* = .484, respectively. Since the PET estimator was excessively negative, we used the trim-and-fill analysis to calculate the publication bias. With the L0 estimator, the analysis filled three studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = 0.33, *SE* = 0.14, 95% CI [0.06; 0.61], *k* = 15, *p* = .017, I^2 = 79.80%. With the R0 estimator, the analysis filled five studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = 0.26, *SE* = 0.14, 95% CI [-0.01; 0.53], *k* = 17, *p* = .058, I^2 = 80.42%. With the Q0 estimator, the analysis filled seven studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = 0.22, *SE* = 0.13, 95% CI [-0.03; 0.47], *k* = 19, *p* = .087, I^2 = 79.52%. To reduce the degree of true heterogeneity, we ran an influential case analysis. The influential case analysis found two influential studies (g = 1.28, SE = 0.31 and g = 1.42, SE = 0.39). After removing the studies, the overall meta-analytic mean was $\bar{g}$ = 0.32, SE = 0.06, 95% CI [0.19; 0.44], k = 10, p < .001. The test of heterogeneity was not significant, Q = 8.95, I^2 = 0.00%, p = .442. PET and PEESE estimators were $\bar{g}$ = 0.01, SE = 0.17, p = .943, and $\bar{g}$ = 0.15, SE = 0.09, p = .124, respectively. With all the three estimators (L0, R0, and Q0), the analysis filled three studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = 0.26, SE = 0.06, 95% CI [0.14; 0.38], k = 13, p < .001, I^2 = 0.01%. Considering that no included study reported a null or negative effect, the PET estimator was considered unreliable. The PEESE estimator and the trim-and-fill estimator were realistic. The trim-and-fill estimator ($\bar{g}$ = 0.26, SE = 0.06) was preferred because the effect size of the studies with the largest sample size was about g = 0.25.

Adults.

The overall meta-analytic mean was $\bar{g}$ = 0.20, SE = 0.04, 95% CI [0.12; 0.28], k = 31, p < .001. The test of heterogeneity was not significant, Q = 38.85, I^2 = 31.99%, p = .129. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 5).

Insert Figure 5 around here

Figure 5. Funnel plot of standard errors and effect sizes (gs) in the meta-analysis of WM training in the adults (near transfer).

PET and PEESE estimators were $\bar{g}$ = 0.15, SE = 0.05, p = .009 and $\bar{g}$ = 0.17, SE = 0.04, p < .001, respectively. The PEESE estimator was selected for the second-order meta-analysis.

Concerning the type of control group (active vs. non-active), there was 93% inter-rater agreement. The two coders solved any discrepancy by talk. The classification of the type of control group was slightly different from the one employed by Melby-Lervåg et al. (2016). Three studies whose control group was considered as active in the original meta-analysis were labeled as "non-active." When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g} = 0.15$, $SE = 0.04$, 95% CI [0.06; 0.23], $k = 20$, $p = .001$. The test of heterogeneity was not significant, $Q = 19.91$, $I^2 = 21.45\%$, $p = .400$. The distribution of the effect sizes in the funnel plot looked asymmetrical (Figure 6).

Insert Figure 6 around here

Figure 6. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the adults (near transfer, active controls).

PET and PEESE estimators were $\bar{g} = 0.15$, $SE = 0.06$, $p = .017$ and $\bar{g} = 0.14$, $SE = 0.04$, $p = .002$, respectively. The PEESE estimator was selected for the second-order meta-analysis.

Older adults.

The overall meta-analytic mean was $\bar{g} = 0.29$, $SE = 0.04$, 95% CI [0.20; 0.37], $k = 35$, $p < .001$. The test of heterogeneity was significant, $Q = 51.14$, $I^2 = 36.58\%$, $p = .030$. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 7).

Insert Figure 7 around here

Figure 7. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the older adults (near transfer).

PET and PEESE estimators were $\bar{g} = 0.09$, $SE = 0.08$, $p = .289$ and $\bar{g} = 0.19$, $SE = 0.05$, $p < .001$, respectively. The influential case analysis found one influential study ($g = 0.81$, $SE = 0.16$). After removing this study, the overall meta-analytic mean was $\bar{g} = 0.25$, $SE = 0.04$, 95% CI [0.18; 0.32], $k = 34$, $p < .001$. The test of heterogeneity was not significant, $Q = 38.83$, $I^2 = 18.27\%$, $p = .224$. PET and PEESE estimators were $\bar{g} = 0.06$, $SE = 0.07$, $p = .371$, and $\bar{g} = 0.16$, $SE = 0.04$, $p < .001$, respectively. The PEESE estimator was preferred over the PET estimator because it was more precise (*SE* = 0.04 vs. *SE* = 0.07). Also, given that the average effect size of the studies with the largest sample size was about *g* = 0.15 and only a few studies reported an effect size close to zero, the adjusted effect size suggested by the PET test seemed too small.

Concerning the type of control group (active vs. non-active), there was 100% inter-rater agreement. When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g}$ = 0.23, *SE* = 0.05, 95% CI [0.12; 0.33], *k* = 19, *p* < .001. The test of heterogeneity was not significant, *Q* = 19.32, I^2 = 20.09%, *p* = .372. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 8).

Insert Figure 8 around here

Figure 8. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the older adults (near transfer, active controls).

PET and PEESE estimators were $\bar{g}$ = 0.12, *SE* = 0.14, *p* = .436 and $\bar{g}$ = 0.18, *SE* = 0.07, *p* = .022, respectively. The PET estimator was selected for the second-order meta-analysis.

Second-order meta-analysis.

Tables 1-4 summarized the results of the second-order meta-analysis of near-transfer effects of WM training. The differences between the first-order meta-analytic means ($\bar{g}_i$) were mostly due to true variance (σ^2) in all the four second-order meta-analytic models. The adjusted overall effect sizes (column 11) were thus very close to the original estimates (column 2).

Table 1

Second-order meta-analysis with the uncorrected (naïve) overall effect sizes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	${S^2}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
TD Children	16	0.45	0.076	5.74							0.41
LD Children	17	0.55	0.180	74.12							0.49
Adults	31	0.20	0.052	31.99							0.23
Older Adults	35	0.29	0.065	36.58							0.29
					0.29	0.00279	0.01113	0.00834	.25	.75	

Note. (1) Number of samples; (2) First-order overall effect size; (3) Variance of the observed *g*s; (4) Degree of true heterogeneity; (5) Second-order grand mean; (6) Second-order sampling error variance; (7) Observed between-first-order-meta-analysis variance; (8) True between-first-order-meta-analysis variance; (9) Proportion of the variance explained by second-order sampling error; (10) Reliability of the first-order overall effect size; (11) Adjusted first-order overall effect size.

Table 2

Second-order meta-analysis with the corrected overall effect sizes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	${S^2}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
TD Children	16	0.50	0.188	5.74							0.42
LD Children	16	0.37	0.081	54.70							0.32
Adults	31	0.17	0.046	31.99							0.18
Older Adults	34	0.16	0.054	18.27							0.18
					0.21	0.00252	0.00912	0.00660	.28	.72	

Note. See Note to Table 1 for abbreviations.

Table 3

Second-order meta-analysis with the uncorrected (naïve) overall effect sizes (only active control groups)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	${S^2}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
TD Children	11	0.43	0.087	0.00							0.37
LD Children	12	0.48	0.142	62.36							0.41
Adults	20	0.15	0.040	21.45							0.17
Older Adults	19	0.23	0.054	20.09							0.23
					0.23	0.00375	0.01307	0.00931	.29	.71	

Note. See Note to Table 1 for abbreviations.

Table 4

Second-order meta-analysis with the corrected overall effect sizes (only active control groups)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	${S^2}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
TD Children	11	0.49	0.180	0.00							0.36
LD Children	13	0.26	0.047	0.01							0.23
Adults	20	0.14	0.030	21.45							0.17
Older Adults	19	0.12	0.399	20.09							0.15
					0.19	0.00380	0.00822	0.00442	.46	.54	

Note. See Note to Table 1 for abbreviation.

Omnibus meta-analysis.

The overall meta-analytic mean was $\bar{g}$ = 0.32, *SE* = 0.03, 95% CI [0.26; 0.38], *k* = 99, *p* < .001. The test of heterogeneity was significant, *Q* = 188.23, I^2 = 51.81%, *p* < .001. The distribution of the effect sizes in the funnel plot looked asymmetrical (Figure 9).

Insert Figure 9 around here

Figure 9. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in all the samples (near transfer).

PET and PEESE estimators were $\bar{g}$ = 0.14, *SE* = 0.05, *p* = .003 and $\bar{g}$ = 0.22, *SE* = 0.03, *p* < .001, respectively. The influential case analysis found three influential studies (*g* = 0.74, *SE* = 0.11; *g* = 1.40, *SE* = 0.22; *g* = 0.81, *SE* = 0.16). After removing these studies, the overall meta-analytic mean was $\bar{g}$ = 0.28, *SE* = 0.03, 95% CI [0.23; 0.33], *k* = 96, *p* < .001. The test of heterogeneity was still significant, and the degree of true heterogeneity was lower, *Q* = 132.42, I^2 = 30.72%, *p* = .007. PET and PEESE estimators were $\bar{g}$ = 0.11, *SE* = 0.04, *p* = .004, and $\bar{g}$ = 0.19, *SE* = 0.02, *p* < .001, respectively. Both the naïve and corrected overall effect sizes ($\bar{g}$ = 0.32 and $\bar{g}$ = 0.19, respectively) were thus very close to the grand means estimated in the second-order meta-analysis ($\bar{\bar{g}}$ = 0.29 and $\bar{\bar{g}}$ = 0.21 for the uncorrected and corrected models, respectively).

When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g}$ = 0.25, *SE* = 0.03, 95% CI [0.19; 0.31], *k* = 62, *p* < .001. The test of heterogeneity was significant, *Q* = 85.39, I^2 = 29.82%, *p* = .021. The distribution of the effect sizes in the funnel plot looked asymmetrical (Figure 10).

Insert Figure 10 around here

Figure 10. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in all the samples (near transfer, active controls).

PET and PEESE estimators were $\bar{g}$ = 0.09, *SE* = 0.05, *p* = .077 and $\bar{g}$ = 0.16, *SE* = 0.03, *p* < .001, respectively. The influential case analysis found three influential studies (*g* = 0.12, *SE* = 0.07; *g* = 0.20, *SE* = 0.06; *g* = 1.28, *SE* = 0.33). After removing these studies, the overall meta-analytic mean was $\bar{g}$ = 0.25, *SE* = 0.03, 95% CI [0.18; 0.32], *k* = 59, *p* < .001. The test of heterogeneity was marginally significant, *Q* = 72.15, I^2 = 23.27%, *p* = .100. PET and PEESE estimators were $\bar{g}$ = 0.05, *SE* = 0.08, *p* = .555, and $\bar{g}$ = 0.16, *SE* = 0.04, *p* < .001, respectively. Again, the PET estimator appeared to be an underestimation of the true effect, especially because the PET test showed the presence of a non-zero effect when the influential cases were included in the model (*p* = .077). On the contrary, the PEESE estimator was more precise and very close to the grand mean in the second-order meta-analysis ($\bar{\bar{g}}$ = 0.19). Finally, in line with the results of the second-order meta-analysis, some true heterogeneity was observed in all the models of the omnibus meta-analysis.

Discussion

The results show that WM training fosters performance in memory tasks in all the reviewed populations. The effect substantially maintains even when only comparisons between trained groups and active controls are considered. The effect is, however, quite heterogeneous across populations. In fact, only a portion of the observed between-meta-analysis variance is due to second-order sampling error (25% to 46%).

TD children seem to benefit the most from the training program. Adults and older adults exhibit much smaller effects. Due to the highly asymmetrical distribution of the effect sizes, the effect is less clear in LD children. As seen, the most probable unbiased effect is about $\bar{g}$ = 0.25. This difference probably reflects the different learning pace in the populations. While TD children learn relatively fast, adults and older adults require greater effort to acquire new skills probably because their cognitive system is less flexible and plastic. LD children exhibit better performance than the adult populations but, as expected, are not as good as TD children.

Thus, in line with previous research, near transfer from WM training to memory often (if not always) takes place. Interestingly, the size of this transfer of skills appears to be moderated by the type of population and, arguably, their particular cognitive profiles. It is reasonable that near transfer represents (i.e., exhibits the same pattern of) the general capability of acquiring new skills by practice. In fact, TD children usually learn faster than their peers suffering from some learning disability, adults, and older adults.

On a final note, it is essential to acknowledge that the participants' boosted performance on memory tasks does not represent evidence of cognitive enhancement. As observed by Shipstead, Redick, and Engle (2012), such an improvement probably denotes the participants' enhanced ability to solve a class of cognitive tasks sharing similar features (e.g., perceptual cues and strategies) with the trained task. Either way, regardless of whether it represents cognitive enhancement, the presence of near transfer seems unquestionable.

Model 2: Working Memory Training (Far Transfer)

Description of The First-Order Meta-Analyses

This section examines the effects of WM training on far transfer measures such as tests of fluid reasoning, cognitive control, processing speed, and language. For the details about the four first-order meta-analyses included, see Model 1 section.

Results

TD children.

The overall meta-analytic mean was $\bar{g}$ = 0.13, *SE* = 0.05, 95% CI [0.03; 0.22], k = 25, p = .010. The test of heterogeneity was not significant, Q = 21.75, I^2 = 9.79%, p = .594. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 11).

Insert Figure 11 around here

Figure 11. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the TD children (far transfer).

PET and PEESE estimators were $\bar{g}$ = 0.08, *SE* = 0.10, p = .441 and $\bar{g}$ = 0.09, *SE* = 0.06, p = .145, respectively. The PET estimator was selected for the second-order meta-analysis.

When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g}$ = 0.01, *SE* = 0.07, 95% CI [-0.12; 0.14], k = 15, p = .879. The test of heterogeneity was not significant, Q = 7.20, I^2 = 0.00%, p = .927. Again, the distribution of the effect sizes in the funnel plot looked symmetrical (Figure 12).

Insert Figure 12 around here

Figure 12. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the TD children (far transfer, active controls).

PET and PEESE estimators were $\bar{g}$ = 0.12, SE = 0.15, p = .440 and $\bar{g}$ = 0.06, SE = 0.08, p = .495, respectively. Neither PET nor PEESE provided any evidence of a non-zero effect (p = .440 and p = .495, respectively). Also, PET estimated a very imprecise overall effect (SE = 0.15) suggesting that the detected publication bias was in fact a statistical artifact. We thus ran the trim-and-fill analysis to check for missing studies right of the mean. With the L0 and Q0 estimators, the analysis filled no study. With the R0 estimator, the analysis filled one study right of the mean. The overall meta-analytic mean was $\bar{g}$ = 0.02, SE = 0.06, 95% CI [-0.10; 0.15], k = 16, p = .717, I^2 = 0.00%. Given that no estimate was significantly different from the null (all ps ≥ .440), we selected the estimate that was the closest to zero ($\bar{g}$ = 0.01, SE = 0.07) as the corrected overall effect size to be included in the second-order meta-analysis.

LD children.

The overall meta-analytic mean was $\bar{g}$ = 0.12, SE = 0.04, 95% CI [0.03; 0.20], k = 18, p = .006. The test of heterogeneity was not significant, Q = 12.34, I^2 = 4.72%, p = .779. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 13).

Insert Figure 13 around here

Figure 13. Funnel plot of standard errors and effect sizes (gs) in the meta-analysis of WM training in the LD children (far transfer).

PET and PEESE estimators were $\bar{g}$ = 0.06, SE = 0.07, p = .350 and $\bar{g}$ = 0.10, SE = 0.03, p = .014, respectively. The PET estimator was selected for the second-order meta-analysis.

When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g}$ = 0.08, SE = 0.06, 95% CI [-0.04; 0.19], k = 12, p = .202. The test of heterogeneity was not significant, Q = 3.50, I^2 = 0.00%, p = .982. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 14).

Insert Figure 14 around here

Figure 14. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the LD children (far transfer, active controls).

PET and PEESE estimators were $\bar{g} = 0.02$, $SE = 0.10$, $p = .870$ and $\bar{g} = 0.03$, $SE = 0.06$, $p = .543$, respectively. The PET estimator was selected for the second-order meta-analysis.

Adults.

The overall meta-analytic mean was $\bar{g} = 0.12$, $SE = 0.03$, 95% CI [0.06; 0.18], $k = 44$, $p < .001$. The test of heterogeneity was not significant, $Q = 39.52$, $I^2 = 6.75\%$, $p = .623$. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 15).

Insert Figure 15 around here

Figure 15. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the adults (far transfer).

PET and PEESE estimators were $\bar{g} = 0.06$, $SE = 0.05$, $p = .298$ and $\bar{g} = 0.08$, $SE = 0.03$, $p = .011$, respectively. The PET estimator was selected for the second-order meta-analysis.

When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g} = 0.09$, $SE = 0.04$, 95% CI [0.01; 0.17], $k = 27$, $p = .032$. The test of heterogeneity was not significant, $Q = 20.06$, $I^2 = 0.00\%$, $p = .789$. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 16).

Figure 16. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the adults (far transfer, active controls).

PET and PEESE estimators were $\bar{g} = 0.00$, $SE = 0.09$, $p = .966$ and $\bar{g} = 0.04$, $SE = 0.05$, $p = .418$, respectively. The PET estimator was selected for the second-order meta-analysis.

Older adults.

The overall meta-analytic mean was $\bar{g}$ = 0.13, *SE* = 0.05, 95% CI [0.03; 0.23], k = 32, p = .010. The test of heterogeneity was significant, Q = 62.19, I^2 = 55.83%, p < .001. The distribution of the effect sizes in the funnel plot looked asymmetrical (Figure 17).

Insert Figure 17 around here

Figure 17. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the older adults (far transfer).

PET and PEESE estimators were $\bar{g}$ = -0.02, *SE* = 0.06, p = .696 and $\bar{g}$ = 0.03, *SE* = 0.05, p = .540, respectively. The influential case analysis found one influential study (g = 0.63, *SE* = 0.14). After removing this study, the overall meta-analytic mean was $\bar{g}$ = 0.10, *SE* = 0.04, 95% CI [0.01; 0.18], k = 31, p = .029. The test of heterogeneity was still significant, and the degree of heterogeneity was lower, Q = 45.05, I^2 = 37.35%, p = .038. The distribution of the effect sizes in the funnel plot looked asymmetrical. PET and PEESE estimators were $\bar{g}$ = -0.04, *SE* = 0.05, p = .453, and $\bar{g}$ = 0.00, *SE* = 0.04, p = .909, respectively. Since the PET estimator was negative, we selected the PEESE estimator for the second-order meta-analysis.

When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g}$ = -0.02, *SE* = 0.03, 95% CI [-0.09; 0.05], k = 16, p = .582. The test of heterogeneity was not significant, Q = 6.59, I^2 = 0.00%, p = .968. The distribution of the effect sizes in the funnel plot looked symmetrical (Figure 18).

Insert Figure 18 around here

Figure 18. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of WM training in the older adults (far transfer, active controls).

PET and PEESE estimators were $\bar{g}$ = 0.02, *SE* = 0.03, *p* = .488 and $\bar{g}$ = 0.01, *SE* = 0.02, *p* = .792, respectively. Since the PET test did not find any evidence of a non-zero effect, we selected the PEESE estimator for the second-order meta-analysis because it was the closest to zero.

Second-order meta-analysis.

Tables 5-8 summarized the results of the second-order meta-analysis of far-transfer effects of WM training. When ruling out placebo effects and publication bias, the differences between the first-order meta-analytic means ($\bar{g}_i$) were mostly due solely to second-order sampling error (i.e., σ^2 = 0; Table 8). We estimated the unbiased overall effect to be $\bar{\bar{g}}$ = 0.01.

Table 5

Second-order meta-analysis with the uncorrected (naïve) overall effect sizes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	$S^2{}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
TD Children	25	0.13	0.060	9.79							0.12
LD Children	18	0.12	0.032	4.72							0.12
Adults	44	0.12	0.041	6.75							0.12
Older Adults	32	0.13	0.084	55.83							0.12
					0.12	0.00163	0.00003	0	1	0	

Note. See Note to Table 1 for abbreviations.

Table 6

Second-order meta-analysis with the corrected overall effect sizes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	$S^2{}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
TD Children	25	0.08	0.244	9.79							0.03
LD Children	18	0.06	0.079	4.72							0.03
Adults	44	0.06	0.125	6.75							0.03
Older Adults	31	0.00	0.048	37.35							0.03
					0.03	0.00302	0.00089	0	1	0	

Note. See Note to Table 1 for abbreviations.

Table 7

Second-order meta-analysis with the uncorrected (naïve) overall effect sizes (only active control groups)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	$S^2{}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
TD Children	15	0.01	0.064	0.00							0.03
LD Children	12	0.08	0.043	0.00							0.04
Adults	27	0.09	0.049	0.00							0.04
Older Adults	16	-0.02	0.018	0.00							0.02
					0.03	0.00207	0.00250	0.0043	.83	.17	

Note. See Note to Table 1 for abbreviations.

Table 8

Second-order meta-analysis with the corrected overall effect sizes (only active control groups)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	$S^2{}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
TD Children	15	0.01	0.064	0.00							0.01
LD Children	12	0.02	0.111	0.00							0.01
Adults	27	0.00	0.213	0.00							0.01
Older Adults	16	0.01	0.009	0.00							0.01
					0.01	0.00185	0.00001	0	1	0	

Note. See Note to Table 1 for abbreviations.

Omnibus meta-analysis.

The overall meta-analytic mean was $\bar{g}$ = 0.12, *SE* = 0.02, 95% CI [0.08; 0.16], k = 119, p < .001. The test of heterogeneity was marginally significant, Q = 138.34, I^2 = 20.89%, p = .097. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 19).

Insert Figure 19 around here

Figure 19. Funnel plot of standard errors and effect sizes (gs) in the meta-analysis of WM training in all the samples (far transfer).

PET and PEESE estimators were $\bar{g}$ = 0.02, *SE* = 0.03, p = .478 and $\bar{g}$ = 0.07, *SE* = 0.02, p < .001, respectively. The influential case analysis found nine influential studies (g = 0.08, *SE* = 0.08; g = -0.04, *SE* = 0.08; g = 0.29, *SE* = 0.09; g = 0.00, *SE* = 0.08; g = 0.63, *SE* = 0.14; g = 0.77, *SE* = 0.21; g = -0.01, *SE* = 0.04; g = -0.04, *SE* = 0.07; g = 0.42, *SE* = 0.13). After removing these studies, the overall meta-analytic mean was $\bar{g}$ = 0.12, *SE* = 0.02, 95% CI [0.08; 0.16], k = 110, p < .001. The test of heterogeneity was not significant, Q = 86.51, I^2 = 0.00%, p = .945. PET and PEESE estimators were $\bar{g}$ = 0.06, *SE* = 0.04, p = .177, and $\bar{g}$ = 0.08, *SE* = 0.02, p < .001, respectively.

When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g}$ = 0.03, *SE* = 0.02, 95% CI [-0.02; 0.07], k = 70, p = .194. The test of heterogeneity was not significant, Q = 42.11, I^2 = 0.00%, p = .996. The distribution of the effect sizes in the funnel plot looked substantially symmetrical (Figure 20).

Insert Figure 20 around here

Figure 20. Funnel plot of standard errors and effect sizes (gs) in the meta-analysis of WM training in all the samples (near transfer, active controls).

PET and PEESE estimators were $\bar{g} = 0.00$, $SE = 0.03$, $p = .909$ and $\bar{g} = 0.01$, $SE = 0.02$, $p = .572$, respectively.

Discussion

The results provided by the second-order meta-analytic models show that the actual impact of WM training on far-transfer measures is null regardless of the population. The small positive effect sizes disappear when placebo effects and publication bias are controlled for (Table 8). That is, the unbiased far-transfer effect exerted by WM training is practically null ($\bar{\bar{g}} = 0.01$) and consistent ($I^2 = 0$). No or low true variance (σ^2) is observed in all the models. Also, it is worth noting that the observed degree of true heterogeneity (I^2; Tables 5 and 6) is entirely accounted for by the type of control group used in the primary studies (Tables 7 and 8). Consequently, comparing effect sizes extracted from different tests of cognitive/academic skills does not add noise to the models. In fact, unbiased far-transfer effects approach zero regardless of the test used to measure them. Finally, the omnibus meta-analysis confirms the findings of the second-order meta-analysis.

Model 3: Other Cognitive-Training Programs

Description of The First-Order Meta-Analyses

This section presents an extension of the meta-analytic models presented in Model 2. Along with the four first-order meta-analyses of WM training, we add other four first-order meta-analyses of the far-transfer effects of action video-game training, music training, chess training, and exergame training.

The action video-game training meta-analysis was a subsample of Sala, Tatlidil, and Gobet (2018). We included only studies examining the effects the training on the adult participants' cognitive skills. Children and older adults were excluded because the number of primary studies was too small. In line with the specific features of this field of research, the controls playing non-action video games were considered active. All the other comparisons (mostly passive controls) were labeled as non-active. There was 100% inter-rater agreement and the classification was the same as in Sala et al. (2018).

The music-training meta-analysis was a subsample of Sala and Gobet (2017c). We excluded two studies that did not administer a pre-test. All the samples consisted of groups of either children or young adolescents. Regarding the active control groups, the classification was the same as in Sala and Gobet (2017c), and the inter-rater agreement was 100%.

To the best of our knowledge, the only meta-analysis of the effects of chess-based interventions on far transfer measures was Sala and Gobet (2016). Like the music-training meta-analysis, the population consisted of children and young adolescents. Sixteen out of 24 of the studies included in this meta-analysis did not meet our inclusion criteria. These studies and effect sizes were excluded because they were derived from (a) questionnaires rather than cognitive/academic tests, (b) self-selected samples, or (c) because they had no pre-test. All the effect sizes were recalculated. Also, a new study (Jerrim, Macmillan, Micklewright, Sawtell, &

Wiggins, 2016) was added because it was the largest experiment in the field (*N* about 4000). The final sample consisted of nine studies. Regarding the type of control group (active vs. non-active), there was 100% inter-rater agreement.

The exergame-training meta-analysis was a subsample of Stanmore, Stubbs, Vancampfort, de Bruin, and Firth (2017). The original meta-analysis included 17 RCTs, most of them analyzing the effect of the intervention on older adults' cognitive function. We excluded the two studies analyzing the effects of the training program in younger populations. We also excluded one study that did not implement any exergame intervention and two studies that did not report data enough to calculate an effect size. All the effect sizes were recalculated. Finally, since this particular type of cognitive-training compares the effects of physical training combined with cognitive-training games, groups who underwent physical activities were considered as active. Concerning the type of control group (active vs. non-active), there was 100% inter-rater agreement and the classification was the same as in Stanmore et al. (2017).[3]

Results

Action video-game training.

The overall meta-analytic mean was $\bar{g} = 0.08$, $SE = 0.05$, 95% CI [-0.01; 0.17], $k = 32$, $p = .094$. The test of heterogeneity was not significant, $Q = 39.76$, $I^2 = 0.00\%$, $p = .135$. The distribution of the effect sizes in the funnel plot looked asymmetrical (Figure 21).

Insert Figure 21 around here

[3] It is worth noting that the results we report here are sometimes different from the ones of the original meta-analyses for a variety of reasons, not only for the exclusion of some studies. First, in the original publications, the authors used different methods to model nested effect sizes. For example, in some cases the authors applied (a) no correction with either merged or unmerged effects (Melby-Lervåg et al., 2016; Sala & Gobet, 2016; Stanmore et al., 2017), (b) corrections based on Cheung and Chan's (2004,2008) method (Sala & Gobet, 2017b,c), or (c) used multiple approaches (Sala et al., 2018). Second, the original meta-analyses employed different publication-bias detection techniques (e.g., p-curve, selection models, and trim-and-fill). Despite these differences, both uncorrected and corrected estimates are very close to the ones reported in the original meta-analyses in most of the cases. The only significant exception is Stanmore et al. (2017). This meta-analysis, however, shows a highly asymmetrical distribution of the effect sizes suggesting the presence of severe publication bias.

Figure 21. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of action video-game training.

PET and PEESE estimators were $\bar{g}$ = -0.42, *SE* = 0.10, $p < .001$ and $\bar{g}$ = -0.11, *SE* = 0.05, *p* = .046, respectively. Since both the corrected estimates were excessively negative, we ran a trim-and-fill analysis. With the L0 estimator, the analysis filled 11 studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = -0.01, *SE* = 0.05, 95% CI [-0.09; 0.08], *k* = 43, *p* = .901, I^2 = 0.00%. With the R0 estimator, the analysis filled seven studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = 0.03, *SE* = 0.05, 95% CI [-0.06; 0.12], *k* = 39, *p* = .468, I^2 = 0.00%. With the Q0 estimator, the analysis filled 15 studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = -0.03, *SE* = 0.05, 95% CI [-0.13; 0.06], *k* = 47, *p* = .519, I^2 = 10.01%. Since no publication-bias corrected estimate provided any evidence of a non-zero effect, the estimate that was the closest to zero ($\bar{g}$ = -0.01, *SE* = 0.05) was selected for the second order meta-analysis.

When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g}$ = 0.10, *SE* = 0.06, 95% CI [-0.02; 0.21], *k* = 25, *p* = .109. The test of heterogeneity was not significant, *Q* = 31.74, I^2 = 12.81%, *p* = .134. The distribution of the effect sizes in the funnel plot looked asymmetrical (Figure 22).

Insert Figure 22 around here

Figure 22. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of action video-game training (active controls).

PET and PEESE estimators were $\bar{g}$ = -0.43, *SE* = 0.09, $p < .001$ and $\bar{g}$ = -0.11, *SE* = 0.05, *p* = .037, respectively. Since both the corrected estimates were excessively negative, we ran a trim-and-fill analysis. With the L0 estimator, the analysis filled nine studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = -0.01, *SE* = 0.06, 95% CI [-0.12; 0.10], *k* = 34, *p* = .919, I^2 = 10.54%. With the R0 estimator, the analysis filled 11 studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = -0.02, *SE* = 0.06, 95% CI [-0.13; 0.09], $k = 36$, $p = .674$, $I^2 = 12.70\%$. With the Q0 estimator, the analysis filled 15 studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = -0.08, *SE* = 0.07, 95% CI [-0.21; 0.04], $k = 40$, $p = .200$, $I^2 = 40.08\%$. Again, the estimate that was the closest to zero ($\bar{g}$ = -0.01, *SE* = 0.06) was selected for the second-order meta-analysis.

Music.

The overall meta-analytic mean was $\bar{g}$ = 0.19, *SE* = 0.05, 95% CI [0.10; 0.29], $k = 36$, $p < .001$. The test of heterogeneity was significant, $Q = 93.00$, $I^2 = 61.20\%$, $p < .001$. The distribution of the effect sizes in the funnel plot looked asymmetrical (Figure 23).

Insert Figure 23 around here

Figure 23. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of music training.

PET and PEESE estimators were $\bar{g}$ = -0.15, *SE* = 0.07, $p = .039$, and $\bar{g}$ = 0.00, *SE* = 0.05, $p = .921$, respectively. The influential case analysis found one influential study ($g = 1.23$, *SE* = 0.26). After removing the study, the overall meta-analytic mean was $\bar{g}$ = 0.16, *SE* = 0.04, 95% CI [0.07; 0.25], $k = 35$, $p < .001$. The test of heterogeneity was still significant, and the degree of heterogeneity was slightly lower, $Q = 72.64$, $I^2 = 51.20\%$, $p < .001$. PET and PEESE estimators were $\bar{g}$ = -0.14, *SE* = 0.06, $p = .039$, and $\bar{g}$ = 0.00, *SE* = 0.04, $p = .938$, respectively. Since the PET estimator was negative, the PEESE estimator was selected for the second-order meta-analysis because it was the closest to zero.

When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g}$ = 0.03, *SE* = 0.06, 95% CI [-0.10; 0.15], $k = 18$, $p = .678$. The test of heterogeneity was marginally significant, $Q = 26.99$, $I^2 = 41.74\%$, $p = .058$. The distribution of the effect sizes in the funnel plot looked substantially symmetrical (Figure 24).

Insert Figure 24 around here

Figure 24. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of music training (active controls).

PET and PEESE estimators were $\bar{g}$ = -0.20, *SE* = 0.08, *p* =.022 and $\bar{g}$ = -0.10, *SE* = 0.06, *p* = .122, respectively. With the L0 estimator, the analysis filled six studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = -0.09, *SE* = 0.07, 95% CI [-0.23; 0.05], *k* = 24, *p* = .189, I^2 = 61.69%. With the R0 estimator, the analysis filled no study. With the Q0 estimator, the analysis filled nine studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = -0.15, *SE* = 0.07, 95% CI [-0.28; -0.01], *k* = 27, *p* = .037, I^2 = 66.38%. The influential case analysis found one influential study (*g* = 0.55, *SE* = 0.22). After removing the study, the overall meta-analytic mean was $\bar{g}$ = -0.02, *SE* = 0.06, 95% CI [-0.13; 0.09], *k* = 17, *p* = .704. The test of heterogeneity was not significant, *Q* = 18.92, I^2 = 28.08%, *p* = .273. PET and PEESE estimators were $\bar{g}$ = -0.20, *SE* = 0.07, *p* = .012, and $\bar{g}$ = -0.11, *SE* = 0.05, *p* = .051, respectively. With the L0 estimator, the analysis filled five studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = -0.10, *SE* = 0.05, 95% CI [-0.20; 0.01], *k* = 22, *p* = .076, I^2 = 29.21%. With the R0 estimator, the analysis filled no study. With the Q0 estimator, the analysis filled eight studies left of the mean. The overall meta-analytic mean was $\bar{g}$ = -0.14, *SE* = 0.06, 95% CI [-0.26; -0.03], *k* = 25, *p* = .012, I^2 = 44.71%. All the corrected and uncorrected estimates were either negative or non-significantly different from zero. We thus selected the estimate that was the closest to zero ($\bar{g}$ = -0.02, *SE* = 0.06) for the second-order meta-analysis.

Chess.

The overall meta-analytic mean was $\bar{g}$ = 0.13, *SE* = 0.07, 95% CI [-0.02; 0.27], *k* = 9, *p* = .089. The test of heterogeneity was significant, *Q* = 56.58, I^2 = 84.22%, *p* < .001. The distribution of the effect sizes in the funnel plot looked symmetrical (Figure 25).

Insert Figure 25 around here

Figure 25. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of chess training.

PET and PEESE estimators were $\bar{g}$ = 0.12, *SE* = 0.11, *p* = .300, and $\bar{g}$ = 0.13, *SE* = 0.07, *p* = .094, respectively. No influential case was found. The PET estimate was selected for the second-order meta-analysis.

When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g}$ = 0.05, *SE* = 0.10, 95% CI [-0.15; 0.25], *k* = 3, *p* = .623. The test of heterogeneity was not significant, *Q* = 0.47, I^2 = 0.00%, *p* = .791. The funnel plot looked slightly asymmetrical (but included only three effect sizes; Figure 26).

Insert Figure 26 around here

Figure 26. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of chess training (active controls).

PET and PEESE estimators were $\bar{g}$ = -0.05, *SE* = 0.22, *p* = .849 and $\bar{g}$ = 0.01, *SE* = 0.10, *p* = .927, respectively. The PEESE estimate was selected for the second-order meta-analysis because it was the closest to zero.

Exergames.

The overall meta-analytic mean was $\bar{g}$ = 0.15, *SE* = 0.08, 95% CI [-0.01; 0.32], *k* = 11, *p* = .071. The test of heterogeneity was not significant, *Q* = 13.96, I^2 = 33.81%, *p* = .175. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 27).

Insert Figure 27 around here

Figure 27. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of exergame training.

PET and PEESE estimators were $\bar{g}$ = -0.03, *SE* = 0.07, *p* = .659 and $\bar{g}$ = 0.03, *SE* = 0.05, *p* = .554, respectively. The PET estimate was selected for the second-order meta-analysis.

When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g}$ = 0.08, *SE* = 0.05, 95% CI [-0.02; 0.17], *k* = 8, *p* = .131. The test of heterogeneity was not significant, *Q* = 10.85, I^2 = 0.01%, *p* = .145. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 28).

Insert Figure 28 around here

Figure 28. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of exergame training (active controls).

PET and PEESE estimators were $\bar{g}$ = -0.02, *SE* = 0.09, *p* = .835 and $\bar{g}$ = 0.03, *SE* = 0.06, *p* = .620 respectively. The PET estimate was selected for the second-order meta-analysis.

Second-order meta-analysis.

Tables 9-12 summarized the results of the second-order meta-analysis of far-transfer effects for Model 3. Like in Model 2, publication-bias corrected estimates in studies implementing an active control group are all around zero. Second-order sampling error accounted for the observed variability (i.e., σ^2 = 0; Table 12). We estimated the unbiased overall effect size to be $\bar{\bar{g}}$ = 0.00.

Table 9

Second-order meta-analysis with the uncorrected (naïve) overall effect sizes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	${S^2}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
WM (TD Children)	25	0.13	0.060	9.79							0.13
WM (LD Children)	18	0.12	0.032	4.72							0.13
WM (Adults)	44	0.12	0.041	6.75							0.13
WM (Older Adults)	32	0.13	0.084	55.83							0.13
Action VG (Adults)	32	0.08	0.073	0.00							0.13
Music (TD Children)	36	0.19	0.087	61.20							0.13
Chess (TD Children)	9	0.13	0.049	84.22							0.13
Exergames (Older Adults)	11	0.15	0.079	33.81							0.13
					0.13	0.00221	0.00086	0	1	0	

Note. See Note to Table 1 for abbreviations.

Table 10

Second-order meta-analysis with the corrected overall effect sizes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	${S^2}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
WM (TD Children)	25	0.08	0.244	9.79							0.02
WM (LD Children)	18	0.06	0.079	4.72							0.02
WM (Adults)	44	0.06	0.125	6.75							0.02
WM (Older Adults)	31	0.00	0.048	37.35							0.02
Action VG (Adults)	43	-0.01	0.087	0.00							0.02
Music (TD Children)	35	0.00	0.067	51.20							0.02
Chess (TD Children)	9	0.12	0.099	84.22							0.02
Exergames (Older Adults)	11	-0.03	0.061	33.81							0.02
					0.02	0.00306	0.00124	0	1	0	

Note. See Note to Table 1 for abbreviations.

Table 11

Second-order meta-analysis with the uncorrected (naïve) overall effect sizes (only active control groups)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	$S^2{}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
WM (TD Children)	15	0.01	0.064	0.00							0.04
WM (LD Children)	12	0.08	0.043	0.00							0.04
WM (Adults)	27	0.09	0.049	0.00							0.04
WM (Older Adults)	16	-0.02	0.018	0.00							0.04
Action VG (Adults)	25	0.10	0.089	12.81							0.04
Music (TD Children)	18	0.03	0.071	41.74							0.04
Chess (TD Children)	3	0.05	0.031	0.00							0.04
Exergames (Older Adults)	8	0.08	0.020	0.01							0.04
					0.04	0.00270	0.00217	0	1	0	

Note. See Note to Table 1 for abbreviations.

Table 12

Second-order meta-analysis with the corrected overall effect sizes (only active control groups)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Population	k_i	$\bar{g}_i$	$S^2{}_{g_i}$	I^2	$\bar{\bar{g}}$	σ_e^2	$\sigma_{\bar{\bar{g}}}^2$	σ^2	Pro_{var}	R_{xx}	$Adj.\bar{g}_i$
WM (TD Children)	15	0.01	0.064	0.00							0.00
WM (LD Children)	12	0.02	0.111	0.00							0.00
WM (Adults)	27	0.00	0.213	0.00							0.00
WM (Older Adults)	16	0.01	0.009	0.00							0.00
Action VG (Adults)	34	-0.01	0.107	10.54							0.00
Music (TD Children)	17	-0.02	0.055	28.08							0.00
Chess (TD Children)	3	0.01	0.032	0.00							0.00
Exergames (Older Adults)	8	-0.02	0.072	0.01							0.00
					0.00	0.00267	0.00011	0	1	0	

Note. See Note to Table 1 for abbreviations.

Omnibus meta-analysis.

The overall meta-analytic mean was $\bar{g} = 0.13$, $SE = 0.02$, 95% CI [0.10; 0.17], $k = 207$, $p < .001$. The test of heterogeneity was significant, $Q = 343.62$, $I^2 = 38.87\%$, $p < .001$. The distribution of the effect sizes in the funnel plot looked slightly asymmetrical (Figure 29).

Insert Figure 29 around here

Figure 29. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of all the samples in Model 3.

PET and PEESE estimators were $\bar{g} = 0.03$, $SE = 0.02$, $p = .224$, and $\bar{g} = 0.07$, $SE = 0.02$, $p < .001$, respectively. The influential case analysis found ten influential studies ($g = 0.63$, $SE = 0.14$; $g = 0.77$, $SE = 0.21$; $g = -0.01$, $SE = 0.04$; $g = 0.08$, $SE = 0.06$; $g = 1.23$, $SE = 0.26$; $g = -0.02$, $SE = 0.05$; $g = -0.25$, $SE = 0.08$; $g = 0.21$, $SE = 0.06$; $g = 0.37$, $SE = 0.05$; $g = 0.00$, $SE = 0.03$). After removing these studies, the overall meta-analytic mean was $\bar{g} = 0.12$, $SE = 0.02$, 95% CI [0.09; 0.15], $k = 197$, $p < .001$. The test of heterogeneity was not significant, $Q = 221.42$, $I^2 = 14.65\%$, $p = .103$. PET and PEESE estimators were $\bar{g} = 0.01$, $SE = 0.03$, $p = .846$, and $\bar{g} = 0.07$, $SE = 0.02$, $p < .001$, respectively.

When considering only studies implementing an active control group the overall meta-analytic mean was $\bar{g} = 0.04$, $SE = 0.02$, 95% CI [0.00; 0.08], $k = 124$, $p = .069$. The test of heterogeneity was not significant, $Q = 117.91$, $I^2 = 10.18\%$, $p = .613$. The distribution of the effect sizes in the funnel plot looked symmetrical (Figure 30).

Insert Figure 30 around here

Figure 30. Funnel plot of standard errors and effect sizes (*g*s) in the meta-analysis of all the samples in Model 3 (active controls).

PET and PEESE estimators were $\bar{g}$ = -0.08, *SE* = 0.03, *p* = .009 and $\bar{g}$ = -0.03, *SE* = 0.02, *p* = .142, respectively.

Discussion

Like in Model 2, when placebo effects and publication bias are controlled for, the actual impact of cognitive-training programs on far-transfer measures is null regardless of the training regimen employed or population examined. In all the models, the differences across the first-order meta-analytic means ($\bar{g}$s) are accounted for by second-order sampling error (σ^2 = 0). Also, all the corrected first-order meta-analytic means are associated with null or low degrees of true heterogeneity (none of them significant). This confirms that far transfer is null regardless the measure employed to assess it. The small or null overall effect sizes and near-zero degree of true heterogeneity in the omnibus meta-analysis corroborate the findings of the second-order meta-analytic models.

General Discussion

The present study is, to the best of our knowledge, the first second-order meta-analysis examining the transfer effects of cognitive-training programs. This large meta-analytic investigation (total $k = 306$) has addressed the question of the impact of a variety of cognitive-training programs on different populations' cognitive/academic skills. Our results confirm previous findings regarding transfer of skills. Near transfer frequently occurs, and, interestingly, it seems to be moderated by the type of population. On the contrary, far transfer is very modest at best. Moreover, once publication bias and placebo effects are ruled out, far-transfer effects are null regardless of the type of far-transfer measure, type of cognitive training program and population. This latter conclusion can be summarized by three equations:

$$Adj.\overline{g_\iota} = \overline{g_\iota} - (PB_i + PE_i) = 0$$

$$I^2_{Adj.\overline{g_\iota}} = 0$$

$$\sigma^2_{Adj.\bar{g}} = 0$$

where $Adj.\overline{g_\iota}$ is the adjusted overall effect size in the *i*th first-order meta-analysis, $\overline{g_\iota}$ is the naïve overall effect size, PB_i is publication bias in the *i*th first-order meta-analysis, PE_i is placebo effects in the *i*th first-order meta-analysis, $I^2_{Adj.\overline{g_\iota}}$ is the degree of true heterogeneity in the *i*th first-order meta-analysis with adjusted overall effect sizes, and $\sigma^2_{Adj.\bar{g}}$ is the true variance between first-order adjusted overall effect sizes.

Beyond first- and second-order meta-analytic evidence, the observed lack of generalized cognitive benefits is consistent with a well-established corpus of findings in other disciplines. For example, although education has been shown to exert an appreciable effect on academic outcomes, its impact on general intelligence or domain-general cognitive skills is substantially null (Detterman, 2016; Finn et al., 2014; Mosing, Madison, Pedersen, & Ullén, 2016; Ritchie et al., 2015). Also, as seen earlier, research into learning and the psychology of expertise has repeatedly shown that far transfer is rare because skill acquisition relies on domain-specific perceptual and conceptual information (Ericsson & Charness, 1994). Furthermore, other non-cognitive-training-based interventions too have failed to induce generalized effects (e.g., Sisk, Burgoyne, Sun, Butler, & MacNamara, 2018). Put together, the convergent insights from different fields of scientific research about far transfer represent a successful example of triangulation (Munafò & Smith, 2018), and lead us towards the conclusion that while human cognition is malleable to training, the benefits are domain-specific.

The implications are profound. From the theoretical point of view, those theories of human cognition predicting minimal or no far transfer of skills are corroborated by our findings (e.g., chunking-based theories; for a review, see Gobet, 2016). Conversely, those theories predicting the generalization of skills acquired by training across multiple domains are refuted (e.g., Bavelier, Green, Pouget, & Schrater, 2012; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Tierney, Krizman, & Kraus, 2015). Regarding practical implications, the obvious conclusion is that professional and educational curricula should focus on domain-specific knowledge rather than general and allegedly transferable skills.

Limitations

As noted in the General Method section, none of the first-order meta-analyses included effects that were corrected for measurement error. However, we think that the practical consequences of this flaw are negligible, especially in Models 2 and 3. In fact, pretty much all the uncorrected and corrected far-transfer overall effects were close or equal to zero. Thus, applying such a correction would leave all these estimates virtually unaltered.

A second limitation concerns the choice of the most appropriate corrected estimate to be included in the second-order meta-analytic models. As seen, our criterion is based on the PET: if the PET-corrected estimate is not significantly ($p < .100$) positive, we select that estimate that is the closest to zero. In our opinion, this criterion is sensible and reflects the rationale of PET (Stanley, 2017). Furthermore, the fact that none of the PET-corrected estimates in Models 2 and 3 are statistically different from zero represents substantial evidence in favor of our hypothesis (i.e., no far transfer regardless of population and training regimen). That being said, preferring a corrected estimate over another one always implies a certain degree of arbitrariness that is impossible to rule out completely. In any case, given the high degree of consistency observed across the findings, we expect the overall results to be the same regardless of the publication-bias corrected estimate employed.

Third, the selection of some of the first-order meta-analyses is, to a certain degree, arbitrary too. Specifically, along with Melby-Lervåg et al. (2016) and Sala and Gobet (2017b), several other recent meta-analyses examining the effects of WM training have been carried out (e.g., Au et al., 2015; Soveri, Antfolk, Karlsson, Salo, & Laine, 2017). These meta-analyses are substantially in line with our conclusions. Soveri et al. (2017) report no far-transfer effects in healthy adults. Au et al. (2015) claim that *n*-back training has a small overall positive impact ($\bar{g}$ = 0.24) on fluid intelligence. However, this effect disappears in studies using active control groups (Dougherty, Hamovits, & Tidwell, 2016). A similar consideration applies to action video-game training. Along with Sala et al. (2018), another meta-analysis examining the effects of action video-game training has been carried out recently (Bediou et al., 2018). While reporting a small to medium overall effect size ($\bar{g}$ = 0.34), this meta-analysis shows a highly asymmetrical distribution of the effect sizes which suggests that the uncorrected effect is an overestimation. This conclusion is upheld by the results of the two publication bias analyses included in the original article. Thus, we think that our findings are robust regardless of the particular meta-analytic study selected for our second-order meta-analyses.

Finally, another issue regards the total number of primary studies included in some of the first-order meta-analyses (e.g., chess and exergames). A small number of effect sizes provides less accurate estimates (large standard errors) and limits the power of publication bias analysis. This problem cannot be overcome until new experiments have been carried out. Nonetheless, the results provided by the omnibus meta-analyses, which do not suffer from low statistical power, confirm the results of the first-order and second-order meta-analyses.

Conclusions and Recommendation for Future Research

This study aimed to test to what extent cognitive-training programs trigger near-transfer and far-transfer effects. Near transfer occurs in all the examined populations but, interestingly, young populations seem to benefit from the treatment (i.e., WM training) more than adult populations. It is also worth remembering that the observed near transfer probably reflects an improvement in the ability to perform memory tasks rather than enhanced cognitive function. On the other hand, there is no evidence of far transfer regardless of the population and type of cognitive-training program. The findings are consistent with substantial research into education, skill acquisition, and expert performance. Based on these results, we think that the lack of far transfer is an invariant in human cognition, at least with regard to the general population.

That being said, further research is needed to extend the present second-order meta-analytic investigation in order to test our hypothesis. First, the investigation should include a meta-analysis of brain-training programs. In our opinion, no eligible first-order meta-analysis on the topic has been published so far. For example, meta-analytic investigations about brain-training programs usually include other cognitive-training programs such as WM training or action video-game training (e.g., Mewborn, Lindbergh, & Miller, 2017). Thus, adding such a meta-analysis would violate the assumption of statistical independence between studies. It is worth noting that, in line with our conclusions, Simons et al. (2016) have shown that no convincing evidence of far transfer has been provided by brain-training experimental studies so far (see also Rebok et al., 2014). Nonetheless, a meta-analysis examining near- and far-transfer effects of brain-training programs would provide valuable additional information to test this claim further. Second, novel cognitive-training programs have been designed in recent years (e.g., Daugherty et al., 2018), and new studies analyzing the effect of old cognitive-training programs on different populations are being currently (or have just been) carried out (e.g., action video-game training in older adults; Ballesteros et al., 2018). Once the number of experimental studies is sufficient to run additional first-order meta-analyses, it will be possible to carry out more extensive second-order meta-analytic models. The same applies to those studies about the effects of cognitive-training programs on those populations that are included in our analyses. New primary studies will contribute to the updating of the first- and second-order meta-analyses.

References

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: A meta-analysis. *Psychonomic Bulletin & Review, 22*, 366-377. doi:10.3758/s13423-014-0699-x

Ballesteros, S., Mayas, J., Prieto, A., Ruiz-Marquez, E., Toril, P., & Reales, J. M. (2017). Effects of video game training on measures of selective attention and working memory in older adults: Results from a randomized controlled trial. *Frontiers in Aging Neuroscience, 9*:354. doi:10.3389/fnagi.2017.00354

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*, 612-637. doi:10.1037//0033-2909.128.4.612

Bavelier, D., Green, C. S., Pouget, A., Schrater, P. (2012). Brain plasticity through the life span: learning to learn and action video games. *Annual Review of Neuroscience, 35*, 391-416. doi:10.1146/annurev-neuro-060909-152832

Bediou, B., Adams, D. M., Mayer, R. E., Tipton, E., Green, C. S., & Bavelier, D. (2018). Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychological Bulletin, 144*, 77-110. doi:10.1037/bul0000130

Bilalić, M., McLeod, P., & Gobet, F. (2009). Specialization effect and its influence on memory and problem solving in expert chess players. *Cognitive Science, 33*, 1117-1143. doi:10.1111/j.1551-6709.2009.01030.x

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science, 8*, 445-454. doi:10.1177/1745691613491271

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.

Burgoyne, A. P., Sala, G., Gobet, F., Macnamara, B. N., Campitelli, G., & Hambrick, D. Z. (2016). The relationship between cognitive ability and chess skill: A comprehensive meta-analysis. *Intelligence, 59*, 72-83. doi:10.1016/j.intell.2016.08.002

Carter, E., Schönbrodt, F., Gervais, W. M., & Hilgard, J. (2017). Correcting for bias in psychology: A comparison of meta-analytic methods. doi:10.17605/OSF.IO/9H3NU. Preprint available at https://psyarxiv.com/9h3nu/

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 1-58). New York: Academic.

Cheung, S. F., & Chan, D. K. (2004). Dependent effect sizes in meta-analysis: Incorporating the degree of interdependence. *Journal of Applied Psychology, 89*, 780-791. doi:10.1037/0021-9010.89.5.780

Cheung, S. F., & Chan, D. K. (2008). Dependent correlations in meta-analysis: The case of heterogeneous dependence. *Educational and Psychological Measurement, 68*, 760-777. doi:10.1177/0013164408315263

Cheung, S. F., & Chan, D. K. (2014). Meta-analyzing dependent correlations: An SPSS macro and an R script. *Behavioral Research Methods, 46*, 331-345. doi:10.3758/s13428-013-0386-2

Daugherty, A. M., Zwilling, C., Paul, E. J., Sherepa, N., Allen, C., Kramer, A. F., … Barbey, A. K. (2018). Multi-modal fitness and cognitive training to enhance fluid intelligence. *Intelligence, 66*, 32-44. doi:10.1016/j.intell.2017.11.001

Detterman, D. K. (1993). The case for the prosecution: Transfer as an epiphenomenon. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 1-24). Westport, CT, US: Ablex Publishing.

Detterman, D. K. (2014). Introduction to the intelligence special issue on the development of expertise: is ability necessary? *Intelligence, 45*, 1-5. doi:10.1016/j.intell.2014.02.004

Detterman, D. K. (2016). Education and intelligence: Pity the poor teacher because student characteristics are more significant than teachers or schools. *The Spanish Journal of Psychology, 19*, E93. doi:10.1017/sjp.2016.88

Donovan, M. S., Bransford, J. D., & Pellegrino, J. W. (1999). *How people learn: Bridging research and practice*. Washington, DC: National Academies Press.

Dougherty, M. R., Hamovits, T, & Tidwell, J. W. (2016). Revaluating the effect of n-back training on transfer through the Bayesian lens: Support for the null. *Psychonomic Bulletin & Review, 23*, 306-316. doi:10.3758/s13423-015-0865-9

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*, 276-284. doi:10.1111/j.0006-341X.2000.00455.x

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49*, 725-747.

Finn, A. S., Kraft, M. A., West, M. R., Leonard, J. A., Bish, C. E., Martin, R. E., … Gabrieli, J. D. E. (2014). Cognitive skills, student achievements test, and schools. *Psychological Science, 25*, 736-744. doi:10.1177/0956797613516008

Gobet, F. (2016). *Understanding expertise: A multi-disciplinary approach*. London: Palgrave/Macmillan.

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31*, 1-40. doi:10.1006/cogp.1996.0011

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*, 39-65. doi:10.1002/jrsm.5

Henmi, M., & Copas, J. B. (2010). Confidence intervals for random effects meta-analysis and robustness to publication bias. *Statistics in Medicine, 29*, 2969-2983. doi:10.1002/sim.4029

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analysis. *British Medical Journal, 327*, 557-560. doi:10.1136/bmj.327.7414.557

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings (2nd ed.)*. Newbury Park, CA: Sage.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 6829-6833. doi:10.1073/pnas.0801268105

Jerrim, J., Macmillan, L., Micklewright, J., Sawtell, M., and Wiggins, M. (2016). Chess in Schools. Evaluation Report and Executive Summary. Education Endowment Foundation. Available online at: https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/EEF_Project_Report_Chess_in_Schools.pdf

Karbach, J., & Schubert, T. (2013). Training-induced cognitive and neural plasticity. *Frontiers in Human Neuroscience, 9*:6. doi:10.3389/fnhum.2013.00048

Kepes, S., & McDaniel, M. A. (2015). The validity of conscientiousness is overestimated in the prediction of job performance. *PLoS ONE, 10*:e0141468. doi:10.1371/journal.pone.0141468

Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level metaanalysis. *Research Synthesis Methods, 2*, 61-76. doi:10.1002/jrsm.35

McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis: An evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science, 11*, 730-749. doi:10.1177/1745691616662243

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of far-transfer: Evidence from a meta-analytic review. *Perspective on Psychological Science, 11*, 512-534. doi:10.1177/1745691616635612

Mewborn, C. M., Lindbergh, C. A., & Miller, L. S. (2017). Cognitive interventions for cognitively healthy, mildly impaired, and mixed samples of older adults: A systematic review and meta-analysis of randomized-controlled trials. *Neuropsychology Review, 27*, 403-439. doi:10.1007/s11065-017-9350-8

Moreno, S. G., Sutton, A. J., Ades, A. E., Stanley, T. D., Abrams, K. R., Peters, J. L., & Cooper, N. J. (2009). Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC medical research methodology, 9*:2. doi:10.1186/1471-2288-9-2

Mosing, M. A., Madison, G., Pedersen, N. L., & Ullén, F. (2016). Investigating cognitive transfer within the framework of music practice: Genetic pleiotropy rather than causality. *Developmental Science, 19*, 504-512. doi:10.1111/desc.12306

Munafò, M. R., & Smith, G. D. (2018). Replication is not enough. *Nature, 553*, 399-401.

Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H.-Y., King, J. W., … Willis, L. S. (2014). Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults. *Journal of the American Geriatrics Society, 62*, 16-24. doi:10.1111/jgs.12607

Rikers, R. M. J. P., Schmidt, H. G., & Boshuizen, H. P. A. (2002). On the constraints of encapsulated knowledge: Clinical case representations by medical experts and subexperts. *Cognition and Instruction, 20*, 27-45. doi:10.1207/S1532690XCI2001_2

Ritchie, S. J., Bates, T. C., & Deary, I. J. (2015). Is education associated with improvements in general cognitive ability, or in specific skills? *Developmental Psychology, 51*, 573-582. doi:10.1037/a0038981

Sala, G., & Gobet, F. (2016). Do the benefits of chess instruction transfer to academic and cognitive skills? A meta-analysis. *Educational Research Review, 18*, 46-57. doi:10.1016/j.edurev.2016.02.002

Sala, G., & Gobet, F. (2017a). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Direction in Psychological Science, 26*, 515-520. doi:10.1177/0963721417712760

Sala, G., & Gobet, F. (2017b). Working memory training in typically developing children: A meta-analysis of the available evidence. *Developmental Psychology, 53*, 671-685. doi:10.1037/dev0000265

Sala, G., & Gobet, F. (2017c). When the music's over. Does music skill transfer to children's and young adolescents' cognitive and academic skills? A meta-analysis. *Educational Research Review, 20*, 55-67. doi:10.1016/j.edurev.2016.11.005

Sala, G., Tatlidil, K. S., & Gobet, F. (2018). Video game training does not enhance cognitive ability: A comprehensive meta-analytic investigation. *Psychological Bulletin, 144*, 111-139. doi:10.1037/bul000013

Schmidt, F. L. (2017). Beyond questionable research methods: The role of omitted relevant research in the credibility of research. *Archives of Scientific Psychology, 5*, 32-41. doi:10.1037/arc0000033

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

Schmidt, F. L., & Oh, I.-S. (2013). Methods for second order meta-analysis and illustrative applications. *Organizational Behavior and Human Decision Processes, 121*, 204-218. doi:10.1016/j.obhdp.2013.03.002

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin, 138*, 628-654. doi:10.1037/a0027473

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J.L., & MacNamara, B. N. (2018). To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analyses. *Psychological Science, 29*, 549-571. doi:10.1177/0956797617739704

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S.E., Chabris, C. F., Hambrick, D. Z., &

Stine-Morrow, E. A. L. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest, 17*, 103-186. doi:10.1177/1529100616661983

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). p-Curve and effect size correcting for publication bias using only significant results. *Perspectives on Psychological Science, 9*, 666-681. doi:10.1177/1745691614553988

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin & Review, 24*, 1077-1096. doi:10.3758/s13423-016-1217-0

Stanley, T. D. (2017). Limitations of PET-PEESE and other meta-analysis methods. *Social Psychological and Personality Science, 8*, 581-591. doi:10.1177/1948550617693062

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods, 5*, 60-78. doi:10.1002/jrsm.1095

Stanley, T. D., & Doucouliagos, H. (2017). Neither fixed nor random: Weighted least squares meta-regression. *Research Synthesis Methods, 8*, 19-42. doi:10.1002/jrsm.1211

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. *Neuroscience and Biobehavioral Reviews, 78*, 34-43. doi:10.1016/j.neubiorev.2017.04.011

Strobach, T., & Karbach, J. (2016). *Cognitive training: An overview of features and applications*. New York: Springer.

Taatgen, N. A. (2013). The nature and transfer of cognitive skills. *Psychological Review, 120*, 439-471. doi:10.1037/a0033138

Taatgen, N. A. (2016). Theoretical models of training and transfer effects. In T. Strobach, & J. Karbach (Eds.), *Cognitive training: An overview of features and applications* (pp. 19-29).

New York: Springer.

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods, 5*, 13-30. doi:10.1002/jrsm.1091

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. (I). *Psychological Review, 8*, 247-261. doi:10.1037/h0074898

Tierney, A. T., Krizman, J., & Kraus, N. (2015). Music training alters the course of adolescent auditory development. *Proceedings of the National Academy of Sciences of the United States of America, 112*, 10062-10067. doi:10.1073/pnas.1505114112

Tracz, S. M., Elmore, P. B., & Pohlmann, J. T. (1992). Correlational meta-analysis: Independent and nonindependent cases. *Educational and Psychological Measurement*, *52*, 879-888. doi:10.1177/0013164492052004007

Vevea J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods, 10*, 428-443. doi:10.1037/1082-989X.10.4.428

Viechtbauer, W. (2010). Conducting meta-analysis in R with the metafor package. *Journal of Statistical Software, 36*, 1-48. Retrieved from http://brieger.esalq.usp.br/CRAN/web/packages/metafor/vignettes/metafor.pdf

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods, 1*, 112-125. doi:10.1002/jrsm.11