UC Davis
UC Davis Previously Published Works

Title

Why Does Cognitive Training Yield Inconsistent Benefits? A Meta-Analysis of Individual Differences in Baseline Cognitive Abilities and Training Outcomes

Permalink

https://escholarship.org/uc/item/6bm9x5sk

Authors

Traut, Hilary J
Guild, Ryan M
Munakata, Yuko

Publication Date

2021

DOI

10.3389/fpsyg.2021.662139

Copyright Information

Peer reviewed

Why Does Cognitive Training Yield Inconsistent Benefits? A Meta-Analysis of Individual Differences in Baseline Cognitive Abilities and Training Outcomes

Hilary J. Traut[1]*, Ryan M. Guild[1] and Yuko Munakata[2]

[1] Cognitive Development Center, Department of Psychology & Neuroscience, University of Colorado Boulder, Boulder, CO, United States, [2] Cognition in Context Lab, Department of Psychology and Center for Mind & Brain, University of California, Davis, Davis, CA, United States

Despite growing interest in improving cognitive abilities across the lifespan through training, the benefits of cognitive training are inconsistent. One powerful contributor may be that individuals arrive at interventions with different baseline levels of the cognitive skill being trained. Some evidence suggests poor performers benefit the most from cognitive training, showing compensation for their weak abilities, while other evidence suggests that high performers benefit most, experiencing a magnification of their abilities. Whether training leads to compensation or magnification effects may depend upon the specific cognitive domain being trained (such as executive function or episodic memory) and the training approach implemented (strategy or process). To clarify the association between individual differences in baseline cognitive ability and training gains as well as potential moderators, we conducted a systematic meta-analysis of the correlation between these two variables. We found evidence of a significant meta-correlation demonstrating a compensatory effect, a negative association between initial ability on a trained cognitive process and training gains. Too few papers met our search criteria across the levels of proposed moderators of cognitive domain and training approach to conduct a reliable investigation of their influence over the meta-analytic effect size. We discuss the implications of a compensatory meta-correlation, potential reasons for the paucity of qualifying papers, and important future directions for better understanding how cognitive trainings work and for whom.

Keywords: executive function, episodic memory, compensation effect, magnification effect, cognitive control

1. INTRODUCTION

There has been steady and growing interest in developing effective methods to manipulate an individual's cognitive abilities. Work on this topic has exploded in recent decades, motivated by rehabilitative and therapeutic goals for impaired or aging populations (Lustig et al., 2009; Dahlin, 2011; van der Donk et al., 2017), by interest in cognitive enhancement for children or the general public (Diamond and Lee, 2011; Diamond, 2012; Simons et al., 2016), as well as by a desire to find means of testing cognitive theories (Cepeda et al., 2001; Chevalier et al., 2014). However, results of cognitive training research have not always been consistent.

While some research finds clear benefits to a trained ability, and even benefits to untrained abilities (both near and far related), other studies yield little to no evidence of benefit from cognitive training. This variation among findings is well-documented and presents serious problems for determining the true effectiveness or utility of cognitive training in both clinical and research contexts (Shipstead et al., 2012; Melby-Lervåg and Hulme, 2013; von Bastian and Oberauer, 2014; Au et al., 2015; Melby-Lervåg et al., 2016; Simons et al., 2016; Rossignoli-Palomeque, 2018). While the design and methodological rigor with which cognitive training research is conducted is an established source of this variation (Noack et al., 2014; Tidwell et al., 2014; Simons et al., 2016; Smoleń et al., 2018) variability remains even when accounting for these concerns. Sources of this remaining variability are of great interest as they could both provide insight into the reasons for the effectiveness or ineffectiveness of cognitive training—illuminating the mechanisms of how trainings work and for whom—and allow for administration of more tailored cognitive trainings in clinical or therapeutic contexts.

Recent evidence suggests that individual differences account for a sizeable amount of the lingering variability in findings over and above methodological differences (Jaeggi et al., 2014; von Bastian and Oberauer, 2014; Borella et al., 2017; Guye et al., 2017). Participants' ages, socio-economic statuses (SES), levels of motivation, and more have all been implicated in influencing the effects of various cognitive trainings (Kliegl et al., 1990; Katz et al., 2014; Segretin et al., 2014, respectively). In particular individual differences in baseline cognitive ability (i.e., an individual's pre-training level of cognitive skill) have been implicated as a prominent source of variation (e.g., Rueda et al., 2005; Jaeggi et al., 2008; Karbach and Unger, 2014; Foster et al., 2017) —suggesting that one way to understand who benefits from what cognitive training is to understand individual abilities upon entering training.

It makes sense that each individual comes into cognitive training with unique abilities, and that baseline could influence effects of training. Multiple theories hypothesize that learning opportunities are most effective when they are tailored to an individual's ability: in child-caregiver interactions (Zone of Proximal Development - Vygotsky, 1978), educational settings (Desirable Difficulty - Bjork and Bjork, 2011), and cognitive interventions (Aptitude by Treatment Interactions - Snow, 1989). Moreover, the influence of other individual difference effects on cognitive training, such as SES and age, appear to be mediated by participants' cognitive baselines (Bürki et al., 2014; Segretin et al., 2014, respectively), suggesting that their influence is actually due in part to baseline cognitive ability. For example, results on individual differences effects of SES suggest that SES does not directly cause differences in cognitive training outcomes, but instead results in a situation that produces different cognitive baseline proficiencies across participants. Given this, one potentially impactful approach to developing a mechanistic understanding of how individual differences give rise to variability in cognitive training outcomes is to focus specifically on individual differences in baseline cognitive ability.

However, we do not have a clear picture of the size, or even the direction, of baseline individual differences effects on cognitive training (see Lövdén et al., 2012; Karbach and Unger, 2014; Au et al., 2015, for review). This makes it difficult to get a handle on the importance and informativeness of this phenomenon. Some training studies yield compensation effects: a negative relationship between a participant's initial cognitive ability and the results of training such that worse performing participants demonstrate the greatest benefits from training. In these cases, training appears to allow poor performers to compensate for their weaknesses, such that overall individual differences among participants at pre-training are reduced post-training (**Figure 1A**). Other training studies yield magnification effects: a positive relationship between participants initial cognitive ability and the results from training. In these scenarios, individuals with the highest levels of performance prior to training experience the most substantial benefits—with individual differences among the training group becoming magnified post-training (**Figure 1B**).

What might lead to these opposing patterns? Compensation effects may result from a need to master certain skills in order to achieve high levels of performance on a task. These skills may be acquired during training by low performers, bolstering their scores, whereas high performers have already mastered them and receive little benefit (e.g., Lustig et al., 2009). In contrast, magnification effects may result from a certain level of skill being necessary to benefit from a cognitive training intervention. Those who have not reached that threshold do not benefit, while their peers who have reached threshold improve. However, neither of these accounts, laid out in their simplest forms, explain why the literature demonstrates both compensation and magnification effect patterns as opposed to just one or the other.

One clue is that these opposing patterns appear to come from qualitatively different types of cognitive training (see Lövdén et al., 2012; Karbach and Unger, 2014, for review). Most magnification effects come from cognitive training of episodic memory, the ability to retrieve specific information in memory such as the definition of a word or location of a favorite cafe (Kliegl et al., 1990; Verhaeghen et al., 1992; Cox, 1994; Brehmer et al., 2007) (cf. Gaultney et al., 1996; Lövdén et al., 2012). In contrast, most compensation effects come from cognitive training of executive function, abilities related to the control and regulation of behavior (Diamond and Lee, 2011; Zinke et al., 2012, 2014; Karbach et al., 2017) (cf. Loosli et al., 2012; Foster et al., 2017; Wiemers et al., 2019). This raises the possibility that episodic memory and executive function require differing initial levels of proficiency to change with training or that change in these domains develops along different growth functions. For instance, change in episodic memory may necessitate individuals to have already achieved a certain level of proficiency to improve from cognitive interventions. Executive function, on the other hand, may develop in such a way that prior to achieving proficiency there is extensive room for improvement via training—but beyond that point, changes in executive function performance may be more nuanced and difficult to achieve via current training approaches or just difficult to detect in behavioral tasks.
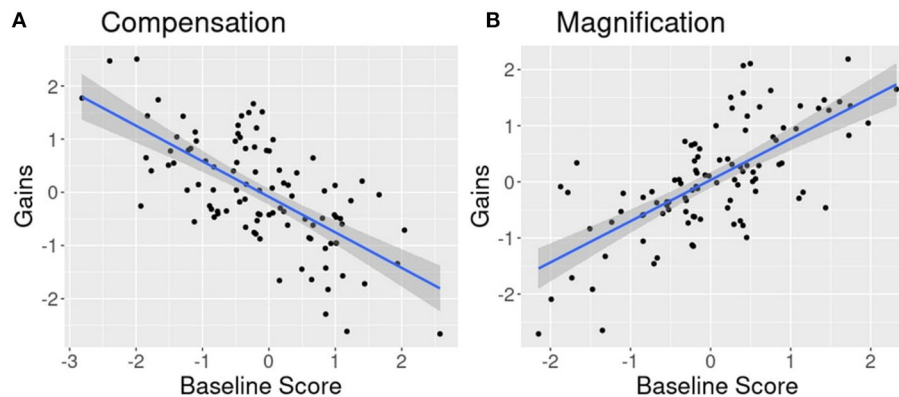
FIGURE 1 | Example of opposing patterns of the relationship between baseline cognitive ability and gains after cognitive training. **(A)** Compensation effect demonstrating lower performing individuals accruing the most gains after training. **(B)** Magnification effect demonstrating higher performing individuals accruing the most gains after training.

Executive function and episodic memory trainings also differ in the type of training approach utilized, raising the question of whether training approach, and not training domain, determines whether the worst performers or the best performers benefit the most (Lövdén et al., 2012; Karbach and Unger, 2014). Cognitive training of episodic memory primarily utilizes a strategy-based approach, during which participants are given explicit strategy instruction (such as Method of Loci) intended to improve their ability (e.g., Kliegl et al., 1990). Cognitive training of executive function tends to utilize a primarily process-based approach, during which participants practice on one or several related executive function tasks—presumably bolstering their ability implicitly through practice (e.g., Jaeggi et al., 2008). Thus, episodic memory trainings may give rise to magnification effects because explicit strategy instruction could require mastery of an initial skill set. Executive function trainings may give rise to compensation effects because high performers have already acquired the skill set provided through process-based training, leaving them little room to improve compared to less skilled participants still mastering that skill set.

It does not appear that any studies have directly tested the effects of training domain (episodic memory vs. executive function), or tested for an interaction of training domain with training approach (process- vs. strategy-based). To our knowledge, only one study has explicitly investigated questions of training approach (strategy- or process-focused) in baseline individual differences effects on cognitive training, conducting a detailed investigation of change across an episodic memory training by sequentially administering strategy followed by process training manipulations. Analyses showed an initial reduction in individual differences post-strategy instruction (a compensation effect), but magnification of those differences after further process based practice (Lövdén et al., 2012). While these results highlight potential differences in training approach within the episodic memory domain, they are inconsistent with the aforementioned pattern of findings in which strategy training leads to magnification and process training leads to

compensation. The design of the study also makes it difficult to rule out the possibility that initial administration of strategy training altered the latter administration of process training via carryover effects. These results highlight that while many studies have tested the association of baseline individual differences with training outcomes, few papers have investigated moderators of this association, leaving many questions unanswered.

Synthesis of investigations into the influence of baseline individual differences is complicated both by conflicting findings and by notable definitional and analysis differences between studies. Studies frequently differ in how they define and operationalize baseline and outcome measures. Precisely what qualifies as "baseline cognitive performance" or "training outcomes" are essential points that change the question ultimately being asked by an analysis. Baseline cognitive ability, for example, is used interchangeably to refer to cognitive skills with varying distance and relation to the exact domain being trained broad and generalized. It could refer to broad cognitive domains operationalized via measures of general cognitive ability or intelligence (e.g., Rueda et al., 2005; Jaeggi et al., 2008; Fu et al., 2017; Hering et al., 2017), but it could also refer to a narrow cognitive domain specific to the to-be-trained cognitive domain, such as a measure of baseline performance on a working memory task administered prior to a working memory training (e.g., Zinke et al., 2012, 2014). Both approaches provide essential information about evaluating who may benefit the most and why, but address conceptually different questions. The former probes how an overall cognitive level may impact training attempts, testing whether a stronger general level of cognitive skill aids training. But, the latter approach asks how ability within the specific cognitive domain or skill targeted by training influences the effects of training.

Near transfer tasks comprise those within the same domain as training but different from the precise task trained (e.g., two different measures of working memory). Far transfer tasks comprise those which measure performance on an ability other than that being trained (e.g., general fluid intelligence

performance after working memory training). While some papers address the degree of transfer with regard to training outcomes and individual differences (e.g., Jaeggi et al., 2014), some analyses or discussions do not distinguish between the specific types of training outcomes. This means that some analyses and discussions of baseline individual differences are based on a distant (and potentially tenuous) relationship between cognitive training and far transfer, while others are based on a close relationship between training and final performance on that training. Studies probing this dimension have suggested compensatory effects on trained tasks or near transfer tasks and magnification effects on far transfer tasks (e.g., Borella et al., 2017), (cf. Karbach et al., 2017), emphasizing the importance of being clear about this distinction.

The actual score of training outcomes also takes on many forms—being analyzed as raw training performance during the cognitive training task, scaled training performance (e.g., rank-ordered or z-scored), assignment to low/high performance groups (e.g., based on a median split of performance or another procedure), or calculated gain scores (e.g., change in performance from baseline to post-test). Each of these approaches has unique drawbacks. The use of only raw or scaled training performance to probe the influence of baseline individual differences may merely capture stable individual differences in general ability among participants, as opposed to differences in training outcome specifically. The use of low/high performance groups reduces the power of analyses (Iacobucci et al., 2015). The use of gain scores can lead to uncovering a compensatory association caused by regression to the mean as well as other possible statistical pitfalls (Smoleń et al., 2018).

Analyzing all of these measurement approaches, as opposed to looking at a single means of capturing outcomes, means contending with all of their associated issues. For example, two papers showing conflicting compensation (Karbach et al., 2017) and magnification (Foster et al., 2017) findings on process-based executive function training were derived from studies using very different methods for interrogating individual differences questions. These studies differed both in the characteristics of baseline data analyzed and analysis approach. One paper utilized a continuous measure of each participant's individual performance on baseline and a latent modeling approach (Karbach et al., 2017)—whereas the other changed baseline into a categorical measure, grouping participants into high and low baseline performance groups for analysis combined with ANCOVA procedures (Foster et al., 2017). These differences in capturing baseline performance and analyzing it makes it difficult to compare the two before even considering methodological differences.

Such differences across studies make it difficult to make sense of who benefits most from which kinds of training. Evaluation of any version of questions about individual differences in cognitive baseline and training outcomes is clouded by the likely differing relationships between various combinations of "cognitive baseline" and "training outcomes." To best leverage individual differences approaches for making theoretical and mechanistic claims about cognitive trainings, we must be

clear about what exactly is being analyzed when we look at individual differences.

The current review seeks to provide a systematic synthesis of the literature on individual differences in cognitive baseline and cognitive training outcomes by conducting a planned meta-analysis. A meta-analytic approach to this problem provides several advantages. First, it brings a systematic method to evaluating currently available findings by requiring clear definitions and eligibility criteria to be used in evaluating findings for inclusion in analysis. Second, it leverages the extensive existing work on cognitive training to make inferences. Third, it allows for a powerful investigation of the known heterogeneity of baseline individual differences influences on cognitive training outcomes—turning the methodological diversity of experimental work into an informative tool as opposed to an obstacle to be overcome. Differences in study approaches allow for an investigation into the reliability of this phenomenon as well as possible moderators: cognitive domain trained, training approach, and their interaction. It allows us to test whether accounting for these differences brings clarity to the nature of baseline individual differences association with training outcomes. Lastly, even in the event of null results, a meta-analysis clarifies the existing literature and can produce a more robust synthesis of that literature than non-systematic reviews. This systematic synthesis can then support a more effective design of causal studies to investigate questions surrounding moderators of the association between cognitive baseline and training outcomes.

However, not all of the variability in the literature is helpful, such as the aforementioned variability in definitions and statistical approach. Differences in the variables used in analyses mean that not every study is testing the same type of baseline/outcome association, such that including all of these approaches above in a meta-analysis could obscure relationships. The present review thus focuses the problem space on studies that are comparable in their definitions of cognitive baseline and training outcome, as well as studies that operationalize training outcomes similarly. This was done to limit the introduction of variability unrelated to the theoretical question into the meta-analysis.

First, this review focuses on baseline ability defined as the pre-training performance in the target cognitive domain manipulated by training (e.g., baseline working memory abilities of participants in a working memory training). It therefore excludes studies that define this measure by some general cognitive ability such as fluid intelligence or studies that evaluate baseline individual differences in a cognitive domain different from that being trained. Second, this review constrains outcomes to measures derived from training or near transfer (provided the near transfer task is of the same cognitive domain as training), excluding far transfer outcomes. Far transfer measures could take on an array of forms and vary in how different they are from the trained ability. Currently there is no clear picture of what the relationship between gains in training, near transfer, and far transfer is (Thompson et al., 2013; Au et al., 2015; Sala and Gobet, 2017)—making it unwise to combine all three in the present meta-analysis. Lastly, this review focuses on studies using

a consistent form of operationalizing training outcomes as gain scores, as opposed to raw or scaled performance, because this was the most readily available scoring procedure across the literature.

As such, the purpose of this meta-analysis is to (a) examine and synthesize evidence for the association of individual differences with baseline cognitive ability on cognitive training outcomes through correlation coefficients of this relationship and (b) investigate how this correlation is moderated. Specifically, these analyses are designed to answer the following research questions:

1. What is the estimate of the average correlation coefficient of the association between baseline abilities and cognitive training outcomes?
2. Is this correlation moderated by the targeted cognitive domain (i.e., episodic memory or executive function)?
3. Is this correlation moderated by the cognitive training approach (i.e., strategy or practice)?
4. Is this correlation moderated by an interaction between domain and training approach?

2. METHODS

This review was conducted in accordance with recommendations outlined by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) group guidelines (Liberati et al., 2009).

2.1. Information Sources & Search Protocol

Based on several practical considerations, all literature searches were conducted through Google Scholar. First, a database was sought that allowed for full text searches (i.e., title, abstract, and body of article) due to (a) the diversity of terms utilized in titles, abstracts, and keywords to describe research on cognitive training, and (b) the frequency with which individual differences analyses are discussed only in the body of an article, given that many of these analyses are exploratory or secondary to the primary question of the article. At the time this study was conducted, this restriction limited potential databases for systematic search to Google Scholar or PsychINFO. Google Scholar was selected because it is both an openly accessible and un-curated database. Unlike other "curated" academic databases such as PsychINFO, PubMed, Web of Science, and others—Google Scholar procures search results through an entirely automated algorithmic search of the internet-at-large, as opposed to searching a limited, pre-selected body of entries. While the downside of this approach is that the nature of this algorithm is unknown, this still arguably makes Google Scholar a more unbiased method of searching the literature. Google Scholar is neither beholden to specific database and publisher relationships nor are its searches influenced by curated database add-ons such as enhanced keywords. It also provides the substantial advantage of crawling a diversity of sources that include dissertations, conference proceedings, posters, unpublished work, and similar "gray" literature because of its broad un-curated search approach. Therefore, the use of Google Scholar helps to circumvent potential issues resulting from both

a lack of common terminology in titles, abstracts, and keywords, as well as issues resulting from bias in search results attributed to the curated and financial nature of curated database services.

As such, all searches were conducted in Google Scholar utilizing the following protocol:

"cognitive training" AND "individual differences"

While the use of such a broad protocol in a full-text search was expected to produce a large initial pool of studies, it was selected to ensure the capture of the majority of studies likely to contain relevant data in order to provide as unbiased a review of the available data as possible. The search protocol was constrained to the years 1980 to 2018 (inclusive) to best capture the modern literature on cognitive training. Searches were conducted within specific year intervals (e.g., 1999 to 1999) to address Google Scholar's 1,000 results limit. This allowed for access to the complete set of articles determined by Google Scholar's search. For individual years with results exceeding 1,000 articles, the search protocol was separated into two parts by searching for articles based on the required inclusion or exclusion of a common word, "baseline":

"cognitive training" AND "individual differences"
AND "baseline"
"cognitive training" AND "individual differences"
NOT "baseline"

Returned search results were extracted from Google Scholar results pages using an automated script implemented in Python (see **Supplementary Materials**). This script extracted title, author, journal, digital access (e.g., website), and related publication information. Complete abstracts for each search result were extracted manually by trained research assistants and saved in a complete document of all search results and retrieved information as Google Scholar search results do not return complete abstracts. All information was recorded in a single tracking document.

2.2. Eligibility Criteria

Eligibility was determined by the first author of this paper (H.J. Traut) over the course of unblinded reviews of retrieved search results. Eligibility criteria were designed along the "PICOS" recommendations of the PRISMA guidelines (Liberati et al., 2009) to account for Participants, Intervention, Comparators, Outcome, and Study design. Additional categories of criteria were included as necessitated by the demands of the present systematic review.

2.2.1. Participants

Studies were not excluded based on the nature of participants provided that participants were human (e.g., excluding cognitive training studies conducted with rodents or other common animal subjects). Participants of any age, neurological status, or other characteristics were eligible for inclusion. This decision was made to allow for the maximum number of results to be retrieved and due to the absence of hypotheses explicitly regarding differences in baseline individual difference effects between typical and atypical individuals.

2.2.2. Intervention (i.e., training)

To be eligible, studies were required to include experiment(s) in which at least one group of participants underwent training with a task intended to manipulate the participants' executive function (EF) or episodic memory (EP) abilities according to the following definitions:

Episodic Memory Training: Training intended to manipulate an individual's ability to retrieve explicit information such as facts, definitions, or experiences (Tulving, 1972).

Executive Function Training: Training intended to manipulate an individual's ability to control and regulate their behavior or the components of that ability, such as inhibition, updating, and shifting (Miyake et al., 2000; Miyake and Friedman, 2012).

Examples of episodic memory training studies would include mnemonic trainings such as Method of Loci (e.g., Cox, 1994). Examples of executive function training studies would include many working memory or inhibition trainings (e.g., von Bastian and Oberauer, 2014). Studies including training on both executive function and episodic memory were excluded as they would not allow for differentiation of the theorized interaction effect between training domain and the influence of baseline individual differences on outcomes.

Trainings were further required to consist of a task or procedure that directly involved a cognitive skill and was either a strategy or process-based training following the definitions provided by Lustig et al. (2009):

Strategy Training: The explicit instruction of a strategy with the intent of helping participants improve their performance on the target domain.

Process Training: The administration of either a single task or set of tasks directly derived from or related to the target domain for participants to practice with.

Studies including trainings that integrated both strategy and process elements were excluded because the inclusion of both approaches would make it impossible to dissociate their effects in the proposed moderator analysis. Similarly, studies administering process or strategy-based trainings alongside additional manipulations were also excluded (i.e., multi-modal interventions). This included, but was not limited to, trainings with manipulations such as cardio-vascular exercise, mindfulness, commercial/popular video-games (e.g., World of Warcraft), social skills training, and emotional skills training. It also included training interventions that involved administration of neurostimulation, neurofeedback, pharmacological interventions, therapeutic interventions (such as cognitive behavioral therapy), or other concurrent additional manipulations to cognitive training. These exclusions were included to ensure that analyzed effects could reasonably be assumed to result from cognitive training manipulations alone and not the influence of additional manipulations not of interest.

2.2.3. Comparator

Studies were required to include at least one eligible treatment group based on the criteria described above. They were not required to include a comparison group of any kind. If present,

the type of comparison group was recorded for all eligible studies (e.g., active or passive controls).

2.2.4. Baseline & Outcome Measures

Studies were required to have measured participants baseline ability in the trained cognitive skill(s) prior to training. For example, if a study tested the effects of a working memory training, a baseline measure of working memory ability prior to training was required for inclusion. This baseline measure could come from the same task that was later used for training or from a different task that assessed the same cognitive skill. Studies were required to have measured participants cognitive training outcomes using gain scores. Gain scores were defined as any calculation of the difference in performance between baseline performance and final performance. To be eligible these outcomes needed to measure the trained cognitive skill(s), and were thus derived from either final training performance or performance on a near transfer task.

2.2.5. Study Design

Studies were required to include a pre- and post- design such that there were measurements of participants performance before and after the cognitive training. Studies needed to be experimental or quasi-experimental in nature. Observational studies, case studies, case-series, qualitative studies, or studies not using a pre-/post-design were excluded.

2.2.6. Other

Studies were not excluded for administration of multiple types of cognitive training across multiple groups, provided that at least one cognitive training group met the above criteria. No exclusion criteria were imposed based on the country of origin, but only papers for which a full-text English version of the article was available were evaluated for eligibility.

2.3. Study Selection

After search results were collected using the search protocol described above, results were saved in a systematized electronic format to be reviewed for eligibility. Prior to review, duplicate papers were flagged through the use of an automated script. To confirm automated results, the first author manually verified duplication of flagged entries.

Post-duplication flagging, results were scanned for inclusion based on titles and abstracts. Results were flagged and removed at this stage if they were obviously ineligible (i.e., pertaining to unrelated topic or clearly not conducting a cognitive training intervention) or failed to meet eligibility criteria (e.g., subjects were not human). Full text documents were obtained for remaining articles. Each full-text document was screened for eligibility based on the established eligibility criteria discussed above. Results are reported below.

2.4. Data Collection & Coding

All papers that passed the eligibility screening processes were coded for relevant data. These data included number of training groups, sample size (broken down by training group), cognitive training characteristics (including trained domain, training approach, training tasks, etc.), and relevant statistical

analyses. From the relevant statistical analyses, baseline measures and outcome measures were recorded. Data were also coded for participant characteristics (e.g., age(s), neurological status, population). Coders were not masked to the article's authors, institution, or place of publication.

2.5. Correlation Coefficient Extraction

Correlations between baseline ability and gain scores were selected for meta-analysis as this presented the most readily available effect size measure across retrieved studies. Studies were allowed to contribute multiple correlation coefficients if more than one was provided. For example, if a study reported multiple eligible correlation coefficients assessing the relationship between cognitive baseline and training outcomes—all of these correlations were recorded. This was done to allow the maximum amount of observations to be included in meta-analysis as well as to avoid introducing bias by determining an *ad-hoc* method of selecting a single effect size from each study for inclusion. Study is included as a level in our multi-level model as described below.

2.6. Analysis Plan

After extraction, correlation coefficients were scaled using a Fisher's Z transformation (Lipsey and Wilson, 2001). Publication bias was evaluated using Egger's regression test and visual inspection for funnel plot asymmetry. This permitted evaluation of the degree to which the characteristics of the funnel plot deviate from normal based on their precision (Egger et al., 1997). Heterogeneity between studies was evaluated using both Cochrane's Q and I^2 measures. If a significant amount of heterogeneity was observed, initial analyses were followed with those described below.

2.6.1. Model 1: Meta-Analytic Correlation

To estimate the mean association between individual differences in baseline and cognitive training gains, a three-level meta-analytic model was implemented. A multi-level model was selected to allow for multiple correlation coefficients to be clustered within a study to both account for similarities between those correlations inherent to having been collected from the same study and control for the possibility of one study with multiple significant observations from over-contributing to results (Konstantopoulos, 2011; Van den Noortgate et al., 2013; Cheung, 2014). This approach also incorporates weighting each observation in analysis according to its respective sensitivity—i.e., sample size—such that effect sizes derived from larger samples contributed more to analysis (Borenstein et al., 2010). The complete model therefore included random effects of both study (Level 3) and correlation coefficient (Level 2). Level 3 captured variance between studies in which effect sizes were clustered. Level 2 captured variance among the effect sizes within these studies. Level 1 captured the sampling variance of the actual correlation coefficient effect sizes themselves. This model is outlined in Equation (1), where *i* refers to individual correlation coefficient observation, *j* refers to individual studies. κ refers to the average correlation coefficient within study *j*, β refers to the average population level correlation coefficient, θ refers to true correlation coefficient, and $\hat{\theta}$ refers to the estimate of observation

i within study *j*.

$$Level1 : \hat{\theta}_{ij} = \theta_{ij} + \epsilon_{ij}$$
$$Level2 : \theta_{ij} = \kappa_j + \xi_{(2)ij} \tag{1}$$
$$Level3 : \kappa_j = \beta_0 + \xi_{(3)j}$$

2.6.2. Model 2: Meta-Regression

To evaluate the prediction that training approach and trained domain moderate the association between baseline ability and training gains, a second three-level mixed effects meta-analytic model was planned. Levels 1 and 2 were identical to those of Model 1. At Level 3, fixed effects for training approach (strategy or process) and trained domain (episodic memory or EF) were included as predictors.

$$Level1 : \hat{\theta}_{ij} = \theta_{ij} + \epsilon_{ij}$$
$$Level2 : \theta_{ij} = \kappa_j + \xi_{(2)ij}$$
$$Level3 : \kappa_j = \beta_0 + \beta_1(CognitiveDomain) + \tag{2}$$
$$\beta_2(TrainingApproach) + \xi_{(3)j}$$

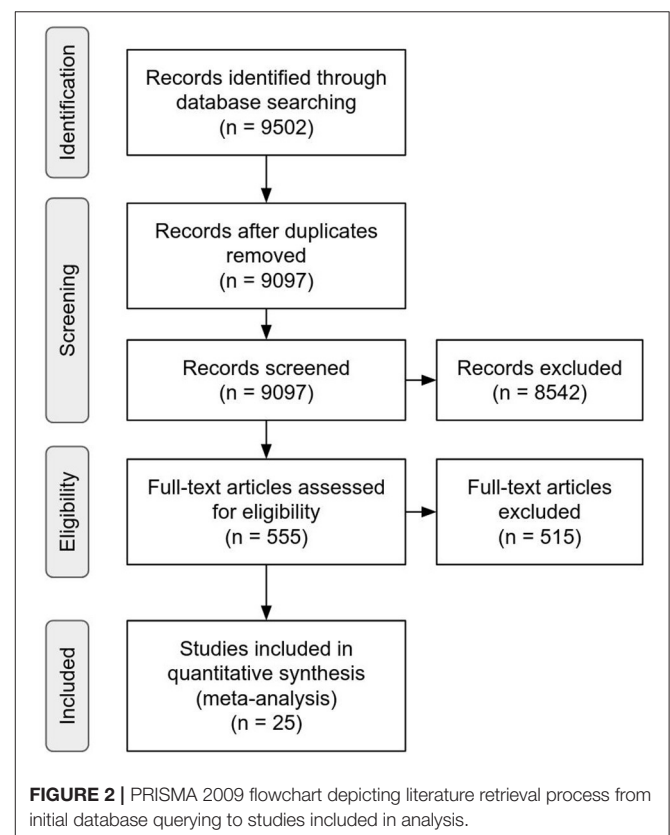Both training approach and trained domain predictors were dummy coded.

FIGURE 2 | PRISMA 2009 flowchart depicting literature retrieval process from initial database querying to studies included in analysis.

3. RESULTS

3.1. Search Results

The initial search protocol within Google Scholar (run between November–December 2018) produced 9,502 documents, of which 405 were identified as duplicates. During the screening of the remaining 9,097 unique documents abstracts and titles, 8,542 documents were excluded. The remaining 555 underwent full text screening for eligibility. A total of 42 papers met the eligibility criteria described above. Of these, 15 reported sufficient information for extraction or calculation of a correlation coefficient. Twenty-seven papers met eligibility criteria but did not report sufficient information for calculation of relevant effect sizes. Authors of these papers were contacted and requested to provide either the requisite correlation or the raw data necessary to calculate the requisite correlations. Ten papers were added via this contact procedure for a final sample of 25 papers (**Figure 2**; **Table 1**). All reported meta-analytic tests were conducted based on a total of 82 correlation coefficients extracted from this set of papers reporting correlation coefficients (see **Table 1**). A range of 1 to 8 correlation coefficients were reported in each paper with an average of 2.5. One study (Lövdén et al., 2012) conducted a two-part training consisting of a strategy approach followed by a process approach. Eligible correlation coefficients were presented after each training, so correlation coefficients presented after the initial strategy training are included whereas the post-process training data are excluded. Of retrieved papers, 3 were episodic memory trainings and 22 were executive function trainings (including 18 working memory, 2 inhibition, 1 task switching, and 1 selective attention). There was a 1:1 ratio of cognitive domain and training approach such that all 3 episodic memory trainings were strategy-based and all 22 executive function trainings were process-based approaches.

3.2. Analysis Results

All analyses were carried out using the R Opensource Framework, R version 3.6.0, (R Core Team, 2020) and the metafor package (Viechtbauer, 2010; Assink and Wibbelink, 2016). Correlation coefficients were standardized via Fisher's Z transformation (Lipsey and Wilson, 2001). For extended analysis reports see **Supplementary Materials**.

3.2.1. Bias

Publication bias was assessed over both individual effect sizes and averaged effect sizes within a study. This was done to most accurately reflect the inclusion of multiple effect sizes from within a single study and adhere to the purpose of bias evaluation methods to be conducted over individual studies (Fernández-Castilla et al., 2020). In both cases, bias was assessed by visual inspection of a funnel plot and testing for funnel plot asymmetry using Egger's regression. Across all effect sizes, Egger's regression suggested asymmetry was present within the corresponding funnel plot at the recommended threshold of below $p = 0.10$ (z = −1.7653; $p = 0.0775$) (Egger et al., 1997). Visual inspection suggested two small n effect sizes in the lower right (both from the paper Hickey, 2018) might be the source of this asymmetry (see **Figure 3**). Evaluation of bias across

effect sizes aggregated within study would suggest this is the case. The aggregated effect size funnel plot (**Figure 4**) shows little indication of asymmetry with visual inspection, revealing a more symmetrical funnel. Egger's regression correspondingly returns non-significant findings for asymmetry (z = −0.4545, $p = 0.6495$). Variability in the upper strata along the x axis suggests notable heterogeneity among observations, which is supported by heterogeneity analyses discussed below, and is not in-and-of itself indicative of asymmetry or bias (Sterne et al., 2011).

3.2.2. Model 1: Meta-Analytic Correlation

The initial model stipulated in Equation (1) estimated the average correlation of baseline ability with training gains. Results demonstrated a significant negative correlation between baseline cognitive performance and training gain ($\beta_0 = -0.4490$, z = −4.1584, $p < 0.0001$) (**Figure 5**). The direction of the correlation is consistent with a compensation effect; individuals who show worse performance prior to training tend to reap greater gains, but the result is inconclusive. A significant amount of heterogeneity between observations was evidenced [$Q_{(81)} = 924.5982$, $p < 0.0001$; $I^2 = 91.17$, $p < 0.0001$] indicating a large amount of variability across the collected literature (Higgins and Thompson, 2002). This high level of heterogeneity was expected due to the diversity of training approaches, ages, and populations included in this review, as well as the hypothesized relevance of the training characteristic predictors for this correlation. The multi-level nature of this model allowed for further separation of this heterogeneity across between- and within-study clusters (Nakagawa and Santos, 2012). This break-down indicated that 20.37% of heterogeneity was between clusters (i.e., between individual studies) and the remaining 70.8% was within clusters (i.e., within individual studies) (**Figure 6**). Complete model results are reported in **Table 2**.

3.2.3. Model 2: Meta-Regression

As analysis of heterogeneity in the base model indicated a significant amount of variability between observations, meta-regression analyses were conducted. However, the characteristics of the collected eligible correlations did not allow for the original model analyzing both the main effects and interaction between cognitive training domain and cognitive training approach to be carried out. As foreshadowed by prior work, executive function and episodic memory training studies in the present sample perfectly aligned with the use of process and strategy training approaches, respectively (see **Table 1**).

$$Level1 : \hat{\theta}_{ij} = \theta_{ij} + \epsilon_{ij}$$
$$Level2 : \theta_{ij} = \kappa_j + \xi_{(2)ij} \quad (3)$$
$$Level3 : \kappa_j = \beta_0 + \beta_1(DomainApproach) + \xi_{(3)j}$$

A large amount of residual heterogeneity remained after the addition of the domain/approach predictor [$QE_{(80)} = 891.1806$, $p < 0.0001$]. There was no evidence for a significant relationship between training domain/approach and the correlation of baseline individual differences with cognitive training gain ($\beta_1 = -0.3082$, z = −0.9537, $p = 0.3402$). The estimated average correlation remained negative and significant ($\beta_0 = -0.4099$, z

TABLE 1 | List of all observations used in reported analyses.

Author (Year)	Correlation ID	Training domain	Training approach	*n*	*r*
Chan et al. (2015)	1	EF	Process	25	−0.72
"	2	EF	Process	25	−0.51
"	3	EF	Process	25	−0.50
"	4	EF	Process	25	−0.81
"	5	EF	Process	25	−0.17
"	6	EF	Process	25	−0.30
"	7	EF	Process	25	−0.53
"	8	EF	Process	25	−0.37
Chooi (2011)	1	EF	process	9	-0.04
"	2	EF	Process	13	−0.48
Fellman et al. (2018)	1	EF	Process	25	−0.62
Foster et al. (2017)	1	EF	Process	30	−0.26
"	2	EF	Process	30	−0.15
"	3	EF	Process	36	−0.34
"	4	EF	Process	36	0.06
Gade et al. (2017)	1	EF	Process	10	−0.86
"	2	EF	Process	16	−0.71
"	3	EF	Process	10	−0.95
"	4	EF	Process	10	−0.74
Gunn et al. (2018)	1	EF	Process	70	−0.01
Hickey (2018)	1	EF	Process	5	0.05
"	2	EF	Process	5	0.70
Jones et al. (2020)	1	EF	Process	41	0.34
Karbach et al. (2017)	1	EF	Process	126	0.81*
Karbach et al. (2015)	1	EF	Process	14	−0.39
Lövdén et al. (2012)	1	EP	strategy	50	−0.85
"	2	EP	Strategy	29	−0.99
"	3	EP	Strategy	29	−0.90
"	4	EP	Process	50	0.33
"	5	EP	Process	29	−0.02
"	6	EP	Process	29	−0.09
McKitrick et al. (1999)	1	EP	Strategy	224	−0.33
"	2	EP	Strategy	224	−0.44
O'Brien et al. (2013)	1	EF	Process	11	−0.63
De Simoni and von Bastian (2018)	1	EF	Process	59	−0.4
"	2	EF	Process	59	−0.26
"	3	EF	Process	59	−0.52
"	4	EF	Process	59	−0.58
"	5	EF	Process	66	−0.28
"	6	EF	Process	66	−0.48
"	7	EF	Process	65	−0.15
"	8	EF	Process	66	0.05
Singer et al. (2003)	1	EP	Strategy	96	0.14
Stepankova et al. (2014)	1	EF	Process	20	−0.01
"	2	EF	Process	20	0.02
Strobach and Huestegge (2017)	1	EF	Process	76	0.15
"	2	EF	Process	76	0.16
"	3	EF	Process	76	0.40
"	4	EF	Process	76	0.39
"	5	EF	Process	76	0.24

(Continued)

Author (Year)	Correlation ID	Training domain	Training approach	n	r
"	6	EF	Process	76	−0.47
Vermeij et al. (2017)	1	EF	Process	21	0.03
"	2	EF	Process	21	−0.58
"	3	EF	Process	20	−0.63
"	4	EF	Process	20	−0.39
"	5	EF	Process	14	−0.16
"	6	EF	Process	14	−0.30
"	7	EF	Process	10	−0.22
"	8	EF	Process	10	−0.50
Volckaert and Noël (2015)	1	EF	Process	24	−0.55
Volckaert and Pascale Noel (2017)	1	EF	Process	16	−0.72
von Bastian and Oberauer (2013)	1	EF	Process	30	−0.27
"	2	EF	Process	30	−0.54
"	3	EF	Process	31	−0.79
de Vries et al. (2018)	1	EF	Process	81	0.51
Weicker et al. (2018)	1	EF	Process	20	−0.13
"	2	EF	Process	20	−0.25
"	3	EF	Process	20	−0.60
"	4	EF	Process	20	−0.31
"	5	EF	Process	20	−0.63
"	6	EF	Process	20	−0.59
"	7	EF	process	20	−0.53
"	8	EF	Process	20	−0.55
"	9	EF	Process	19	−0.56
Zinke et al. (2012)	1	EF	Process	20	−0.66
"	2	EF	Process	20	−0.59
"	3	EF	Process	20	−0.65
"	4	EF	Process	20	−0.89
"	5	EF	Process	20	−0.80
Zinke et al. (2014)	1	EF	Process	40	−0.46
"	2	EF	Process	40	−0.17
"	3	EF	Process	40	−0.32
"	4	EF	Process	40	0.10
"	5	EF	Process	40	−0.64
"	6	EF	Process	40	−0.30

EF, executive function; EP, episodic memory.
*Indicates correlation was multiplied by -1 for interpretability purposes in analysis.

$= -3.5574$, $p = 0.0004$). Complete model results are reported in **Table 3**.

3.2.4. Model 3: Exploratory Meta-Analytic Correlation of Executive Function Papers Only

Given the paucity of qualifying episodic memory/strategy cognitive training papers, a third exploratory analysis was conducted using only executive function/process cognitive training papers. This analysis evaluated the meta-analytic correlation coefficient for this subset of the retrieved effect sizes using the same equations described in Equation (1). A total of 22 papers with 76 effect sizes were included. Complete model results are reported in **Table 4**.

Evaluation of publication bias in this restricted dataset demonstrated similar patterns to the complete dataset with significant results in an Egger's test for funnel plot asymmetry ($z = -2.5614$, $p = 0.0104$). This restricted dataset again demonstrated a significant negative association between baseline ability and gains in that ability as a result of training ($\beta_0 = -0.4101$, $z = -4.1074$, $p < 0.0001$) with a significant amount of heterogeneity [$Q_{(75)} = 676.4997$, $p < 0.0001$] (**Figure 7**).

4. DISCUSSION

We highlight three main findings from this systematic meta-analysis of the cognitive training literature. First, baseline

FIGURE 3 | Contoured funnel plot of all effect sizes. Fisher's Z transformed correlation coefficients for each reported effect size are plotted against the standard error. The dark dashed line indicates the summary effect size on which the funnel is centered. The lighter dashed line indicates a null effect. The funnel is centered on the summary effect size. Egger's regression suggests mild asymmetry present in the plot ($z = -1.7653$; $p = 0.0775$).

individual differences in the cognitive ability trained yield significant compensatory effects. Participants with initially weaker targeted abilities gain more from cognitive training than those who are initially more proficient. This meta-analytic finding is consistent with multiple prior accounts of the influence of individual differences on cognitive training gains, par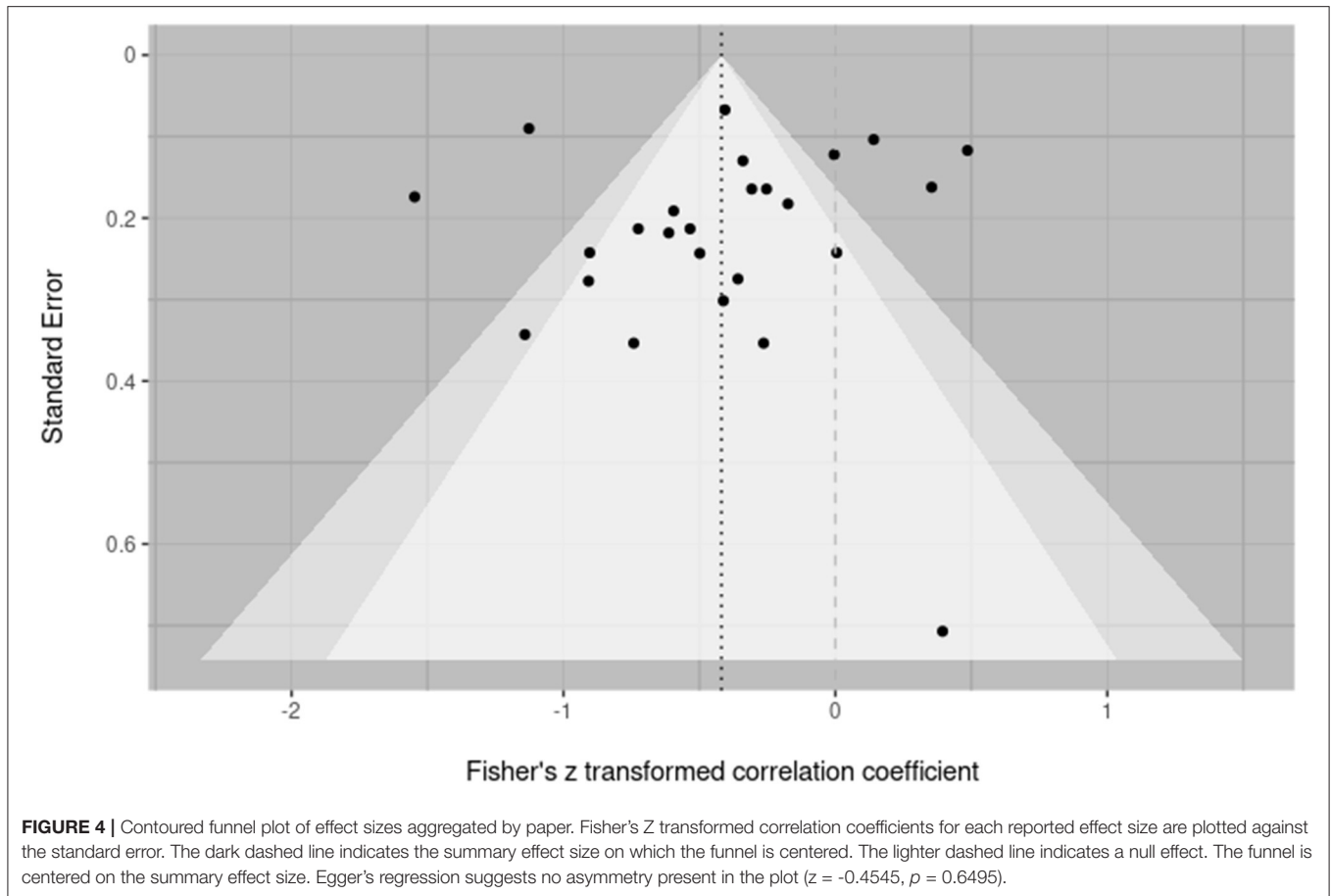ticularly for executive function and process based cognitive training approaches (Lövdén et al., 2012; Karbach and Unger, 2014), and is the first to demonstrate it via a systematic literature synthesis. Second, the literature did not support an interrogation of how this effect might differ based on cognitive domain being trained (executive function vs. episodic memory) and/or training approach (process- vs. strategy-based). Our review retrieved too few eligible papers reporting results of episodic memory or strategy-based training to compare correlations to those of executive function or process-based cognitive trainings—raising questions as to why these papers in particular did not qualify for inclusion. This was accompanied by little variation in the use of process-based approaches with executive function trainings and strategy-based approaches with episodic memory trainings (with a handful of notable exceptions, e.g., Lövdén et al., 2012). Third, despite our sample being predominantly executive function and process-based, a significant amount of heterogeneity was evidenced, suggesting a range of moderators outside of cognitive domain and training

approach are at play. We discuss each of these important findings in turn.

4.1. A Compensatory Meta-Correlation

Why would the weak gain the most? Cognitive training may close the proficiency gap by providing exposure to the targeted cognitive skill, bolstering individuals' mastery of fundamentals necessary for succeeding with that skill—fundamentals that more advanced participants may have already mastered. Relatedly, compensatory patterns could be the result of ceiling effects, with higher performing individuals having little to no room to improve on existing measures with further cognitive training due to their already high abilities. A ceiling-effect account of compensatory effects could be explored in future systematic reviews that analyze the extent to which effective ceilings were reached across different measures and potential moderating effects of this measurement factor.

Compensatory effects could also arise from participants' expectations about how much they will improve from training, if participants with worse initial abilities hold stronger beliefs that an intervention will be beneficial to them—in comparison to more proficient participants who might feel confident in their existing abilities. This possibility highlights the importance of measuring or

FIGURE 4 | Contoured funnel plot of effect sizes aggregated by paper. Fisher's Z transformed correlation coefficients for each reported effect size are plotted against the standard error. The dark dashed line indicates the summary effect size on which the funnel is centered. The lighter dashed line indicates a null effect. The funnel is centered on the summary effect size. Egger's regression suggests no asymmetry present in the plot ($z = -0.4545$, $p = 0.6495$).

manipulating expectations to address their role in training gains (Boot et al., 2013) and their potential interactions with initial ability.

Lastly, compensatory effects could be the result of the specific effect size coefficients extracted for this study: correlation coefficients reflecting the association between baseline ability and gain. Gain scores are most commonly calculated as the difference between performance at test and that original baseline score. The use of this method has been argued to be subject to demonstrating regression to the mean, leading to a greater chance of yielding a compensatory effect (Smoleń et al., 2018). A growing body of work on individual differences in cognitive training gains has moved toward analyzing individual differences effects on cognitive training gains through alternative metrics. These studies test the impact of training via approaches such as latent growth curve modeling or structural equation modeling, which account for variability to support better estimates of the relationship between baseline ability and gain (e.g., Lövdén et al., 2012; Foster et al., 2017; Guye et al., 2017, and others). Given that correlations were the dominant available coefficient from which to calculate a meta-analytic effect, this review highlights the need for both consistency and deeper methodological consideration in future analyses of effects of individual differences in baseline ability on benefits from cognitive training.

4.2. Consequences of Eligibility Criteria

There are multiple potential reasons for the difficulty in retrieving eligible papers across our moderators of interest and their combinations. First, our eligibility requirements were stringent. An unexpected difficulty during eligibility screening of retrieved papers was the number of studies utilizing a multi-modal training approach that did not permit for a controlled analysis of the influence of cognitive training in the absence of other approaches (e.g., group therapies, aerobic exercise, and other educational interventions). Future work might include such multi-modal studies to increase the number of usable observations for analysis and more thoroughly represent the current characteristics of the cognitive training literature. While it was important to narrow our scope of inquiry to clearly identify sources of variation as was done here, it will also be valuable to evaluate the possible true effect size of baseline individual differences within the types of studies that are frequently conducted.

Similarly, the eligibility criteria used in this systematic review revealed a paucity of studies meeting requirements for the specific baseline cognition and training gains association being analyzed. Eligibility for inclusion required studies to have investigated the association between individual differences in baseline cognitive ability, with that cognitive ability defined as being of the same
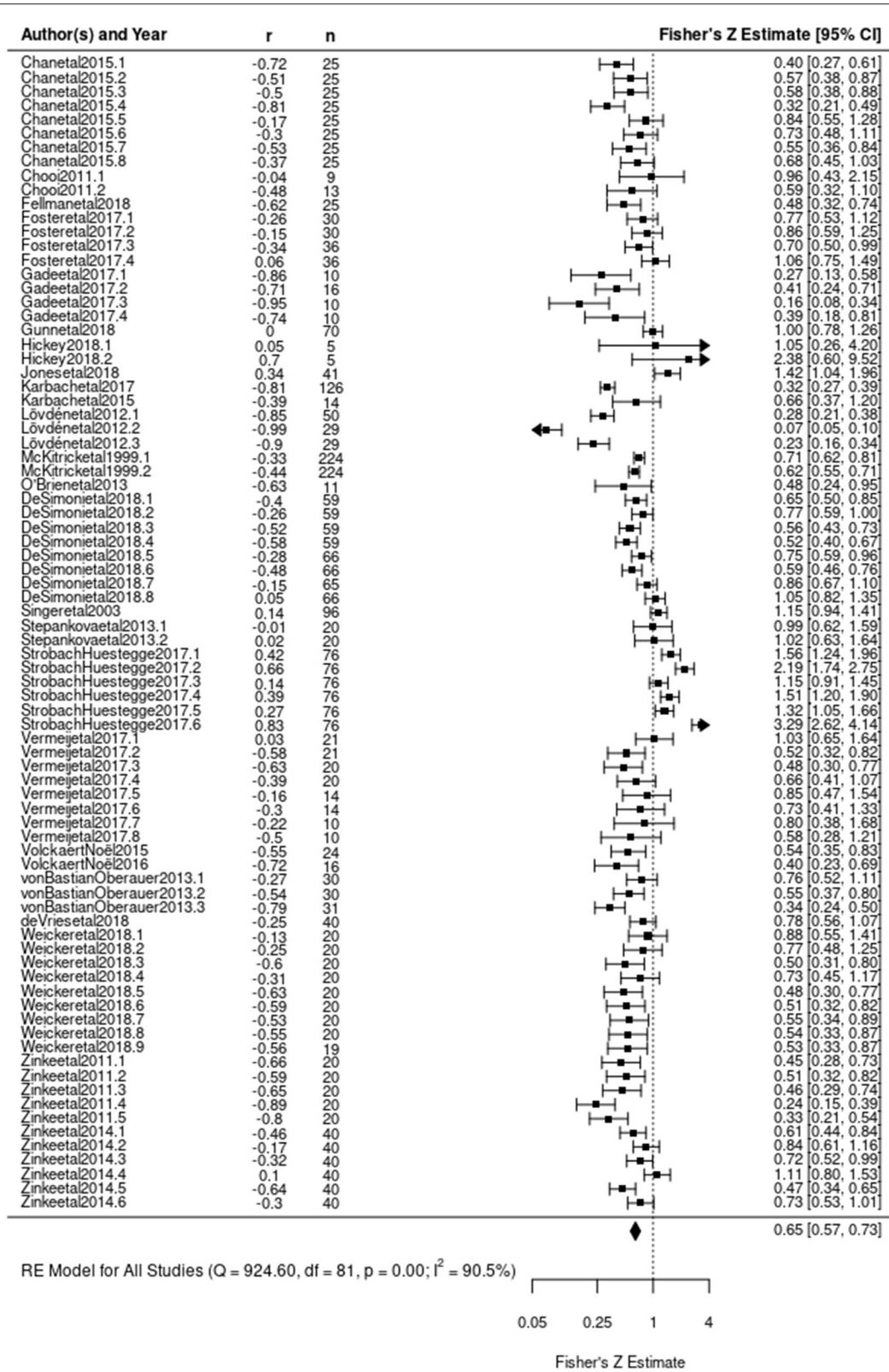
FIGURE 5 | Forest plot of Model 1 results. Each correlation coefficient is plotted according to its Fisher's Z transformation. Size of square indicates weight assigned to the observation during analysis. Length of lines extending from square indicates variance associated with observation based on sample size. Each observation's Fisher's Z transformation and confidence interval are presented on the right. The diamond at the bottom indicates the estimated Fisher's Z transformed correlation coefficient derived from Model 1.

FIGURE 6 | Meta-correlation (Model 1) variance breakdown by level. The majority of heterogeneity within this model results from Level 3.
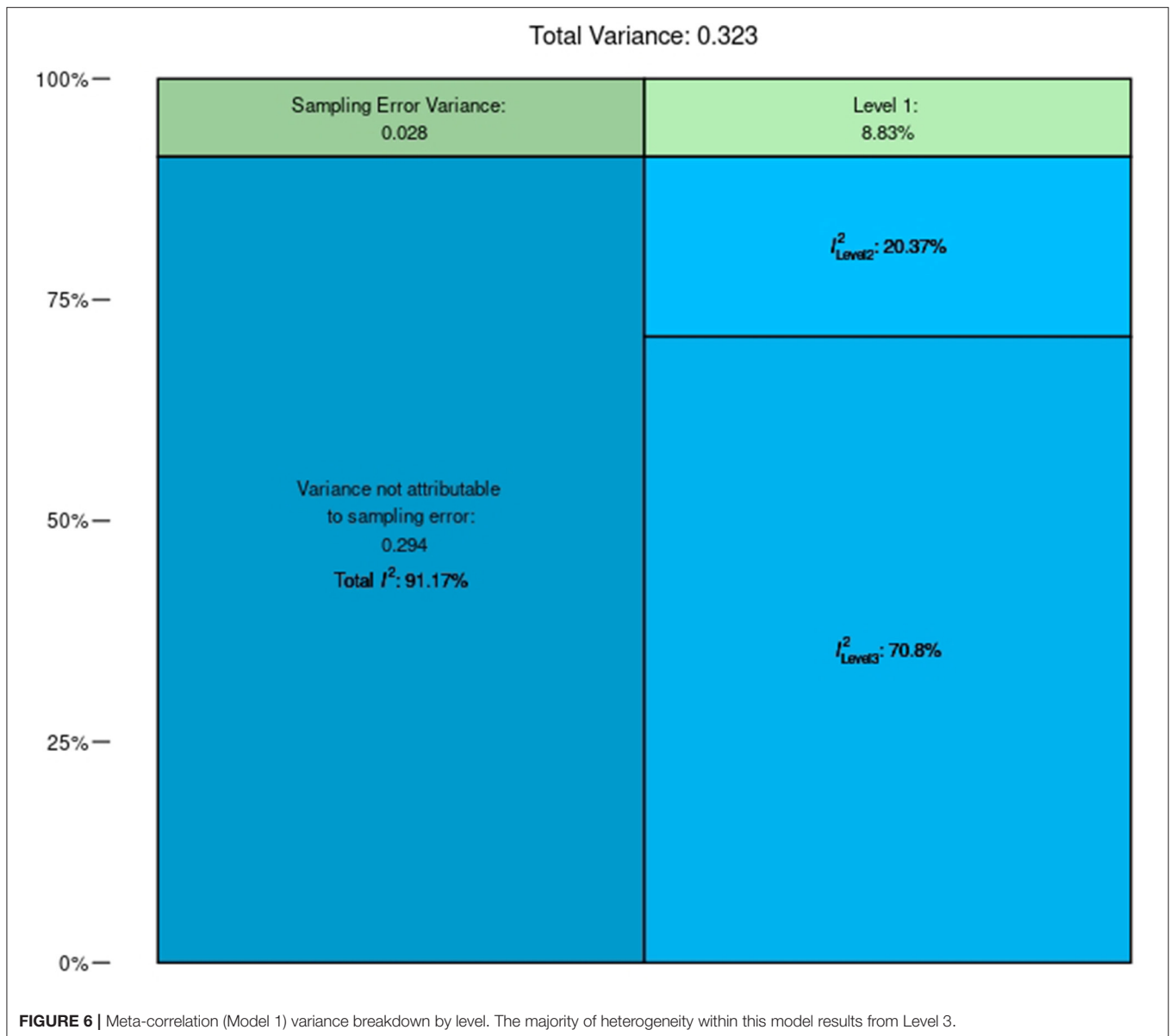
type as the one trained, and with cognitive training gains derived from either the final training performance or near transfer performance. This isolated the specific relationship in question to initial ability in cognitive construct "X" and gains in "X" from training in "X." Many papers were disqualified for deviating from either or both of the baseline cognitive measure and training outcome requirements, highlighting inconsistencies in what these common terms actually represent across the literature. For example, one of the most cited works on this topic analyzed the association of a baseline measure of general fluid intelligence with far transfer outcomes on another fluid intelligence task, after working memory training (Jaeggi et al., 2008). This work is relevant to understanding individual differences in cognitive training, and many papers talk about results such as these

in tandem with the types of studies included in this meta-analysis (e.g., Lustig et al., 2009; Lövdén et al., 2012; Karbach and Unger, 2014; Borella et al., 2017; Karbach et al., 2017). However, individual differences in baseline abilities may predict the effectiveness of training in different ways depending on whether baseline, training, and gains target the same ability (e.g., working memory) or target different abilities that require transfer (e.g., training working memory and testing fluid intelligence gains and baseline). Cognitive training could show greater magnification effects in the context of transfer tasks (e.g., Borella et al., 2017; cf. Karbach et al., 2017), given the need to not only acquire a relevant skill but to apply it to another domain, which might require sufficient initial proficiency in that domain. Future work could systematically assess this question, with the

TABLE 2 | Model 1: Meta-correlation results.

Model Estimates

	Estimate	SE	*p*-value	Lower bound	Upper bound
Correlation value	−0.4490	0.1080	<0.0001	−0.661	−0.2374
Variance					
	Estimate	Sqrt	Levels		
Level 1: σ^2	0.0657	0.2564	82		
Level 2: σ^2	0.2284	0.4479	25		
Heterogeneity					
	Q(df)	I^2			
	924.598(81)	91.17			

TABLE 3 | Model 2: Meta-regression results.

Model Estimates

	Estimate	SE	*p*-value	Lower bound	Upper bound
Correlation value	−0.4099	0.1152	0.0004	−0.6357	−0.1841
Factor value	−0.4490	0.1080	0.3402	−0.9416	0.3520
Variance					
	Estimate	Sqrt	Levels		
Level 1: σ^2	0.0666	0.2580	82		
Level 2: σ^2	0.2264	0.4758	25		
Heterogeneity					
	QM(df)	QE(df)			
	0.9096(1)	891.181(80)			

challenge of comparing transfer measures that span a wide range of distances from trained cognitive abilities.

4.3. Alternative Moderators

While we were unable to sufficiently explore the hypothesized moderators of cognitive domain and training approach, analyses indicated a significant amount of heterogeneity among correlations in both our complete dataset and our exploratory EF-process only dataset. Other moderators not investigated here likely influence whether baseline individual differences predict cognitive training benefits. For example, different age groups respond differently to training, which has been interpreted through the lens of age group differences in cognitive baseline (Lövdén et al., 2012; Bürki et al., 2014; Borella et al., 2017; Karbach et al., 2017). We did not explore age effects within dataset as many observations incorporated multiple age groups (see **Supplementary Table 1**).

Responsiveness to training is likely also influenced by typical or atypical cognitive status—with specific aspects of an individual's condition potentially impacting the potency of baseline individual differences effects. Many meta-analyses of cognitive training effects limit eligibility to healthy populations for this reason (e.g., Au et al., 2015). Our dataset included a range of atypical populations including individuals with ADD/ADHD, autism, Parkinson's disease and other conditions (see **Supplementary Table 1**), making it difficult to pose a

TABLE 4 | Model 3: Exploratory Meta-correlation results with executive function/process studies only.

Model Estimates

	Estimate	SE	*p*-value	Lower bound	Upper bound
Correlation value	−0.4101	0.0998	<0.0001	−0.6057	−0.2144
Variance					
	Estimate	Sqrt	Levels		
Level 1: σ^2	0.0521	0.2283	76		
Level 2: σ^2	0.1629	0.4036	22		
Heterogeneity					
	Q(df)	I^2			
	676.499(75)	86.51%			

unified hypothesis for how atypical groups with disparate etiologies might be differently affected by cognitive training when compared to healthy samples. An exploration of typical vs. atypical cognitive status yielded no significant effect on the meta-analytic correlation ($\beta_1 = -0.1758$, z = -0.8930, p < 0.3719), likely because of the small number of atypical observations and the range of conditions within our sample. Future work might focus on the question of how typical vs. atypical groups' baseline abilities influence cognitive training outcomes within a specific condition (e.g., ADD/ADHD) to probe the nuances of magnification or compensation effects based on specific patient characteristics.

Outside of participant characteristics, compensation and magnification effects may also vary based on characteristics of cognitive trainings, such as the specific cognitive components trained. The present dataset, for example, included trainings of distinct executive function processes. Because of the small number of non-working memory executive functions trainings retrieved we were not able to examine the influence of specific executive function effects within the present dataset. Our work sought to determine the meta-analytic effect most representative of existing studies including analysis of baseline individual differences effects on cognitive training outcomes without regard for specific executive function components. This was done to be consistent with the prevailing discussion on the nature of these effects—which typically does not distinguish between specific executive function components. Future work should delve into the potential differences that might arise in this effect between different types of executive functions, as has been done in meta-analytic work looking at the overall effect of cognitive training specific to component types (e.g., Au et al., 2015).

Lastly, the overall potency of a cognitive training (i.e., how effective that training is compared to an active control) could influence who benefits from that training, with weaker interventions tending to benefit those who need them most and stronger interventions showing more equitable benefits. Heterogeneity in the present sample might be driven by moderating effects of any of these factors. While our investigation sought to clarify the pattern of compensation and magnification results seen across executive function and episodic memory training studies using process- and strategy-based approaches,
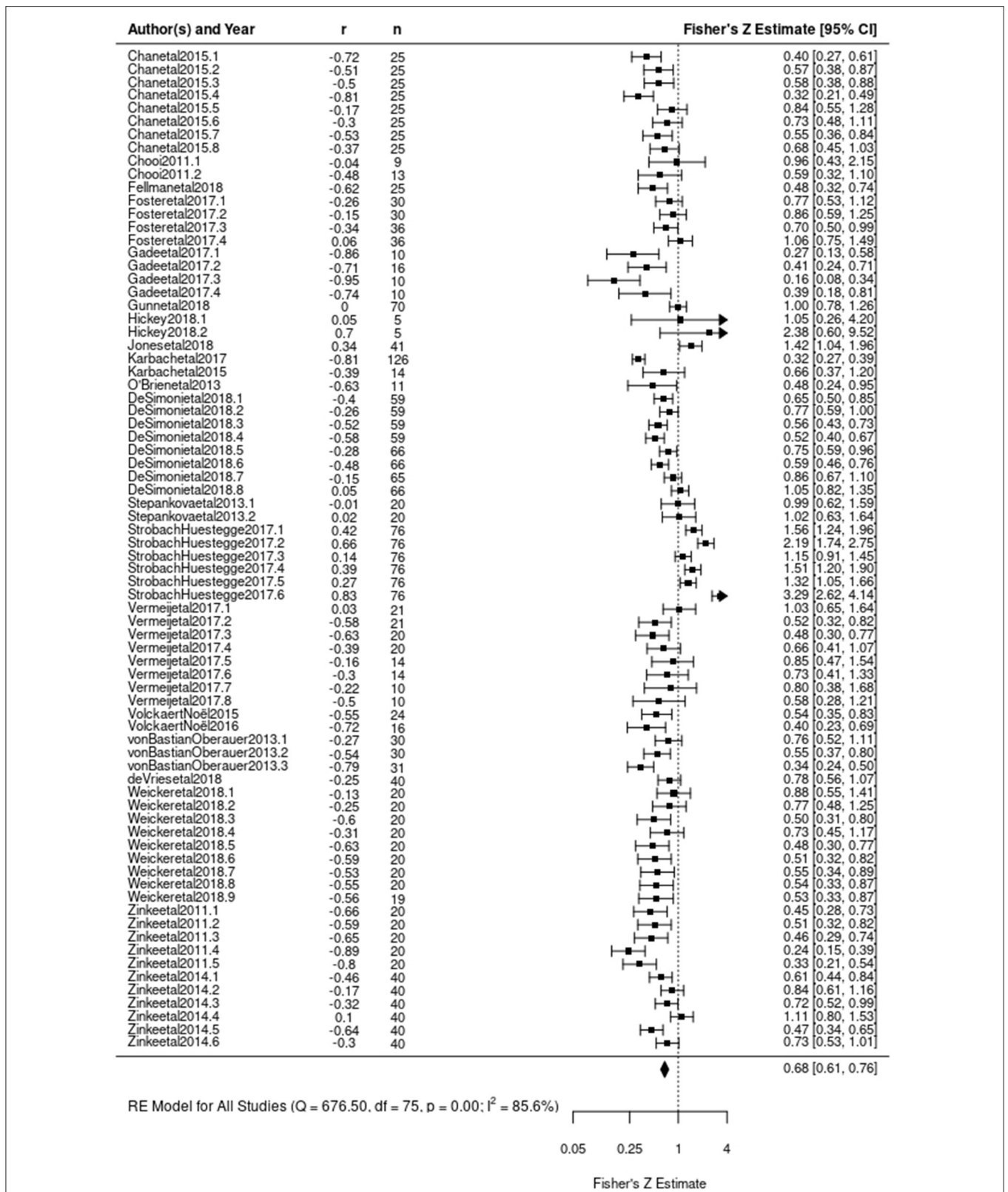
Author(s) and Year	r	n	Fisher's Z Estimate [95% CI]
Chanetal2015.1	-0.72	25	0.40 [0.27, 0.61]
Chanetal2015.2	-0.51	25	0.57 [0.38, 0.87]
Chanetal2015.3	-0.5	25	0.58 [0.38, 0.88]
Chanetal2015.4	-0.81	25	0.32 [0.21, 0.49]
Chanetal2015.5	-0.17	25	0.84 [0.55, 1.28]
Chanetal2015.6	-0.3	25	0.73 [0.48, 1.11]
Chanetal2015.7	-0.53	25	0.55 [0.36, 0.84]
Chanetal2015.8	-0.37	25	0.68 [0.45, 1.03]
Chooi2011.1	-0.04	9	0.96 [0.43, 2.15]
Chooi2011.2	-0.48	13	0.59 [0.32, 1.10]
Fellmanetal2018	-0.62	25	0.48 [0.32, 0.74]
Fosteretal2017.1	-0.26	30	0.77 [0.53, 1.12]
Fosteretal2017.2	-0.15	30	0.86 [0.59, 1.25]
Fosteretal2017.3	-0.34	36	0.70 [0.50, 0.99]
Fosteretal2017.4	0.06	36	1.06 [0.75, 1.49]
Gadeetal2017.1	-0.86	10	0.27 [0.13, 0.58]
Gadeetal2017.2	-0.71	16	0.41 [0.24, 0.71]
Gadeetal2017.3	-0.95	10	0.16 [0.08, 0.34]
Gadeetal2017.4	-0.74	10	0.39 [0.18, 0.81]
Gunnetal2018	0	70	1.00 [0.78, 1.26]
Hickey2018.1	0.05	5	1.05 [0.26, 4.20]
Hickey2018.2	0.7	5	2.38 [0.60, 9.52]
Jonesetal2018	0.34	41	1.42 [1.04, 1.96]
Karbachetal2017	-0.81	126	0.32 [0.27, 0.39]
Karbachetal2015	-0.39	14	0.66 [0.37, 1.20]
O'Brienetal2013	-0.63	11	0.48 [0.24, 0.95]
DeSimonietal2018.1	-0.4	59	0.65 [0.50, 0.85]
DeSimonietal2018.2	-0.26	59	0.77 [0.59, 1.00]
DeSimonietal2018.3	-0.52	59	0.56 [0.43, 0.73]
DeSimonietal2018.4	-0.58	59	0.52 [0.40, 0.67]
DeSimonietal2018.5	-0.28	66	0.75 [0.59, 0.96]
DeSimonietal2018.6	-0.48	66	0.59 [0.46, 0.76]
DeSimonietal2018.7	-0.15	65	0.86 [0.67, 1.10]
DeSimonietal2018.8	0.05	66	1.05 [0.82, 1.35]
Stepankovaetal2013.1	-0.01	20	0.99 [0.62, 1.59]
Stepankovaetal2013.2	0.02	20	1.02 [0.63, 1.64]
StrobachHuestegge2017.1	0.42	76	1.56 [1.24, 1.96]
StrobachHuestegge2017.2	0.66	76	2.19 [1.74, 2.75]
StrobachHuestegge2017.3	0.14	76	1.15 [0.91, 1.45]
StrobachHuestegge2017.4	0.39	76	1.51 [1.20, 1.90]
StrobachHuestegge2017.5	0.27	76	1.32 [1.05, 1.66]
StrobachHuestegge2017.6	0.83	76	3.29 [2.62, 4.14]
Vermeijetal2017.1	0.03	21	1.03 [0.65, 1.64]
Vermeijetal2017.2	-0.58	21	0.52 [0.32, 0.82]
Vermeijetal2017.3	-0.63	20	0.48 [0.30, 0.77]
Vermeijetal2017.4	-0.39	20	0.66 [0.41, 1.07]
Vermeijetal2017.5	-0.16	14	0.85 [0.47, 1.54]
Vermeijetal2017.6	-0.3	14	0.73 [0.41, 1.33]
Vermeijetal2017.7	-0.22	10	0.80 [0.38, 1.68]
Vermeijetal2017.8	-0.5	10	0.58 [0.28, 1.21]
VolckaertNoël2015	-0.55	24	0.54 [0.35, 0.83]
VolckaertNoël2016	-0.72	16	0.40 [0.23, 0.69]
vonBastianOberauer2013.1	-0.27	30	0.76 [0.52, 1.11]
vonBastianOberauer2013.2	-0.54	30	0.55 [0.37, 0.80]
vonBastianOberauer2013.3	-0.79	31	0.34 [0.24, 0.50]
deVriesetal2018	-0.25	40	0.78 [0.56, 1.07]
Weickeretal2018.1	-0.13	20	0.88 [0.55, 1.41]
Weickeretal2018.2	-0.25	20	0.77 [0.48, 1.25]
Weickeretal2018.3	-0.6	20	0.50 [0.31, 0.80]
Weickeretal2018.4	-0.31	20	0.73 [0.45, 1.17]
Weickeretal2018.5	-0.63	20	0.48 [0.30, 0.77]
Weickeretal2018.6	-0.59	20	0.51 [0.32, 0.82]
Weickeretal2018.7	-0.53	20	0.55 [0.34, 0.89]
Weickeretal2018.8	-0.55	20	0.54 [0.33, 0.87]
Weickeretal2018.9	-0.56	19	0.53 [0.33, 0.87]
Zinkeetal2011.1	-0.66	20	0.45 [0.28, 0.73]
Zinkeetal2011.2	-0.59	20	0.51 [0.32, 0.82]
Zinkeetal2011.3	-0.65	20	0.46 [0.29, 0.74]
Zinkeetal2011.4	-0.89	20	0.24 [0.15, 0.39]
Zinkeetal2011.5	-0.8	20	0.33 [0.21, 0.54]
Zinkeetal2014.1	-0.46	40	0.61 [0.44, 0.84]
Zinkeetal2014.2	-0.17	40	0.84 [0.61, 1.16]
Zinkeetal2014.3	-0.32	40	0.72 [0.52, 0.99]
Zinkeetal2014.4	0.1	40	1.11 [0.80, 1.53]
Zinkeetal2014.5	-0.64	40	0.47 [0.34, 0.65]
Zinkeetal2014.6	-0.3	40	0.73 [0.53, 1.01]

0.68 [0.61, 0.76]

RE Model for All Studies (Q = 676.50, df = 75, p = 0.00; I^2 = 85.6%)

0.05   0.25   1   4

Fisher's Z Estimate

FIGURE 7 | Forest plot of Model 3 results including only executive function trainings. Each correlation coefficient is plotted according to its Fisher's Z transformation. Size of square indicates weight assigned to the observation during analysis. Length of lines extending from square indicates variance associated with observation based on sample size. Each observation's Fisher's Z transformation and confidence interval are presented on the right. The diamond at the bottom indicates the estimated Fisher's Z transformed correlation coefficient derived from Model 3.

an important direction for future work will be to systematically test the role of these other types of factors to more fully understand to what extent baseline abilities predict benefits from cognitive training.

4.4. Recommendations for Future Work

The availability of data limited both the number of eligible papers and the type of effect size coefficients that could be extracted for meta-analysis. To aid in future meta-analytic endeavors, our minimum recommendation is that papers include the following in their main or supplementary materials: complete correlation tables by group as well as complete descriptive statistics by group for all recorded measures. Complete correlation tables would aid in extraction of correlation coefficients (as was meta-analyzed in the present paper) and complete descriptive statistics would aid in the calculation of Cohen's d for meta-analysis of group mean comparisons. Explicit separation of both types of information by group (if multiple groups are present) is essential to extracting information pertinent to cognitive training.

Further, the inclusion of open data sets would greatly assist in the pursuit of meta-analytic work. Within the context of the present study, open datasets would allow for calculation of gain scores when otherwise eligible papers did not report results in this manner. Moving forward, they would allow for the additional calculation of alternative measures and effect sizes not reported in a specific paper. For example, open data sets could support the calculation of indices designed to account for ceiling effects, such as scaling gain scores by dividing them by how much an individual could have improved (i.e., the difference between the maximum scoring of the test and the score obtained by the individual at baseline). In addition, in cases where studies include treatment and control groups, open datasets could allow for *post-hoc* calculation of treatment effects controlling for baseline covariates. For example, baseline scores could be included as a covariate in an ANCOVA model, with group as an independent variable and outcome score as the dependent variables, to account for individual differences in cognitive training effects (Bonate, 2000; Nunes et al., 2011). Open data sets could further support contemporary recommendations for the investigation of individual differences in cognitive training effects including the use of latent variables (e.g., Lövdén et al., 2012; von Bastian and Oberauer, 2014; Guye et al., 2017) (for extensive discussion see Smoleń et al., 2018).

5. CONCLUSION

Using a focused, systematic approach to evaluate the literature, the present meta-analysis demonstrates evidence for a compensatory effect of initial baseline ability on cognitive training gains, suggesting that those best positioned to benefit from cognitive trainings are those who are relatively weaker in ability. Despite finding that few episodic memory or strategy-based studies met our stringent eligibility criteria, preventing us

from evaluating the influence of cognitive domain and training type, we found that a substantial amount of heterogeneity still remained among our primarily executive function and process-based studies. This indicates that even in this refined sample, other pertinent moderators of the baseline individual differences effect are at play. The prolific complexity in this field was evident in this study. Future work should tackle defining the problem space of individual differences in initial ability and cognitive training gains. What initial abilities are we interested in—those directly related to the cognitive skill being trained, those nearly related, or those with a far relationship? The same question applies for change after cognitive training. What are the most productive definitions of cognitive training to use—the narrow ones implemented here or broader definitions including multi-modal trainings? What are the most appropriate means of capturing these questions statistically? The present meta-analysis indicates a productive direction forward in targeting cognitive training toward lower performing individuals, and reaffirms the urgency of continuing systematic work on this topic.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

HT conceived of the idea with feedback from YM. HT and RG carried out data collection. HT performed all analyses. All authors discussed the results and interpretation. HT took the lead in writing the manuscript. YM provided critical revisions. RG provided input. All authors gave final approval of the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.662139/full#supplementary-material

REFERENCES

Assink, M., and Wibbelink, C. J. (2016). Fitting three-level meta-analytic models in R: a step-by-step tutorial. *Quant. Methods Psychol.* 12, 154–174. doi: 10.20982/tqmp.12.3.p154

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., and Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: a meta-analysis. *Psychon. Bull. Rev.* 22, 366–377. doi: 10.3758/s13423-014-0699-x

Bjork, E. L., and Bjork, R. A. (2011). "Making things hard on yourself, but in a good way: creating desirable difficulties to enhance learning," in *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*, eds M. A. Gernsbacher, R. W. Pew, L. M. Hough, and J. R. Pomerantz (New York: Worth Publishers), 59–68.

Bonate, P. L. (2000). *Analysis of Pretest-Posttest Designs.* Boca Raton, FL: Chapman and Hall/CRC Press.

Boot, W. R., Simons, D. J., Stothart, C., and Stutts, C. (2013). The pervasive problem with placebos in psychology: why active control groups are not sufficient to rule out placebo effects. *Perspect. Psychol. Sci.* 8, 445–454. doi: 10.1177/1745691613491271

Borella, E., Carbone, E., Pastore, M., De Beni, R., and Carretti, B. (2017). Working memory training for healthy older adults: the role of individual characteristics in explaining short- and long-term gains. *Front. Hum. Neurosci.* 11:99. doi: 10.3389/fnhum.2017.00099

Borenstein, M., Hedges, L. V., Higgins, J. P., and Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res. Synth. Methods* 1, 97–111. doi: 10.1002/jrsm.12

Brehmer, Y., Li, S.-C., Müller, V., von Oertzen, T., and Lindenberger, U. (2007). Memory plasticity across the life span: uncovering children's latent potential. *Dev. Psychol.* 43:465. doi: 10.1037/0012-1649.43.2.465

Bürki, C. N., Ludwig, C., Chicherio, C., and de Ribaupierre, A. (2014). Individual differences in cognitive plasticity: an investigation of training curves in younger and older adults. *Psychol. Res.* 78, 821–835. doi: 10.1007/s00426-014-0559-3

Cepeda, N. J., Kramer, A. F., and Gonzalez de Sather, J. (2001). Changes in executive control across the life span: examination of task-switching performance. *Dev. Psychol.* 37:715. doi: 10.1037/0012-1649.37.5.715

Chan, J. S., Wu, Q., Liang, D., and Yan, J. H. (2015). Visuospatial working memory training facilitates visually-aided explicit sequence learning. *Acta Psychol.* 161, 145–153. doi: 10.1016/j.actpsy.2015.09.008

Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychol. Methods* 19:211. doi: 10.1037/a0032968

Chevalier, N., Chatham, C. H., and Munakata, Y. (2014). The practice of going helps children to stop: the importance of context monitoring in inhibitory control. *J. Exp. Psychol. Gen.* 143:959. doi: 10.1037/a0035868

Chooi, W. T. (2011). *Improving intelligence by increasing working memory capacity* (Electronic Thesis or Dissertation). Case Western Reserve University, Cleveland, OH. Retrieved from: https://etd.ohiolink.edu/

Cox, B. D. (1994). Children's use of mnemonic strategies: variability in response to metamemory training. *J. Genet. Psychol.* 155, 423–442. doi: 10.1080/00221325.1994.9914792

Dahlin, K. I. E. (2011). Effects of working memory training on reading in children with special needs. *Read. Writ.* 24, 479–491. doi: 10.1007/s11145-010-9238-y

De Simoni, C., and von Bastian, C. C. (2018). Working memory updating and binding training: Bayesian evidence supporting the absence of transfer. *J. Exp. Psychol. Gen.* 147:829. doi: 10.1037/xge0000453

de Vries, M., Verdam, M. G., Prins, P. J., Schmand, B. A., and Geurts, H. M. (2018). Exploring possible predictors and moderators of an executive function training for children with an autism spectrum disorder. *Autism* 22, 440–449. doi: 10.1177/1362361316682622

Diamond, A. (2012). Activities and programs that improve children's executive functions. *Curr. Dir. Psychol. Sci.* 21, 335–341. doi: 10.1177/0963721412453722

Diamond, A., and Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science* 333, 959–964. doi: 10.1126/science.1204529

Egger, M., Smith, G. D., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj* 315, 629–634. doi: 10.1136/bmj.315.7109.629

Fellman, D., Salmi, J., Ritakallio, L., Ellfolk, U., Rinne, J. O., and Laine, M. (2018). Training working memory updating in Parkinson's disease: a randomised controlled trial. *Neuropsychol. Rehabil.* 30, 673–708. doi: 10.1080/09602011.2018.1489860

Fernández-Castilla, B., Declercq, L., Jamshidi, L., Beretvas, N., Onghena, P., and Van den Noortgate, W. (2020). Visual representations of meta-analyses of multiple outcomes: extensions to forest plots, funnel plots, and caterpillar plots. *Methodology* 16, 299–315. doi: 10.5964/meth.4013

Foster, J. L., Harrison, T. L., Hicks, K. L., Draheim, C., Redick, T. S., and Engle, R. W. (2017). Do the effects of working memory training depend on baseline ability level? *J. Exp. Psychol. Learn. Mem. Cogn.* 43, 1677–1689. doi: 10.1037/xlm0000426

Fu, L., Maes, J. H., Kessels, R. P., and Daselaar, S. M. (2017). To boost or to CRUNCH? Effect of effortful encoding on episodic memory in older adults is dependent on executive functioning. *PLoS ONE* 12:e0174217. doi: 10.1371/journal.pone.0174217

Gade, M., Zoelch, C., and Seitz-Stein, K. (2017). Training of visual-spatial working memory in preschool children. *Adv. Cogn. Psychol.* 13:177. doi: 10.5709/acp-0217-7

Gaultney, J. F., Bjorklund, D. F., and Goldstein, D. (1996). To be young, gifted, and strategic: advantages for memory performance. *J. Exp. Child Psychol.* 61, 43–66.

Gunn, R. L., Gerst, K. R., Wiemers, E. A., Redick, T. S., and Finn, P. R. (2018). Predictors of effective working memory training in individuals with alcohol use disorders. *Alcohol. Clin. Exp. Res.* 42, 2432–2441. doi: 10.1111/acer.13892

Guye, S., De Simoni, C., and von Bastian, C. C. (2017). Do individual differences predict change in cognitive training performance? A latent growth curve modeling approach. *J. Cogn. Enhanc.* 1, 374–393. doi: 10.1007/s41465-017-0049-9

Hering, A., Meuleman, B., Bürki, C., Borella, E., and Kliegel, M. (2017). Improving older adults' working memory: the influence of age and crystallized intelligence on training outcomes. *J. Cogn. Enhanc.* 1, 358–373. doi: 10.1007/s41465-017-0041-4

Hickey, A. (2018). *Improving academic outcomes for children in foster care through tutoring or working memory training: three randomized trials* (Electronic Thesis or Dissertation). University of Ottawa, Ottawa, ON, Canada. doi: 10.20381/ruor-21886

Higgins, J. P. T., and Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Stat. Med.* 21, 1539–1558. doi: 10.1002/sim.1186

Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J., and Popovich, D. L. (2015). Toward a more nuanced understanding of the statistical properties of a median split. *J. Consum. Psychol.* 25, 652–665. doi: 10.1016/j.jcps.2014.12.002

Jaeggi, S. M., Buschkuehl, M., Jonides, J., and Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6829–6833. doi: 10.1073/pnas.0801268105

Jaeggi, S. M., Buschkuehl, M., Shah, P., and Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Mem. Cogn.* 42, 464–480. doi: 10.3758/s13421-013-0364-z

Jones, M. R., Katz, B., Buschkuehl, M., Jaeggi, S. M., and Shah, P. (2020). Exploring N-Back cognitive training for children with ADHD. *J. Atten. Disord.* 24, 704–719. doi: 10.1177/1087054718779230

Karbach, J., Könen, T., and Spengler, M. (2017). Who benefits the most? Individual differences in the transfer of executive control training across the lifespan. *J. Cogn. Enhanc.* 1, 394–405. doi: 10.1007/s41465-017-0054-z

Karbach, J., Strobach, T., and Schubert, T. (2015). Adaptive working-memory training benefits reading, but not mathematics in middle childhood. *Child Neuropsychol.* 21, 285–301. doi: 10.1080/09297049.2014.899336

Karbach, J., and Unger, K. (2014). Executive control training from middle childhood to adolescence. *Front. Psychol.* 5:390. doi: 10.3389/fpsyg.2014.00390

Katz, B., Jaeggi, S., Buschkuehl, M., Stegman, A., and Shah, P. (2014). Differential effect of motivational features on training improvements in school-based cognitive training. *Front. Hum. Neurosci.* 8:242. doi: 10.3389/fnhum.2014.00242

Kliegl, R., Smith, J., and Baltes, P. B. (1990). On the locus and process of magnification of age differences during mnemonic training. *Dev. Psychol.* 26:894–904. doi: 10.1037/0012-1649.26.6.894

Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Res. Synth. Methods* 2, 61–76. doi: 10.1002/jrsm.35

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., et al. (2009). The PRISMA statement for reporting

systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J. Clin. Epidemiol.* 62, e1–e34. doi: 10.1016/j.jclinepi.2009.06.006

Lipsey, M. W., and Wilson, D. B. (2001). *Applied social research methods series, Vol. 49. Practical Meta-Analysis.* Thousand Oaks, CA: Sage Publications, Inc.

Loosli, S. V., Buschkuehl, M., Perrig, W. J., and Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychol.* 18, 62–78. doi: 10.1080/09297049.2011.575772

Lövdén, M., Brehmer, Y., Li, S.-C., and Lindenberger, U. (2012). Training-induced compensation versus magnification of individual differences in memory performance. *Front. Hum. Neurosci.* 6:141. doi: 10.3389/fnhum.2012.00141

Lustig, C., Shah, P., Seidler, R., and Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: a review and future directions. *Neuropsychol. Rev.* 19, 504–522. doi: 10.1007/s11065-009-9119-9

McKitrick, L. A., Friedman, L. F., Brooks, J. O., Pearman, A., Kraemer, H. C., and Yesavage, J. A. (1999). Predicting response of older adults to mnemonic training: who will benefit? *Int. Psychogeriatr.* 11, 289–300. doi: 10.1017/S1041610299005852

Melby-Lervåg, M., and Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Dev. Psychol.* 49:270. doi: 10.1037/a0028228

Melby-Lervåg, M., Redick, T. S., and 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. *Perspect. Psychol. Sci.* 11, 512–534. doi: 10.1177/1745691616635612

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Nakagawa, S., and Santos, E. S. (2012). Methodological issues and advances in biological meta-analysis. *Evol. Ecol.* 26, 1253–1274. doi: 10.1007/s10682-012-9555-5

Noack, H., Lövdén, M., and Schmiedek, F. (2014). On the validity and generality of transfer effects in cognitive training research. *Psychol. Res.* 78, 773–789. doi: 10.1007/s00426-014-0564-6

Nunes, E. V., Pavlicova, M., Hu, M.-C., Campbell, A. N., Miele, G., Hien, D., et al. (2011). Baseline matters: the importance of covariation for baseline severity in the analysis of clinical trials. *Am. J. Drug Alcohol Abuse* 37, 446–452. doi: 10.3109/00952990.2011.596980

O'Brien, J. L., Edwards, J. D., Maxfield, N. D., Peronto, C. L., Williams, V. A., and Lister, J. J. (2013). Cognitive training and selective attention in the aging brain: an electrophysiological study. *Clin. Neurophysiol.* 124, 2198–2208. doi: 10.1016/j.clinph.2013.05.012

R Core Team (2020). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/

Rossignoli-Palomeque, T., Perez-Hernandez, E., and González-Marqués, J. (2018). Brain training in children and adolescents: is it scientifically valid? *Front. Psychol.* 9:565. doi: 10.3389/fpsyg.2018.00565

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., and Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14931–14936. doi: 10.1073/pnas.0506897102

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

Segretin, M. S., Lipina, S. J., Hermida, M. J., Sheffield, T. D., Nelson, J. M., Espy, K. A., et al. (2014). Predictors of cognitive enhancement after training in preschoolers from diverse socioeconomic backgrounds. *Front. Psychol.* 5:205. doi: 10.3389/fpsyg.2014.00205

Shipstead, Z., Redick, T. S., and Engle, R. W. (2012). Is working memory training effective? *Psychol. Bull.* 138, 628–654. doi: 10.1037/a0027473

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., et al. (2016). Do "brain-training" programs work? *Psychol. Sci. Public Interest* 17, 103–186. doi: 10.1177/1529100616661983

Singer, T., Lindenberger, U., and Baltes, P. B. (2003). Plasticity of memory for new learning in very old age: a story of major loss? *Psychol. Aging* 18:306. doi: 10.1037/0882-7974.18.2.306

Smoleń, T., Jastrzebski, J., Estrada, E., and Chuderski, A. (2018). Most evidence for the compensation account of cognitive training is unreliable. *Mem. Cogn.* 46, 1315–1330. doi: 10.3758/s13421-018-0839-z

Snow, R. E. (1989). "Aptitude-treatment interaction as a framework for research on individual differences in learning," in *A Series of Books in Psychology. Learning and Individual Differences: Advances in Theory and Research*, eds P. L. Ackerman, R. J. Sternberg, and R. Glaser (New York, NY: W. H. Freeman).

Stepankova, H., Lukavsky, J., Buschkuehl, M., Kopecek, M., Ripova, D., and Jaeggi, S. M. (2014). The malleability of working memory and visuospatial skills: a randomized controlled study in older adults. *Dev. Psychol.* 50, 1049–1059. doi: 10.1037/a0034913

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., et al. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *Bmj* 343:d4002. doi: 10.1136/bmj.d4002

Strobach, T., and Huestegge, L. (2017). Evaluating the effectiveness of commercial brain game training with working-memory tasks. *J. Cogn. Enhanc.* 1, 539–558. doi: 10.1007/s41465-017-0053-0

Thompson, T. W., Waskom, M. L., Garel, K.-L. A., Cardenas-Iniguez, C., Reynolds, G. O., Winter, R., et al. (2013). Failure of working memory training to enhance cognition or intelligence. *PLoS ONE* 8:e63614. doi: 10.1371/journal.pone.0063614

Tidwell, J. W., Dougherty, M. R., Chrabaszcz, J. R., Thomas, R. P., and Mendoza, J. L. (2014). What counts as evidence for working memory training? Problems with correlated gains and dichotomization. *Psychon. Bull. Rev.* 21, 620–628. doi: 10.3758/s13423-013-0560-7

Tulving, E. (1972). *Episodic and Semantic Memory [w:] E. Tulving, W. Donaldson (Red.), Organisation of Memory (s. 381–403).* New York, NY: Academic Press.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., and Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behav. Res. Methods* 45, 576–594. doi: 10.3758/s13428-012-0261-6

van der Donk, M. L., van Viersen, S., Hiemstra-Beernink, A.-C., Tjeenk-Kalff, A. C., van der Leij, A., and Lindauer, R. J. (2017). Individual differences in training gains and transfer measures: an investigation of training curves in children with attention-deficit/hyperactivity disorder. *Appl. Cogn. Psychol.* 31, 302–314. doi: 10.1002/acp.3327

Verhaeghen, P., Marcoen, A., and Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: a meta-analytic study. *Psychol. Aging* 7, 242–251. doi: 10.1037//0882-7974.7.2.242

Vermeij, A., Kessels, R. P. C., Heskamp, L., Simons, E. M. F., Dautzenberg, P. L. J., and Claassen, J. A. H. R. (2017). Prefrontal activation may predict working-memory training gain in normal aging and mild cognitive impairment. *Brain Imaging Behav.* 11, 141–154. doi: 10.1007/s11682-016-9508-7

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 36, 1–48. doi: 10.18637/jss.v036.i03

Volckaert, A. M., and Pascale Noel, M. (2017). Externalizing behavior problems in preschoolers: impact of an inhibition training. *J. Psychol. Abnorm.* 5, 1–11. doi: 10.4172/2471-9900.1000154

Volckaert, A. M. S., and Noël, M.-P. (2015). Training executive function in preschoolers reduce externalizing behaviors. *Trends Neurosci. Educ.* 4, 37–47. doi: 10.1016/j.tine.2015.02.001

von Bastian, C. C., and Oberauer, K. (2013). Distinct transfer effects of training different facets of working memory capacity. *J. Mem. Lang.* 69, 36–58. doi: 10.1016/j.jml.2013.02.002

von Bastian, C. C., and Oberauer, K. (2014). Effects and mechanisms of working memory training: a review. *Psychol. Res.* 78, 803–820. doi: 10.1007/s00426-013-0524-6

Vygotsky, L. (1978). "Interaction between learning and development," in *Mind in Society: The development of higher psychological processes*, eds M. Cole, V. John-Steiner, S. Scribner, and E. Souberman (Cambridge: Harvard University Press).

Weicker, J., Hudl, N., Frisch, S., Lepsien, J., Mueller, K., Villringer, A., et al. (2018). WOME: theory-based working memory training — A placebo-controlled, double-blind Evaluation in older adults. *Front. Aging Neurosci.* 10:247. doi: 10.3389/fnagi.2018.00247

Wiemers, E. A., Redick, T. S., and Morrison, A. B. (2019). The influence of individual differences in cognitive ability on working memory training gains. *J. Cogn. Enhanc.* 3, 174–185. doi: 10.1007/s41465-018-0111-2

Zinke, K., Zeintl, M., Eschen, A., Herzog, C., and Kliegel, M. (2012). Potentials and limits of plasticity induced by working memory training in old-old age. *Gerontology* 58, 79–87. doi: 10.1159/000324240

Zinke, K., Zeintl, M., Rose, N. S., Putzmann, J., Pydde, A., and Kliegel, M. (2014). Working memory training and transfer in older adults: effects of age, baseline performance, and training gains. *Dev. Psychol.* 50, 304–315. doi: 10.1037/a0032982