Immediate and Long-Term Efficacy of Executive Functions Cognitive Training in Older
Adults: A Systematic Review and Meta-Analysis

Lan Nguyen¹, Karen Murphy¹, Glenda Andrews^{1,2}

Griffith University

¹School of Applied Psychology, Griffith University, Gold Coast, QLD, 4222, Australia

²Menzies Health Institute, Queensland, Australia

Author's Note

Corresponding author: Lan Nguyen, School of Applied Psychology, Griffith University, Gold Coast Campus, Southport, QLD, 4222, Australia. E-mail: lan.nguyen3@griffithuni.edu.au

This work was supported by the Australian Government Research Training Program Scholarship.

©American Psychological Association, 2019. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. Please do not copy or cite without author's permission. The final article and supplementary materials is available, upon publication, at: <http://dx.doi.org/10.1037/bul0000196>

Please cite this article as:

Nguyen, L., Murphy, K., & Andrews, G. (2019). Immediate and long-term efficacy of executive functions cognitive training in older adults: A systematic review and meta-analysis. *Psychological Bulletin*. doi:10.1037/bul0000196

Abstract

There has been growing interest in enhancing cognition in older adulthood via computerized cognitive training (CCT), though, there is controversy surrounding the efficacy of CCT in promoting improvements to functional everyday activities. As core executive-functions (EFs)—cognitive-flexibility, inhibition, working memory—are applicable to most aspects of daily living, CCT targeting these processes would likely promote gains on trained tasks, and potentially on similar untrained tasks (near-transfer), and general cognitive performance (far-transfer). We report two meta-analyses investigating the immediate (pre-test to post-test) and long-term efficacy (pre-test to follow-up) of core-EF CCT in improving cognition amongst older adults. Sixty-four studies (encompassing 3,594 participants) included an eligible CCT intervention targeting at least one core-EF (e.g., working memory training). Both immediate and long-term efficacy analyses revealed significant, large training effects for trained outcomes, and significant, small training effects for near-transfer and far-transfer outcomes. When comparing the same studies, effect sizes from immediate and long-term efficacy analyses were comparable, suggesting that CCT gains were maintained over time. Further analyses of immediate efficacy revealed significant, small training effects for performance on executive functioning, fluid intelligence, memory, and visuospatial domains, but not for attention or processing speed. After adjusting for publication bias, the training effect for fluid intelligence was nonsignificant, whereas processing speed was significant. It is recommended that future studies employ adaptive multidomain training as these studies were shown to produce significant training effects at each transfer level. Overall, core-EF CCT interventions show promise in promoting immediate and long-term improvements in cognitive performance amongst older adults.

Keywords: cognitive training, executive functions, older adults, cognitive aging, meta-analysis

Significance Statement

This meta-analytic investigation suggests that cognitive training targeting multiple core executive functions (working memory, inhibition, cognitive flexibility) promotes cognitive enhancement in older adulthood. Improvements in cognitive performance are shown not only on trained tasks, but on many untrained tasks as well (near- and far-transfer), though these transfer effects are small and limited to select cognitive domains. These improvements in cognitive performance appear to be maintained over time.

Immediate and Long-Term Efficacy of Executive Functions Cognitive Training in Older Adults: A Systematic Review and Meta-Analysis

With the rapid global growth of the older population, higher rates of age-related cognitive decline and more severe pathological decline such as mild cognitive impairment (MCI) and dementia are anticipated. Approximately 50 million people currently suffer from dementia worldwide, and the World Health Organization (2017) predicts that this number will increase threefold by 2050. This will impact aspects of everyday living such as quality of life (e.g., loss of autonomy for those suffering from cognitive decline and their carers) and economic spending at both individual and societal levels (e.g., increased individual and government spending on age-related health care and long-term aged care; Kendig, McDonald, & Piggott, 2016). Considering these projections for the aging population, there has been an urgent call for interventions which might delay or attenuate cognitive decline.

In response, there has been a surge in the number of studies investigating the efficacy of cognitive interventions for older adults—computerized cognitive training (CCT) interventions in particular. CCT requires participants to frequently and repeatedly engage in cognitively challenging tasks targeting specific cognitive domains (e.g., working memory, attention, processing speed) on a technological device such as a computer or tablet. The value of cognitive training is rooted in the premise that frequent, repeated practice on a task improves performance on the same, and similar tasks (Thorndike, 1906). Thus, if proficiency on these cognitive tasks predicts real-world abilities and success, and if that success is contingent upon cognitive abilities, then repeated practice on these tasks should improve both task outcomes and performance in everyday activities (Simons et al., 2016). The topic of controversy, however, is whether cognitive training interventions can indeed promote general learning, and in extension, enhance general cognitive functioning and real-world abilities (Oei & Patterson, 2014), or whether it merely induces task-specific learning—a pattern of

findings that Green and Bavelier (2008) term the ‘curse of specificity.’

There has been mixed empirical evidence concerning the efficacy of cognitive training. Many researchers have found that CCT is effective in improving cognition amongst healthy older adults (e.g., Anguera et al., 2013; Heinzl et al., 2014; Karbach & Kray, 2009). It has also been suggested that implementing such interventions during the prodromal stage of MCI or dementia might prevent or delay the progress of pathological cognitive decline (Vermeij, Claassen, Dautzenberg, & Kessels, 2016). However, findings from other studies suggest that, although CCT improves performance on trained tasks, these gains do not generalize or ‘transfer’ to untrained tasks (e.g., Bellander et al., 2016; Hering, Meuleman, Bürki, Borella & Kliegel, 2017; Lange & Süß, 2015; Souders et al., 2017).

Despite numerous investigations into the efficacy of CCT, there is still heated debate regarding whether CCT is effective in improving cognition, and which training conditions facilitate and produce the largest training and transfer effects. Meta-analyses are therefore extremely useful in this field of research because they use statistical tests to synthesize information from many related studies (e.g., CCT studies) to address mixed empirical findings. Meta-analytic techniques also allow for the identification of moderating variables which can offer insight into whether any particular study characteristics or training conditions appear to be more effective in improving cognitive performance. Ironically, meta-analyses have also reported mixed findings regarding the generalizability of cognitive training (e.g., Au et al., 2015; Au, Buschkuehl, Duncan, & Jaeggi, 2016; Karbach & Verhaeghen, 2014; Melby-Lervåg & Hulme, 2013, 2016). However, some meta-analyses that have reported positive transfer to untrained tasks have received criticism regarding methodology (for a detailed discussion, see Melby-Lervåg & Hulme, 2016). The purpose of this meta-analytic investigation was to therefore build upon, and improve the methodology of, previous meta-analyses (discussed further in the ‘*Previous Syntheses of Cognitive Training Research*

Targeting Older Adults' section) in order to shed light on the efficacy of CCT interventions in the older adult population, and identify potential study characteristics that appear to be effective in producing broad transfer effects.

Elements of Computerized Cognitive Training Interventions

The typical CCT study implements a pre-test–intervention–post-test design and includes at least one training condition and one control condition. An effect of training or ‘training effect’ is established where the trained group shows a significantly larger improvement on an outcome measure (better performance at post-test than at pre-test) compared to the control group (Schmiedek, 2016). The efficacy of a particular CCT intervention is typically evaluated in terms of whether training effects were observed for: a) trained outcomes, b) near-transfer outcomes (untrained tasks similar to, but measuring the same construct as, the trained task), and c) far-transfer outcomes (tasks measuring a different cognitive construct; Hogrefe, Studer-Luethi, Kodzhabashev, & Perrig, 2017; Karbach & Verhaeghen, 2014; Könen, Strobach, & Karbach, 2016; Shipstead, Redick, & Engle, 2010).

Controversies and inconsistencies in findings within the CCT literature are likely due to differences between studies in design and methodology. Aspects which have differed across CCT interventions include characteristics pertaining to the intervention such as the type of training administered (e.g., focal cognitive domains being trained, whether training is adaptive, whether training is process- or strategy-based), and the training schedule (e.g., length of training sessions, total training sessions, total training duration, number of sessions administered per week). Variations of these characteristics can affect the outcome of a study. For example, adaptive training (training which continually adjusts task difficulty based on participants’ performance level) has been reported to be more beneficial than non-adaptive training in terms of training gains and neural efficiency (e.g., Brehmer et al., 2011; Brehmer, Westerberg, & Bäckman, 2012; Cuenen et al., 2016).

Characteristics pertaining to study design also affect study outcomes, such as type of control condition(s) included (active, passive, or both). Participants in passive control conditions only take part in the testing phases of a study. Researchers who only include a passive condition risk overestimating training and transfer effects as non-focal sources of improvement (e.g., expectancy, motivation, general cognitive stimulation, social interaction) cannot be controlled (Guye, Röcke, Mérillat, von Bastian, & Martin, 2016). In contrast, participants within active control conditions engage in activities or ‘training’ that produces the same non-focal effects but does not target the focal cognitive domain of interest (Schmiedek, 2016). This type of control is preferred as it allows for a more accurate estimate of training and transfer effects. Other variable study design characteristics include training location, randomization, sample size, outcome measures, and researchers’ operationalization of ‘transfer distance.’

Transfer distance refers to the taxonomy describing similarities in the nature of, and processes that underlie, outcome measures compared to the trained task(s) (Noack, Lövdén, Schmiedek, & Lindenberger, 2009). The majority of training studies typically categorize outcome measures in terms of trained outcomes (or criterion tasks), near-transfer outcomes, or far-transfer outcomes, though, this classification has been defined and implemented differently across studies. For example, in terms of working memory training, some researchers define near-transfer as untrained measures strictly from the same trained domain (i.e., untrained working memory measures; e.g., Goghari & Lawlor-Savage, 2017; Könen et al., 2016; Li et al., 2008), whereas other researchers define near-transfer as untrained measures tapping the same or similar cognitive processes to the trained domain (e.g., working memory and short-term memory measures; Borella, Carretti, Riboldi, & De Beni, 2010; Schwaighofer, Fischer, & Bühner, 2015). Such differences in classification of transfer tasks

complicate the comparison of findings across studies.¹

Whilst it is clear that there are numerous characteristics that vary across cognitive training studies, the ultimate goal of these interventions is to promote far-transfer effects—i.e., apply training gains to a broad range of (untrained) everyday functional activities. With the various design characteristics used in the large number of cognitive training studies conducted to date, there is a clear need for studies that focus on synthesizing and summarizing findings across the literature (e.g., meta-analytic investigations). Information from such studies would be valuable to cognitive aging researchers and provide guidance for future research regarding the optimal training conditions and study characteristics that are likely to produce large training and transfer effects.

How Does Cognitive Training Work?

It has long since been acknowledged that repeated practice on any task improves performance on that same task, however, there are many different accounts of the cognitive and neural mechanisms underlying this training effect, particularly throughout adult development. One explanation emphasizes that, by frequently practicing a cognitive task, one will become quicker and more accurate on that task as the underlying cognitive and neural processes become more automatic, allowing for more efficient processing (e.g., Jonides, 2004).

Another account—more relevant to aging—implicates the brain's role in reorganizing neural resources. In their Scaffolding Theory of Aging and Cognition (STAC), Park and Reuter-Lorenz (2009) posit that learning and skill acquisition promotes compensatory 'scaffolding' in the aging brain, that is, the recruitment of additional neural circuitry to supplement, complement, or provide an alternative way to achieve a particular cognitive goal

¹ For example, two working memory training studies might find improvements to untrained working memory measures and measures closely related to working memory, such as short-term memory. Although both studies observed the same results, Study 1 might report that they found near-transfer effects and far-transfer effects, whereas Study 2 might only report finding near-transfer effects.

or task. This scaffolding occurs prominently within the prefrontal cortex—specifically, within the inferior, lateral, and rostral regions—and is an adaptive response to deficits within other areas of the brain such as the ventral visual cortex, hippocampus, and more posterior regions. The authors suggest that training establishes scaffolds to enhance performance of tasks that were previously carried out inefficiently. However, they also acknowledge that, where older adults already rely on scaffolds (evidenced by overactivation of neural resources), training may help lower reliance on scaffolding (reduce activation in scaffolded areas), and improve efficiency of honed networks. In their revised model (STAC-r), the authors additionally consider life course factors that serve to enhance or deplete neural resources, and incorporate the possibility that cognitive interventions may directly influence brain structure and function (Reuter-Lorenz & Park, 2014).

On a similar note, Lövdén, Bäckman, Lindenberger, Schaefer, and Schmiedek (2010) emphasize plasticity as a mechanism underlying training effects. In their framework of adult cognitive plasticity, the authors distinguish flexibility from plasticity, where flexibility denotes an individual's capacity to optimize performance using existing neural resources, and plasticity as the capacity for changes in flexibility. For plasticity to occur, there must be a prolonged 'supply-demand mismatch,' that is, a discrepancy between functional supply (current range of flexibility, or current ability) and environmental demands (e.g., task difficulty). Thus, when learning a new task, there is an initial mismatch between functional supply and environmental demands; through repeated practice, the supply gradually improves to meet the demands, and eventually when they are matched, cognitive performance remains stable as there is no room for further improvement.

Although enhanced performance on trained tasks are well established, this does not necessarily imply enhancements in other cognitive functions and everyday abilities (e.g., Noack et al., 2009; Simons et al., 2016). To address this matter, cognitive researchers

typically assess performance on a variety of transfer tasks in addition to the training task to measure the extent to which the training improvements are generalizable. However, to date, relatively little is known about the mechanisms underlying transfer (Katz, Shah, & Meyer, 2018; Lustig, Shah, Seidler, & Reuter-Lorenz, 2009; Wolf et al., 2018).

One of the earliest doctrines of training and transfer stems from Thorndike (1906) who proposed the *Identical Elements Theory*, whereby improvements on one cognitive process transfers to another, only if the two cognitive processes share identical features. This foundational concept paved the way for modern frameworks of transfer, including Taatgen's (2013, 2016) *Primitive Elements (PRIMs) Theory*. The overarching premise of this theory posits that the learning of general cognitive skills is a by-product of learning specific cognitive skills. Central to this framework are primitive information processing elements (the smallest possible building blocks of information)—some of which are task-specific, whilst others are task-general. After repeated practice on a task, task-specific elements can be combined which leads to more efficient task performance, whereas task-general elements can be reused for other tasks and skills. Thus, the PRIMs model predicts that far-transfer can occur even if the knowledge and skills gained from training on a certain cognitive task or domain is not directly applicable to a different task or domain as task-general elements can be reused. In particular, Taatgen suggests that training programs which promote proactive over reactive strategies may more successfully produce far-transfer. Whereas reactive strategies are driven by experimental stimuli, proactive strategies are driven by internal task-related processing or preparation. Using the Stroop task as an example, the reactive strategy involves waiting for the stimulus, attending to the whole stimulus (both word and ink color), then naming the ink color. The proactive strategy involves preparing for the upcoming stimulus which consequently alters the 'default' attending step and replacing it with a more efficient step which only focuses on the ink color, thereby reducing interference from the 'Stroop

effect.’ The PRIMs model suggests that proactive strategies are more complex than reactive strategies in that they require additional processing elements to be learnt. However, proactive strategies appear to be task-general and the learning of such strategies likely promotes transfer (for a modelled simulation of the PRIMs model, see Taatgen, 2013).

Building upon these premises, researchers have investigated the neural mechanisms underlying training and transfer. One highly influential account posited that transfer is more likely to occur when the training and transfer tasks share related information processing components and engage similar neural circuitry (e.g., Jonides, 2004; Karbach & Kray, 2016; Lövdén et al., 2010). Dahlin, Stigsdotter Neely, Larsson, Bäckman, and Nyberg (2008) were the first to provide both behavioral and neuroimaging evidence to support this overlap principle of training and transfer. After training young adults on a working memory updating task (letter memory) for five weeks, the researchers found that the trained group (compared to the control group) demonstrated transfer to another updating task (n -back), but not to a task requiring inhibitory control (Stroop task). Their imaging data indicated that, at pre-test, only the letter memory and n -back tasks shared overlapping striatal activation, whereas the Stroop task did not. Additionally, an increase in activation was observed in this same region for both updating tasks as a function of training. When conducting the same experiment amongst older adults, however, no transfer effects were observed. Their neuroimaging data showed no overlapping pre-test striatal activation between training and transfer tasks, nor any common training-related neural changes from pre-test to post-test. The authors concluded that a prerequisite for transfer is that the training and transfer tasks engage overlapping brain regions (at pre-test), and that the response to training (neural changes) in these regions is similar for both tasks. In other words, gains in cognitive performance are likely to transfer to other untrained tasks, only if training and transfer tasks activate and show the same pattern of alterations in the same brain regions.

This notion has also been supported by recent training studies employing neuroimaging techniques. Heinzl et al. (2016) employed *n*-back training in older adults, reporting significant, albeit limited, transfer to the Sternberg task (a different working memory task requiring updating and maintenance processes). They found overlapping bilateral frontoparietal activation in both training and transfer tasks at pre-test, and a common training-related decrease in neural activity within the medial frontal gyrus and caudal superior frontal sulcus for both tasks. Interestingly, no striatal activation was reported in this study which may explain the limited transfer effects observed (improvements in only 1 out of 4 conditions of the Sternberg task)—similar to the findings of the older participants in the Dahlin et al. (2008) study.

Salminen, Kühn, Frensch, and Schubert (2016) implemented a dual *n*-back training program to assess near-transfer to a similar, untrained updating task amongst young adults. Participants were assigned to one of three groups. The training group received dual *n*-back training (verbal and visuospatial *n*-back tasks presented simultaneously), the active control group received single *n*-back training (verbal and visuospatial *n*-back tasks presented consecutively), and the passive control group did not partake in any training activities. Although the authors found no significant transfer to the single updating transfer tasks (all groups showed similar improvement), only the training group showed significant transfer to the dual updating transfer task. Their neuroimaging data revealed that all groups showed neural overlap between the training and transfer dual task within the frontoparietal network and striatum at pre-test. However, at post-test, only the training group showed a reduction in overlapping activation within the frontal cortex between the training and transfer task. The authors suggested that this decrease in activity may have signified a more efficient use of relevant task processes. Additionally, compared to the two control groups who showed little or no change, the training group exhibited similar patterns of neural changes on the training

and transfer dual tasks from pre-test to post-test—specifically, decreased frontoparietal activation and increased striatal activation. As the training group only showed significant transfer on the dual *n*-back task and not the single *n*-back tasks, the authors suggested that, in addition to the prerequisites outlined by Dahlin et al. (2008), the training and transfer tasks must also share very similar processes for transfer to occur.

The Case for Training Core Executive Functions

Executive functions (EFs) refer to a group of top-down mental processes that are essential for performing everyday activities (Diamond, 2013). Three core-EFs are most often represented in the extant literature: *working memory* (WM; the ability to simultaneously hold and actively manipulate stimuli in mind before recalling the appropriate information), *inhibition* (the ability to suppress automatic responses or irrelevant stimuli), and *cognitive flexibility* (the ability to switch between multiple tasks, mental sets, or perspectives). Although these EFs are distinct constructs, they share underlying processes and are, therefore, not completely independent of one another (Karr et al., 2018; Miyake et al., 2000). To illustrate this point, Diamond (2013) explained that, to know which stimuli to inhibit, task goals must be held and updated in WM and to hold and update information in WM, irrelevant stimuli must be inhibited. Further, both inhibition and WM contribute to successful cognitive flexibility, such that, switching between tasks is facilitated when irrelevant information is inhibited and relevant information or goals are maintained and updated in WM (for more recent findings regarding the construct of executive functions, see Karr et al., 2018 and Miyake & Friedman, 2012).

These core-EFs form the foundations for higher-order cognitive functions (e.g., planning, reasoning, problem solving, decision making), and are heavily reliant on the prefrontal cortex (Otero & Barker, 2014)—the anterior cingulate cortex and dorsolateral prefrontal cortex, in particular (Goldstein, Naglieri, Princiotta, & Otero, 2014). The prefrontal

cortex, however, is particularly vulnerable to the aging process, being one of the first areas of the brain to deteriorate (Princiotta, DeVries, & Goldstein, 2014). Distinct regions within the prefrontal cortex project to a variety of cortical regions via white matter pathways, however, these pathways weaken and become ‘disconnected’ with increasing age (Yuan & Raz, 2014). The strength of these pathways has been related to performance on tasks of executive functioning, whereby poorer integrity of white matter pathways is associated with worse performance (Madden, Bennett, & Song, 2009). Accordingly, other cognitive processes reliant upon EFs also show early decline (Reuter-Lorenz, Festini, & Jantz, 2016). Therefore, training EFs should be beneficial in improving cognition amongst older adults by strengthening the pathways in and around the prefrontal cortex, through mechanisms such as scaffolding (Park & Reuter-Lorenz, 2009) or plasticity (Lövdén et al., 2010).

Furthermore, core-EFs have been reported to be critical in the execution of a number of cognitive processes which are also dependent upon the prefrontal cortex. For instance, prospective memory—the ability to remember intentions and planned activities—is a process critical for functional independence in older adulthood (e.g., remembering to take medications, remembering to pay bills). Neuroimaging studies have implicated the anterior prefrontal cortex in prospective memory performance (e.g., Simons, Schölvinck, Gilbert, Frith, & Burgess, 2006). Additionally, behavioral studies have shown that EFs including working memory and cognitive flexibility are critical in successful prospective memory performance (e.g., Kliegel, Martin, McDaniel, & Einstein, 2002; Martin, Kliegel, & McDaniel, 2003; Schnitzspahn, Stahl, Zeintl, Kaller, & Kliegel, 2013). Based on the overlap principle discussed previously, training core-EFs is likely to promote transfer to other cognitive processes such as prospective memory, as they seemingly engage overlapping processing mechanisms and brain regions.

As EFs are essential to almost all aspects of daily living (e.g., mental and physical

health, public safety, occupational and marital success; Diamond, 2013), declines in these processes have measurable everyday consequences. For example, lower scores on EF tasks have been linked to poorer driving performance (Adrian, Postal, Moessinger, Rasclé, & Charles, 2011) and poorer decision making (e.g., Brand, Recknor, Grabenhorst, & Bechara, 2007). There is clearly a strong case for cognitive training targeting core-EFs, as improving these processes would likely enhance multiple cognitive functions critical for functional independence in older adulthood.

Multidomain Executive Functions Training

Thorndike (1906) posited that “the most common and surest source of general improvement of a capacity is to train it in many particular connections” (p. 248). Applying this notion to CCT, multidomain training (training multiple cognitive domains) should exercise a wide range of cognitive processes, activate a large network of neural circuitry, and improve multiple different skills (Guye et al., 2016). Thus, as aging is associated with declines in numerous cognitive domains, taking a multidomain approach to cognitive training is therefore more likely to promote broader transfer than training a single cognitive domain alone. This notion has been supported by Binder et al. (2016) who trained older adults in either inhibition, visuomotor functions, spatial navigation, or all three processes (comprising three single-domain and one multidomain group). After training, the multidomain group demonstrated an increase in visuomotor function and spatial navigation—comparable to corresponding single-domain groups. However, only the multidomain group demonstrated a significant gain in far-transfer measures of attention as well.

Despite the promise of multidomain training, this type of intervention has received relatively little attention compared to single-domain training—WM training in particular—which has been the focus of many research studies, meta-analytic investigations, and reviews. The numerous meta-analyses that have examined WM training have produced varied

outcomes, with some authors concluding that training WM promotes far transfer in young and older adults (Au et al., 2015; Karbach & Verhaeghen, 2014) and even improvements to inattention in daily life amongst children and adults (Spencer-Smith & Klingberg, 2015), whereas others have found negligible far transfer effects in a range of different populations (Melby-Lervåg & Hulme, 2013, 2016; Schwaighofer et al., 2015). Thus, the purpose of the current study was not to reproduce yet another meta-analysis of WM training, but rather, to examine all types of EF training (i.e., CCT interventions that have included at least one EF component), and to further investigate whether multidomain-EF training is more effective in promoting transfer effects compared to single-domain EF training interventions. These findings would provide the foundations for future research to empirically examine whether single-domain or multidomain-EF training promotes larger training and transfer effects in older adults.

Previous Syntheses of Cognitive Training Research Targeting Older Adults

There have been a variety of different cognitive training interventions designed to improve cognition in older adulthood. One of the most popular types of interventions concerns video game training. Numerous meta-analyses have investigated the efficacy of video game training in improving cognition, though mixed findings have been reported. Toril, Reales, and Ballesteros (2014) explored whether video game training enhanced cognition in older adults. These authors reported small but significant, positive training effects on cognitive outcomes including speed, memory, attention, and overall cognitive functioning. In contrast, Sala, Tatlidil, and Gobet (2018) reported nonsignificant and negative effect sizes for video game training on older adults' cognitive performance. Similarly, Bediou et al. (2018) reported a nonsignificant, negative overall effect size for action video game training on cognition for older adults. These authors suggested that, as action video games are tailored to the younger population, training of this type may be too challenging for older

adults who, therefore, are not likely to benefit from action video game training. Findings from these meta-analytic studies suggest that training on some types of video games (e.g., action video games) might not be as effective in enhancing cognition in older adults compared to others (e.g., commercial ‘brain training’ games). In line with this notion, a meta-analysis conducted by Tetlow and Edwards (2017) reported that commercialized cognitive training programs promoted small to moderate transfer effects to numerous cognitive domains including attention, processing speed, visuospatial memory, and self-reported measures of everyday functioning amongst older adults.

The current study, however, did not focus solely on video game training, but examined all types of computerized cognitive interventions targeting core-EFs (including video game training and general cognitive training). The efficacy of general cognitive training in healthy older adults has been examined by a number of recent meta-analyses. Karbach and Verhaeghen (2014) examined the efficacy of EF training across 49 studies, though their meta-analysis did not include multidomain interventions. Further, the studies included in their analyses were not limited to CCT training, and some studies were not ‘training’ interventions, but rather, once-off practice sessions. These authors reported significant effect sizes on trained, near-transfer, and far-transfer outcomes, in favor of EF training. They also found no difference in effect sizes between WM training and other EF training interventions. It should be noted, however, that this meta-analysis has received criticism for i) including studies with no control group, ii) not accounting for group differences at pre-test (baseline), and iii) inadequately addressing the differences in training effects between studies utilizing active versus passive control groups (Melby-Lervåg & Hulme, 2016).

A recent meta-analysis conducted by Sala et al. (in press) focusing on WM training in older adults reported significant, large training effects for trained tasks and small training

effects for near-transfer tasks. However, significant far-transfer was only observed for studies comparing training effects to a passive control group; far-transfer effects for studies including an active control group were nonsignificant. Similarly, other WM training meta-analyses that have not specifically focused on older adults have reported similar findings, whereby far-transfer effects were largely observed for studies including passive control groups compared to those including active controls (e.g., Melby-Lervåg, Redick, & Hulme, 2016).

Another meta-analysis (Lampit, Hallock, & Valenzuela, 2014) examined the efficacy of CCT interventions across 52 studies (not limited to EF training, though multidomain interventions were included). The results of their analyses indicated that CCT produced significant, small training effects for memory, processing speed, working memory, visuospatial skills, and for all cognitive outcomes combined amongst health older adults. Transfer effects were not limited to studies including a passive control group. Studies including active control groups demonstrated significant, and larger transfer effects to verbal memory, non-verbal memory, WM, and visuospatial domains compared to studies including passive controls. However, studies including passive controls obtained significant training effects for processing speed and executive functions which were nonsignificant in studies including an active control group. It should be noted that the studies included in the Lampit et al. (2014) meta-analysis were restricted to randomized controlled trials. This likely eliminated numerous potential interventions as a large number of CCT studies implement pseudo-randomization methods in order to ensure that the intervention and control conditions are equivalent in terms of age, baseline intelligence, and cognitive status—factors which potentially influence the effect of cognitive training gains (e.g., Borella, Carbone, Pastore, De Beni, & Carretti, 2017).

Conclusions from these meta-analyses suggest that EF training and CCT training have the potential to produce small to moderate gains in cognitive performance amongst older

adults. However, it is not sufficient to examine cognitive training efficacy immediately after training. It is also important to examine the durability of training effects to determine whether continued enrichment is necessary and when this should be provided to promote the maintenance of functional independence in older adulthood. Meta-analyses in this area of research should therefore not be limited to examining the immediate efficacy of cognitive training, but should also explore whether any improvements in cognitive performance, as a function of training, persists over time. Accordingly, this study included an analysis examining long-term efficacy of CCT.

The Current Meta-Analytic Investigation

With the increasing interest in cognitive aging, the number of CCT studies continues to grow, warranting updated analyses that examine and incorporate more recent empirical evidence regarding the efficacy of CCT. The primary aim of this meta-analytic investigation was to investigate the immediate and long-term efficacy of core-EF CCT in improving cognitive performance amongst older adults. Specifically, it was of interest to examine whether EF training produced gains in near-transfer outcomes (untrained measures similar to the trained task from the same construct) and far-transfer outcomes (measures from different constructs).²

This study extended upon previous meta-analyses (Karchach & Verhaeghen, 2014; Lampit et al., 2014) by examining the efficacy of both single-domain and multidomain CCT targeting EFs, not limiting studies to randomized controlled trials, and addressing methodological shortcomings that have been identified previously (e.g., Melby-Lervåg & Hulme, 2016). This study also sought to explore whether any factors pertaining to the intervention or study design moderated the efficacy of core-EF CCT across the different levels of transfer—in particular, whether single-domain or multidomain EF training produced

² This classification of transfer distance was adopted in the meta-analysis conducted by Karchach and Verhaeghen (2014) and was therefore implemented in the current meta-analysis to maintain consistency and allow for easy interpretation and comparison of findings.

larger training and transfer effects. Other factors pertaining to the intervention, training schedule, and study design were also considered as potential moderators. Findings from the moderator analyses will provide a guide for future studies regarding potential training conditions and study characteristics that promote broad transfer to untrained cognitive domains.

Method

Procedures were conducted, and findings were reported in accordance with the Cochrane Handbook for Systematic Reviews of Interventions (Higgins & Green, 2011) and the Meta-Analysis Reporting Standards (MARS; American Psychological Association, 2008).

Eligibility Criteria

Types of studies. Articles examining the effects of a cognitive training intervention were considered for inclusion. Articles that reported follow-up data only were also included if the original training study met all eligibility criteria.

Studies must have been published in a peer-reviewed journal to ensure that the studies included in the analysis were of an adequate caliber. Studies must have comprised a pre-test–intervention–post-test design and included at least one intervention and one appropriate control condition (active or passive). Cross-sectional studies and studies without control conditions were excluded from the analysis.

An active control (AC) group was defined as any ‘training condition’ that aimed to produce the same non-focal effects as the intervention condition (e.g., motivation, social contact, enjoyment, cognitive stimulation) but did not comprise the focal cognitive component(s) being trained (i.e., EFs). A passive control (PC) group was defined as one that did not engage in a cognitive program or cognitive activities (e.g., waitlist, watching lectures or videos without answering questions about content). Control groups comprising any physical training components were deemed inappropriate for this meta-analysis as physical

activity training has generally been implemented as a separate intervention in itself. Reviews and meta-analyses of physical activity training suggest that physical exercise can improve cognitive functioning (e.g., Bherer, Erickson, & Liu-Ambrose, 2013; Colcombe & Kramer, 2003). Therefore, including control groups with physical activity components may confound the investigation for improvements in cognition as a result of cognitive training. Control conditions engaging in a simpler variation of the same cognitive training received by the training group (e.g., non-adaptive training, shorter training duration) were also excluded as both groups would have received the same domain EF training.

Types of participants. Studies must have included healthy older adults (mean sample age ≥ 60 years).³ Studies comprising older adults with functional or neurological impairments, MCI, or diagnoses of dementia were excluded unless the study also included intervention and control conditions comprising healthy older adults (in which case, only data for healthy older adults were examined). Samples comprising older adults with medical or physical conditions (e.g., hypertension, insomnia, balance impairment, hearing impairment) were also excluded due to variability in cognitive performance within these populations (e.g., Haimov & Shatil, 2013).

Types of training. Training must have comprised a process-based CCT intervention targeting at least one of the three core-EFs: cognitive flexibility, inhibition, or WM. Specifically, trained participants must have repeatedly practiced tasks via engagement with a technological device such as a computer, tablet, gaming console, or mobile device. Video game training interventions exercising at least one core-EF were also included. Studies utilizing pure pencil-and-paper cognitive training or a projector to administer stimuli as the

³Age cut-off varies across studies targeting older adults—some studies include participants over 50 years of age, and some only include participants over 60 or 65 years. Thus, the criterion concerning age in the current study (i.e., mean sample age of at least 60 years) was specified to be inclusive of studies with samples including older adults in the 50-year range. This is in accordance with similar reviews and meta-analyses that have specified the same criterion in terms of participant age (e.g., Karbach & Verhaeghen, 2014; Lampit et al., 2014; Reijnders, van Heugten, & van Boxtel, 2013).

main delivery method for the training were excluded.

At least three sessions must have been included in the intervention to ensure that participants had practiced on the training tasks.⁴ Further, the intervention must have had a relatively regular schedule—that is, participants should not have had the option to choose the duration and frequency of training sessions haphazardly throughout the intervention.⁵ Studies too lenient regarding adherence to training (i.e., did not have a regular schedule) were therefore excluded from the analysis.

Interventions that combined cognitive training with physical exercise training, brain stimulation, or pharmacological components were excluded due to potential confound effects. Where a study included a cognitive training only condition in addition to other combined training conditions (e.g., cognitive training + physical activity), only the pure cognitive training and control conditions were examined. As this meta-analytic investigation focused on examining the efficacy of CCT targeting core-EFs, interventions only included cognitive training components and no other.

Types of outcome measures. Only cognitive outcomes were examined in this meta-analysis—non-cognitive measures were not included (e.g., subjective questionnaires, physical outcomes, neural outcomes, etc.). Additionally, at least one untrained task must have been included as an outcome measure to examine the efficacy of cognitive training on cognition. Studies that examined trained outcomes only were excluded because while they yield conclusions about the magnitude of improvement on a task after repeated practice, they do not address cognitive transfer. Cognitive performance from pre- to post-test was analyzed, in addition to long-term outcomes (pre-test to follow-up) where available.

⁴There is no definitive rule regarding how many sessions constitutes “training.” However, some researchers have acknowledged that training interventions should comprise a minimum number of three sessions (e.g., Jaeggi & Buschkuhl, 2014; von Bastian & Oberauer, 2014). We therefore specified a minimum number of three training sessions for this criterion.

⁵If participants within a single study completed different amounts of training (e.g., Participant A opted to train for 1 hour, but Participant B trained for 24 hours), performance on outcome measures would be considerably variable, thereby producing unreliable findings pertaining to the efficacy of that particular intervention.

Search Methods for Identification of Studies

Electronic searches. Two comprehensive searches were conducted by one of the authors on two separate occasions to ensure that all eligible studies were included in the review. The first search was completed on 12th October 2017 and the second search on 7th January 2018. Results from the searches were similar as the same search terms were applied to the same databases (1,995 compared to 2,085). Differences in search results likely emerged from articles which were published after the time of the first search.

Key search terms included “cognitive training” OR “cognitive intervention” OR “computerized training” OR “game training” OR “executive function* training” OR “executive control training” OR “multi-domain training” AND “older adult” OR “aging”. The optimal set of key terms used to search each database was computed after running numerous different searches and refining keyword combinations. This procedure was carried out with the consultation of an academic librarian from the university.

Databases searched include Scopus, ProQuest, Science Direct, Ovid, and PubMed. Searches were limited to articles reported in English but not limited in terms of publication dates. Two additional studies were identified from Google Scholar Alerts after searching was completed.

Searching other resources. Additional studies were identified by searching through other resources such as reference books related to cognitive aging, and forward and backwards searching through reference lists of relevant articles.

Data Collection and Analysis

Selection of studies. Initial search results were screened for eligibility based upon the article title, abstract, and document type (reviews and conference or presentation abstracts were excluded). The full-text of each remaining article was then screened to be further assessed for eligibility. This screening process was conducted twice by one author after each

search (on two separate occasions). There were no discrepancies in the final set of eligible articles between the two searches. Two new articles which were published after the first search had been conducted and were also included in the final set of studies. A random subset of articles was provided to a second independent reviewer who was involved in screening 50% of title and abstracts, and 50% of full-text articles. In accordance with the guideline outlined by Orwin and Vevea (2009), there was excellent inter-rater reliability ($\kappa = .83$). The two reviewers resolved all disagreements after discussing the basis for inclusion or exclusion of each study. The eligibility of the final set of studies included in the review were assessed and agreed upon by the two reviewers. There was some overlap of studies included in the current study and previous meta-analytic studies (11 from Karbach & Verhaeghen, 2014; and 16 from Lampit et al., 2014).

Data extraction and management. A comprehensive record of data and study characteristics were coded using an electronic data collection form. A second independent coder was assigned a subset (50%) of eligible studies to code using the same electronic form. There was excellent inter-rater reliability ($\kappa = .86$). All outcome measures from each study were categorized into their respective cognitive domains by the two independent reviewers (100% agreement). The included studies employed measures from the following cognitive domains: cognitive flexibility, inhibition, WM, executive functioning, attention, fluid intelligence, language, memory, processing speed, and visuospatial ability. These domains were therefore examined in the meta-analyses (see Table 1 for additional information regarding cognitive domains and categorization of tasks). Outcome measures were also categorized by transfer distance: trained, near-transfer, and far-transfer.

[INSERT TABLE 1 ABOUT HERE]

Comprehensive Meta-Analysis Version 3.3 (CMA; Biostat, Englewood, NJ) was used to conduct the meta-analyses. Most studies reported change score data to examine whether

there were any changes in cognitive performance as a function of training group. These data were entered into CMA (pre-test and post-test means and standard deviations, and sample size of training and control groups). The correlations between pre-test and post-test measures were not reported in any of the included studies. Borenstein et al. (2009) suggests that when the correlation value is unknown, relevant studies in the field can be used to identify a plausible range for the correlation. Soveri, Antfolk, Karlsson, Salo, and Laine (2017) recently conducted a meta-analysis examining pre-post correlations within WM training studies. These authors reported similar pre-post correlations for training groups ($r = .66$), AC groups ($r = .65$), and PC groups ($r = .66$). Similarly, a meta-analysis conducted by Lampit et al. (2014)—examining CCT efficacy in healthy elderly—applied a pre-post correlation value of .60. In line with these meta-analyses, the current study imputed .60 as the pre-post correlation for all studies. Sensitivity analyses indicated that findings were robust to pre-post correlation values (see section on *Sensitivity Analyses*).

Where required, raw data were converted (e.g., standard errors or confidence intervals to standard deviations) using formulas provided in the Cochrane Handbook (Higgins & Deeks, 2011). In cases where study characteristics were unclear or sufficient data were not available from the full-text article or supplementary materials, original authors were contacted to request additional information. Nine out of the ten authors contacted provided additional data and/or clarification of methodology.

Most studies employed a variety of tasks to measure cognitive performance. Multiple effect sizes were therefore extracted from each study. To correct for this, aggregate effect sizes were calculated for each individual study.

Measures of training efficacy. Effect sizes were computed using Hedge's g which corrects for small sample sizes. A positive mean effect size indicated larger performance gains in the CCT condition compared to the control condition, whereas the opposite held for a

negative effect size. In accordance with Cohen (1988), an effect size of 0.2 was interpreted as a small effect, 0.5 as a moderate effect, and 0.8 as a large effect.

Studies with non-standard designs. Similar to previous meta-analyses (e.g., Lampit et al., 2014), where studies included both AC and PC groups, only data from AC groups were included. Including AC over PC groups provides more accurate effect size estimates for CCT efficacy (as AC groups control for non-focal training effects such as expectancy and social contact). Exceptions included AC groups which comprised a simpler variant of the intervention (as in Cuenen et al., 2016) or AC groups which incorporated physical training components (as in Gajewski & Falkenstein, 2012), wherein only data from the PC group were included. Sensitivity analyses were conducted to examine whether there were any differences in effect size estimates between the type of control groups included (see section on *Sensitivity Analyses*).

Where studies included multiple intervention conditions (provided that the interventions were similar and trained the same cognitive domains), data were pooled to form one intervention group. Pooled data were calculated using the formula provided by the Cochrane Handbook (Higgins & Deeks, 2011). Conversely, some studies that included multiple intervention conditions were split into separate studies. Borella et al. (2014) trained two older adult age-groups: young-old and old-old. Both age-groups comprised an intervention and control condition. Data from each age-group were therefore treated as two separate studies. Additionally, three studies each included two eligible intervention conditions targeting different core-EFs (Binder et al., 2016; Stepankova et al., 2014; Zając-Lamparska & Trempała, 2016). Intervention conditions were not pooled for these studies. Instead, for each paper, the two eligible intervention conditions were treated as two separate studies (e.g., Binder Study 1 – Multidomain, Binder Study 2 – Inhibition). These two independent comparisons were entered into CMA and the sample size of the control

condition was split between both comparisons (Higgins, Deeks, & Altman, 2011).

Lastly, four separate articles reported different outcome measures from the same participants and were therefore collated into one single study (Gajewski & Falkenstein, 2012, in press; Küper, Gajewski, Frieg, & Falkenstein, 2017; Wild-Wall, Falkenstein, & Gajewski, 2012). See Supplementary Materials for additional details regarding data extraction (pooling, splitting, and collating) for all included studies.

Assessment of bias and heterogeneity. Risk of bias was assessed individually for each study using the Cochrane Risk of Bias Tool (Higgins et al., 2011). The sources of bias or ‘domains’ were adapted to assess the methodology and design of a typical cognitive training study. These domains included: a) random sequence generation, b) blinding of participants, c) blinding of outcome assessment/assessor, d) incomplete outcome data, e) selective reporting, and f) other bias. The ‘allocation concealment’ domain from the original Risk of Bias Tool was omitted as this source of bias was not described by any of the eligible studies. That is, no studies “describe[d] the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment” (Higgins, Altman, & Sterne, 2011, p. 9). Instead, emphasis was placed on ‘blinding of participants and personnel’ to assess whether participants had knowledge of the intervention received.

Summary assessments of risk of bias were evaluated with respect to important outcomes within each study. A rating of low, unclear, or high risk of bias was appraised for each source of bias in each study. The summary rating for each study (high or low overall risk of bias) was evaluated on the basis of bias concerning blinding (of participants and outcome assessors—sources b and c, respectively). These two sources of bias were identified as ‘key domains’ in determining the summary risk of bias as participant or assessor knowledge of the intervention received may greatly affect performance during testing,

thereby affecting a study's findings and overall conclusions. Thus, a study was considered to have a low risk of bias if the blinding domains were evaluated as 'low risk.' A study was considered to have a high risk of bias if 'unclear risk' was rated for both blinding domains or 'high risk' was rated for either domain. Additionally, where at least four out of the six sources of bias were evaluated as 'high risk,' the study was also considered to have a high risk of bias as numerous methodological flaws might also greatly affect a study's findings and outcomes. Risk of bias evaluations were undertaken by two independent reviewers. All risk of bias evaluations were congruent between the reviewers. Refer to Supplementary Materials for risk evaluations for each source of bias in all included studies.

Publication bias refers to the issue that small and moderate-sized studies reporting significant findings or large effect sizes are more likely to be published than those obtaining null findings (Borenstein, Hedges, Higgins, & Rothstein, 2009). Consequently, such bias within the literature likely inflates the perceived efficacy of an intervention in question. However, due to the heated controversy surrounding the efficacy of cognitive training, it is likely that this field of research encompasses a more representative sample of findings and effect sizes. Even so, publication bias was assessed for each analysis.

The funnel plot was visually inspected for asymmetry, and Egger's test (Egger, Smith, Schneider, & Minder, 1997) was reported to support these visual evaluations. Within the funnel plots, studies are plotted in terms of effect size (x-axis) by standard error (y-axis); large studies or studies with high precision generally cluster around the mean effect size and appear toward the top of the graph, whereas smaller studies or studies with low precision tend to be spread out toward the bottom of the graph, forming the funnel pattern (Borenstein et al., 2009). Duval and Tweedie's (2000) *trim and fill* analysis provides the best estimate of an unbiased effect size by imputing missing studies to generate a symmetrically distributed funnel plot. This effect size adjusted for bias was also reported when trim and fill analyses

indicated potentially missing studies. Borenstein et al. (2009) suggests that, if the adjusted effect size remains largely unchanged, the impact of bias is trivial. Where the adjusted effect size shifts, but the key finding (i.e., the overall mean effect size) remains unchanged, this may reflect modest impact of bias. Lastly, where the adjusted effect size alters the key findings, this may reflect substantial impact of bias.

Each analysis was screened for potential outliers by visual inspection of the forest plot. Study effect sizes that appeared to deviate from the overall pattern were further inspected to determine whether the inclusion of the deviating effect sizes substantially influenced the results. Influential studies were identified using the one-study-removed function on CMA, which provides overall effect size estimates with each study removed. If the inclusion of a study made a considerable impact on the outcome of the overall mean effect size (i.e., altered the interpretation of the outcome by changing the size of the training effect), that study was excluded from the analysis.

The Q -statistic for heterogeneity was used to examine the variance among all study effect sizes (Card, 2015). This statistic tests the assumption that all studies share a common effect size, where a significant Q -test indicates that the true effect size varies substantially from study to study (Borenstein et al., 2009). The I^2 statistic was also reported to gauge the degree of heterogeneity between studies. I^2 expresses the proportion of 'true' or between-study variance (as opposed to sampling error, or within-study variance) to total variance (Deeks, Higgins, & Altman, 2011). In accordance with Higgins, Deeks, and Altman (2003), I^2 was interpreted as follows: 25% = low heterogeneity, 50% = moderate heterogeneity, and 75% = high heterogeneity. A higher degree of heterogeneity reflects a larger proportion of true variance that can potentially be explained by covariates through moderator analyses (Borenstein et al., 2009).

Data synthesis. As many methodological aspects differ across cognitive training

studies (e.g., study characteristics, training type and intensity), the effect sizes underlying these studies are likely to differ. All meta-analyses were therefore based on the random-effects model which accounts for variation of the true effect across studies. The alpha level was set at .05 for all analyses. Bonferroni corrections were applied to any post-hoc analyses.

Planned analyses.

Immediate efficacy (pre-post). To examine the efficacy of core-EF CCT in improving cognitive performance, analyses were grouped and run by transfer distance: trained, near-transfer, and far-transfer. An overall effect size was computed for each transfer distance. Effect sizes were also computed for individual (near- and far-transfer) cognitive domains to investigate which domains were likely to improve as a function of core-EF CCT. Individual cognitive domains for trained outcomes were not examined due to the limited number of studies available in this analysis.

Long-term efficacy (pre-test to follow-up). Long-term improvements in cognition were examined by running analyses comparing performance from pre-test to follow-up. All studies providing follow-up data (in original article or separate follow-up article) were included in these analyses. Overall effect sizes were computed for each transfer distance. Analyses were not conducted at the level of individual cognitive domains due to the limited number of studies available.

Moderator analyses. Moderators of CCT efficacy were examined by means of meta-regression and subgroup analyses. Meta-regression was only performed for the continuous moderator variable *age (of mean sample)*. This analysis was not performed for any other continuous variable as the data were not normally distributed (see Melby-Lervåg & Hulme, 2013). Normality was assessed using the Shapiro-Wilk statistic which tests the null hypothesis that the population is normally distributed; a significant statistic indicates non-normality (Shapiro & Wilk, 1965): length of training session ($W = .90, p < .001$), total

number of training sessions ($W = .83, p < .001$), total amount of time spent training ($W = .71, p < .001$), age of mean sample ($W = .98, p = .291$). Subgroups for non-normally-distributed variables were therefore generated in a way that allowed for a relatively equal distribution of studies between each subgroup. Planned subgroup analyses were conducted for: a) type of training received (single-domain/multidomain), b) whether training was adaptive (adaptive/non-adaptive), c) training location (home/lab), d) type of control condition included (active/passive), e) length of training session (up to 30 minutes/over 30 minutes), f) total number of training sessions (up to 12 sessions/over 12 sessions) g) number of weeks spent training (less than four weeks/over four weeks), and h) total amount of time spent training (up to 12 hours/over 12 hours).

Moderator analyses were not reported for individual far-transfer cognitive domains or long-term efficacy data as the number of studies in each subgroup would have been disproportionate due to the small sample size of studies in each analysis. However, these analyses are provided in Supplementary Materials.

Sensitivity analyses. Sensitivity analyses were conducted to determine whether findings were robust to the different assumptions and decisions made when conducting the meta-analyses. This type of analysis constitutes an informal test to evaluate whether the same conclusions can be drawn after using different methods or excluding certain studies (Borenstein et al., 2009; Deeks et al., 2011). Findings were therefore considered robust if the conclusion or interpretation of the analyses remained the same across different assumptions or methods. For instance, the overall mean effect size did not differ as a function of model type: fixed-effects (small effect, $g = 0.33, p < .001$) versus random-effects (small effect, $g = 0.34, p < .001$). Under both assumptions, the overall effect size remained small and significant, indicating that the findings were robust to the type of model used.

Due to the variety of study designs employing different control groups, data were

entered for the most appropriate control group from each study (which varied between AC and PC groups). A sensitivity analysis was run to examine whether the type of control condition (AC, PC, or mixed) affected the overall effect size estimate. The overall mean effect size remained small and significant, regardless of the type of control included. The final analyses were therefore not limited to studies examining a specific type of control group, but included a mix of the most appropriate control groups from each study (AC or PC).

Analyses were conducted assuming different pre-post correlation values ($r = .00$, $r = .60$, and $r = .90$). Irrespective of the imputed correlation value, the same conclusions were drawn for each level of transfer (i.e., significant large effects for trained outcomes, and significant, small effects for near- and far-transfer), indicating that outcomes were robust to the pre-post correlation value. A value of .60 was therefore imputed for each study, in accordance with prior cognitive training meta-analyses (Lampit et al., 2014; Soveri et al., 2017). Lastly, sensitivity analyses were conducted with respect to risk of bias (comparing all studies, only low-risk studies, and only high-risk studies). The overall mean effect size remained small and significant, regardless of risk of bias. All studies were therefore included in the final analyses. Refer to Appendix A for summary results of all sensitivity analyses.

Results

Description of Studies

Search results. Figure 1 illustrates the study flow diagram for this review. From a total of 2,088 articles, 64 studies were included in the meta-analysis—data from 61 studies were analyzed to examine immediate CCT efficacy (pre-test to post-test), and data from 16 studies were analyzed to examine long-term efficacy (pre-test to follow-up).

[INSERT FIGURE 1 ABOUT HERE]

Characteristics of included studies.

The majority of CCT studies were conducted in Europe (59%), followed by North America (24%), Asia (14%), then Australia (3%). Refer to Appendix B for a summary of intervention details, study characteristics, and overall effect sizes for each study included in the quantitative review and meta-analyses.

Participants. Excluding repeated samples from follow-up studies, the number of participants across the studies totaled 3,591 healthy older adults (1,991 trained participants and 1,600 control participants). Most studies utilized screening tools to assess intact cognitive functioning (MMSE range = 24 to 30, or MoCA range = 22 to 30). The mean age of participants across the included studies was 69.32 years ($SD = 3.95$, ranging from 53 to 95). Eight studies included participants younger than 60 years of age.

Study conditions. The number of active and passive control groups were comparable across studies (41% and 48%, respectively), with only few studies including both types of control conditions (11%). Five different types of core-EF CCT interventions were identified: cognitive flexibility training (8%), inhibition training (9%), WM training (34%), multi-cognitive training (training multiple cognitive domains; 23%), and video game training (26%). Multi-cognitive training studies were separated from the video game training studies as having a gamified component may increase motivation and engagement, and thereby produce different training and transfer effects (e.g., Boot et al., 2016). Due to the limited numbers of studies for some types of training, the main subgroup (moderator) analyses compared single-domain training (comprising cognitive flexibility, inhibition, and WM training) and multidomain training (comprising multi-cognitive and video game training).

Interventions. Total time spent training ranged from 2.08 to 100 hours ($M = 17.38$, $SD = 17.11$). Interventions averaged 22.59 training sessions ($SD = 17.06$, range = 3 to 100), with session length ranging from 15 to 90 minutes ($M = 46.52$, $SD = 20.26$). Interventions spanned an average of 7.02 weeks ($SD = 4.68$, range = 1 to 27), and training schedules ranged

from one to seven sessions per week. The majority of training programs were adaptive in nature (77%) rather than non-adaptive (23%). More interventions took place in a laboratory (67%) than at home (33%).

Study design. Most studies employed randomization methods (65%), followed by pseudo-randomization methods (24%), with no randomization methods being least common (11%). Forty-eight percent of studies were classified as having a high risk of bias and fifty-two percent as having a low risk of bias.

Correlations were examined to evaluate collinearity between the moderator variables concerning study and design characteristics. All moderators were found to be independent of one another—all intercorrelations between moderator variables were below the suggested .90 cut-off value for multicollinearity (Card, 2015; Cohen, Cohen, West, & Aiken, 2003). See Appendix C for intercorrelations of regression coefficients between all moderator variables (computed using CMA).

Follow-up studies. Sixteen studies were included in the analyses examining long-term efficacy of core-EF CCT. Of these studies, three comprised WM training, three cognitive flexibility training, four multi-cognitive training, and six video game training. Follow-up length ranged from three weeks to eighteen months ($M_{\text{months}} = 6.83$, $SD = 5.09$). These studies included a total of 1,060 participants (585 participated in a training condition and 475 in a control condition).

Efficacy of Core-EF CCT Interventions

Immediate efficacy (pre-test to post-test). Table 2 presents an overview of summary and moderator effect sizes for all outcomes (including overall and individual cognitive domains) concerning immediate efficacy of core-EF CCT.

[INSERT TABLE 2 ABOUT HERE]

Trained outcomes. Performance on trained outcomes was examined in 24 studies

(see Figure 2). Between these studies, there was a total of 678 trained participants ($M_{\text{sample size}} = 28.25$, $SD = 21.09$) and 514 control participants ($M_{\text{sample size}} = 21.42$, $SD = 13.14$). Most studies in this analysis employed WM training (15 out of 24). There was a significant, large mean effect size in favor of CCT for trained outcomes, $g = 1.00$, 95% CI [0.79, 1.20], $p < .001$. Significant, moderate heterogeneity was evident between studies, $Q(23) = 58.52$, $p < .001$, $I^2 = 60.70\%$. The funnel plot displayed no substantial asymmetry, confirmed by Egger's test ($\beta = 1.97$, $SE = 1.32$, $p = .148$). The trim and fill analysis imputed five studies with negative effect sizes of low precision (in favor of no training effect), though the adjusted effect size was smaller than the observed effect size, $g_{\text{adjusted}} = 0.83$, 95% CI [0.61, 1.05]. Although there was a drop in the effect size adjusted for bias, the key finding (i.e., overall size of the training effect) did not change. According to Borenstein et al. (2009), this may reflect modest impact of publication bias.

[INSERT FIGURE 2 ABOUT HERE]

There were no significant moderators for trained outcomes; the effect sizes were comparable across all moderator variables and subgroups. This indicates that, regardless of intervention or study characteristic, participants who had received repeated training on a task performed better on that same task than those who had not. Moderator analyses for this outcome were therefore not reported here, but a summary of these statistics is presented in Table 2 and in the Supplementary Materials (Table S1). The Supplementary Materials also provides the dataset and presents additional analyses and figures not reported in this paper.

Near-transfer outcomes.

Overall. A near-transfer outcome was included in 55 studies encompassing 1,782 participants in a training group ($M_{\text{sample size}} = 32.40$, $SD = 34.15$) and 1,509 participants in a control group ($M_{\text{sample size}} = 27.44$, $SD = 32.65$). A small, significant mean effect size was observed in favor of CCT for near-transfer outcomes, $g = 0.26$, 95% CI [0.17, 0.34], $p < .001$,

with marginally significant, low heterogeneity between studies, $Q(54) = 68.52, p = .058, I^2 = 31.19\%$ (refer to Table 2 and Figure 3). Although a clear outlier was identified (Borella et al., 2014—Young-Old; $g_{\text{study}} = 1.74$), the one-study-removed analysis revealed that the omission of this study did not substantially impact the overall effect size estimate ($g_{\text{one study removed}} = 0.23$). A sensitivity analysis also indicated that the exclusion of this study did not impact the findings for any subgroup analyses of near-transfer outcomes (see Appendix A, Table A1). The funnel plot was visually symmetrical, confirmed by Egger's test ($\beta = 0.64, SE = 0.43, p = .144$). No studies were imputed in the trim and fill analysis.

Moderator analyses were conducted to further examine the variance of effect sizes between studies for near-transfer outcomes. The subgroup analysis revealed that type of training did not moderate the efficacy of core-EF CCT, $Q(1) = 1.66, p = .198$. As significant effect sizes were produced by both single-domain training ($k = 28, g = 0.32, 95\% \text{ CI } [0.19, 0.46], p < .001$) and multidomain training studies ($k = 27, g = 0.21, 95\% \text{ CI } [0.11, 0.31], p < .001$), an exploratory analysis was conducted to examine whether any specific training types were most effective. The exploratory analyses indicated that only WM training ($k = 21, g = 0.37, 95\% \text{ CI } [0.20, 0.55], p < .001$), multi-cognitive training ($k = 13, g = 0.15, 95\% \text{ CI } [0.04, 0.26], p = .009$), and video game training studies ($k = 14, g = 0.27, 95\% \text{ CI } [0.13, 0.41], p < .001$) produced significant effect sizes after Bonferroni corrections were applied ($.05/5 = .01$). See Figure 3 for the effect sizes of near-transfer outcomes stratified by type of training.

The efficacy of core-EF CCT on near-transfer outcomes was not moderated by training location, $Q(1) = 0.29, p = .591$, or control type, $Q(1) = 1.16, p = .282$. Training adaptiveness was not a significant moderator, $Q(1) = 1.39, p = .239$, however, only adaptive training produced a significant effect size ($k = 47, g = 0.27, 95\% \text{ CI } [0.19, 0.36], p < .001$). Non-adaptive training did not yield a significant effect size ($k = 8, g = 0.13, 95\% \text{ CI } [-0.08, 0.35], p = .224$). Similarly, no training schedule characteristics moderated CCT efficacy for

near-transfer outcomes: length of training session, $Q(1) = 0.00, p = .995$, total number of training sessions, $Q(1) = 0.15, p = .694$, total weeks of training, $Q(1) = 1.35, p = .246$, or total time spent training, $Q(1) = 0.58, p = .446$. All subgroups within these analyses yielded significant and comparable effect sizes (refer to Table 2). Further, there was no significant association between mean sample age and training efficacy ($\beta = 0.02, Q(1) = 2.60, p = .107$).

[INSERT FIGURE 3 ABOUT HERE]

Cognitive flexibility. Of the studies including a near-transfer outcome, 12 measured cognitive flexibility. A significant, small training effect was obtained for untrained cognitive flexibility outcomes, $g = 0.16, 95\% \text{ CI } [0.02, 0.31], p = .031$ (as shown in the Supplementary Materials, Table S2). Exploratory analyses indicated that no types of training produced significant training effects for untrained cognitive flexibility outcomes after Bonferroni corrections were applied ($.05/3 = .017$): cognitive flexibility training ($k = 3, g = 0.33, 95\% \text{ CI } [-0.01, 0.67], p = .056$), multi-cognitive training ($k = 1, g = 0.20, 95\% \text{ CI } [-0.30, 0.69], p = .434$), video game training ($k = 8, g = 0.12, 95\% \text{ CI } [-0.06, 0.29], p = .202$).

Inhibitory control. Thirteen studies assessed inhibitory control as a near-transfer outcome. The analysis revealed a small, significant mean effect size for inhibitory control outcomes, ($k = 13, g = 0.13, 95\% \text{ CI } [0.01, 0.28], p = .048$). Exploratory analyses indicated that no training types produced a significant training effect: inhibition training ($k = 4, g = 0.15, 95\% \text{ CI } [-0.04, 0.44], p = .099$), multi-cognitive training ($k = 4, g = 0.14, 95\% \text{ CI } [-0.11, 0.39], p = .271$), video games training ($k = 5, g = 0.11, 95\% \text{ CI } [-0.13, 0.35], p = .354$).

Memory. In total, 45 studies included a near-transfer memory outcome (assessing WM or short-term memory). These studies produced a significant, small training effect ($k = 45, g = 0.29, 95\% \text{ CI } [0.19, 0.39], p < .001$). Exploratory analyses indicated that WM training ($k = 21, g = 0.37, 95\% \text{ CI } [0.20, 0.55], p < .001$) and video game training ($k = 13, g = 0.36, 95\% \text{ CI } [0.22, 0.50], p < .001$) produced a significant training effect for near-transfer memory

outcomes (even after Bonferroni corrections applied, $.05/3 = .017$), whereas multi-cognitive training did not ($k = 11$, $g = 0.11$, 95% CI $[-0.09, 0.30]$, $p = .278$). Refer to Table S2 in the Supplementary Materials for additional statistics regarding this analysis.

Far-transfer outcomes.

Overall. At least one far-transfer outcome was assessed by 57 studies⁶, encompassing 1,896 participants in a training group ($M_{\text{sample size}} = 33.26$, $SD = 33.71$) and 1,534 in a control group ($M = 26.91$, $SD = 32.16$). No influential studies were detected or removed from the analysis. There was a significant, small training effect for far-transfer outcomes overall, $g = 0.22$, 95% CI $[0.15, 0.29]$, $p < .001$ (see Figure 4 for a forest plot depicting overall far-transfer effect sizes). Heterogeneity between studies was nonsignificant, $Q(56) = 46.92$, $p = .801$, $I^2 = 0.00\%$. Egger's test indicated no significant asymmetry within the funnel plot ($\beta = -0.23$, $SE = 0.36$, $p = .525$), though the trim and fill analysis imputed 16 studies with positive effect sizes of low and high precision (in favor of CCT), $g_{\text{adjusted}} = 0.33$, 95% CI $[0.26, 0.40]$. Although there was an increase in the effect size adjusted for bias, the size of the training effect was left unchanged, suggesting modest impact of publication bias.

Figure 4 illustrates the effect sizes of far-transfer outcomes stratified by training type. Subgroup analyses revealed that type of training moderated the training effect, $Q(1) = 4.56$, $p = .033$, indicating that multidomain training studies produced a larger training effect ($k = 26$, $g = 0.29$, 95% CI $[0.20, 0.37]$, $p < .001$) than single-domain training ($k = 31$, $g = 0.13$, 95% CI $[0.03, 0.24]$, $p = .013$). An exploratory analysis indicated that significant training effects were only obtained for multi-cognitive ($k = 14$, $g = 0.27$, 95% CI $[0.16, 0.39]$, $p < .001$) and video game training studies ($k = 12$, $g = 0.30$, 95% CI $[0.16, 0.45]$, $p < .001$) after Bonferroni corrections were applied ($.05/5 = .01$). Also refer to Table 2 for all subgroup effect sizes.

⁶ Note that this overall far-transfer analysis aggregates numerous measures from different cognitive domains which vary in 'distance' from the trained domains (e.g., attention may share a higher degree of overlap with EFs than fluid intelligence). Findings from this overall analysis should therefore be interpreted with caution. The analyses concerning individual cognitive domains provide more informative details regarding far-transfer.

Training adaptiveness did not moderate the training effect, $Q(1) = 0.09, p = .766$, though studies that employed adaptive training produced a significant effect size ($k = 47, g = 0.23, 95\% \text{ CI } [0.15, 0.30], p < .001$), while in studies employing non-adaptive training the effect approached significance ($k = 10, g = 0.20, 95\% \text{ CI } [-0.01, 0.38], p = .056$). Training location was not a significant moderator, $Q(1) = 0.29, p = .593$, nor was type of control group, $Q(1) = 0.04, p = .852$.

No training schedule characteristics moderated the training effect: length of training session, $Q(1) = 2.44, p = .119$, number of training sessions, $Q(1) = 0.08, p = .782$, total weeks of training, $Q(1) = 0.60, p = .438$, total time spent training, $Q(1) = 0.50, p = .479$. However, there was a significant, positive association between mean sample age and training efficacy ($\beta = 0.02, Q(1) = 4.90, p = .027$), indicating that studies comprising older participants demonstrated larger training improvements.

[INSERT FIGURE 4 ABOUT HERE]

Attention. Twenty-four studies included a far-transfer measure of attention. In total, these studies included 690 trained participants ($M_{\text{sample size}} = 28.75, SD = 17.35$) and 526 control participants ($M_{\text{sample size}} = 21.92, SD = 15.28$). The analysis revealed a small, nonsignificant mean effect size for attention outcomes, $g = 0.11, 95\% \text{ CI } [-0.04, 0.27], p = .155$, with significant, low heterogeneity between studies, $Q(23) = 38.18, p = .024, I^2 = 39.76\%$. The funnel plot was considerably asymmetrical, confirmed by Egger's test ($\beta = 2.63, SE = 0.90, p = .008$) and the trim and fill analysis which imputed seven additional studies with negative effect sizes, $g_{\text{adjusted}} = -0.05, 95\% \text{ CI } [-0.22, 0.12]$. The adjusted mean effect size was nonsignificant but negative, indicating larger improvements for control groups compared to trained groups on attention outcomes.

Executive functioning. Far-transfer measures of executive functioning were assessed in 38 studies encompassing 1,303 intervention participants ($M_{\text{sample size}} = 34.30, SD = 38.68$)

and 1,077 control participants ($M_{\text{sample size}} = 28.34$, $SD = 38.59$). These far-transfer outcomes included tasks which measured global executive functioning or different EFs to those which were trained (e.g., WM training completing an inhibition measure). There was a small, significant mean effect size for executive functioning outcomes in favor of CCT, $g = 0.18$, 95% CI [0.09, 0.26], $p < .001$. Heterogeneity between studies was nonsignificant, $Q(37) = 28.62$, $p = .837$, $I^2 = 0.00\%$. The included studies were distributed symmetrically across the funnel plot ($\beta = 0.25$, $SE = 0.38$, $p = .515$), however five studies with positive effect sizes were imputed to adjust for potential publication bias, $g_{\text{adjusted}} = 0.21$, 95% CI [0.14, 0.29]. The effect size adjusted for bias remained largely unchanged, reflecting trivial impact of publication bias.

Of these 38 studies, 18 included measures of global executive functioning, producing a small, significant training effect ($g = 0.20$, 95% CI [0.10, 0.31], $p < .001$). Exploratory analyses indicated that no types of training produced a significant training effect after Bonferroni corrections were applied ($.05/3 = .017$): WM training ($k = 2$, $g = 0.15$, 95% CI [-0.23, 0.54], $p = .438$), multi-cognitive training ($k = 6$, $g = 0.26$, 95% CI [-0.01, 0.53], $p = .060$), video game training ($k = 4$, $g = 0.19$, 95% CI [-0.06, 0.44], $p = .134$). For all statistics concerning subgroup analyses for each EF outcome, refer to the Supplementary Materials, Table S5.

The studies assessing cognitive flexibility as a far-transfer outcome yielded a nonsignificant training effect ($k = 12$, $g = 0.04$, 95% CI [-0.10, 0.19], $p = .561$). No types of training were effective in promoting improvements on far-transfer cognitive flexibility measures (all $ps \geq .117$). Similarly, the studies assessing inhibitory control as a far-transfer measure did not obtain a significant training effect ($k = 14$, $g = 0.13$, 95% CI [-0.02, 0.27], $p = .092$). No types of training produced significant training effects on these inhibition outcomes (all $ps \geq .067$).

Twelve studies included a far-transfer outcome assessing working memory. A small, significant training effect was obtained ($g = 0.23$, 95% CI [0.01, 0.45], $p = .038$), though no types of training produced improvements on far-transfer working memory measures (all $ps \geq .132$). In sum, far-transfer to executive functioning was observed primarily on global EF and WM tasks.

Fluid intelligence. Far-transfer measures of fluid intelligence were assessed by 33 studies comprising 1,103 intervention participants ($M_{\text{sample size}} = 33.42$, $SD = 21.84$) and 811 control participants ($M_{\text{sample size}} = 24.58$, $SD = 14.23$). The analysis revealed a small but significant mean effect size in favor of CCT, $g = 0.13$, 95% CI [0.04, 0.22], $p = .005$, and no significant heterogeneity between studies, $Q(32) = 31.55$, $p = .489$, $I^2 = 0.00\%$. Although Egger's test indicated no substantial asymmetry ($\beta = 0.84$, $SE = 0.73$, $p = .257$), the trim and fill analysis imputed seven studies with negative effect sizes of low precision to balance out the funnel plot. Notably, the mean effect size for fluid intelligence outcomes adjusted for potential publication bias was nonsignificant, $g_{\text{adjusted}} = 0.05$, 95% CI [-0.06, 0.15]. This reflects substantial impact of publication bias as the key findings changed (became nonsignificant) when adjusting for publication bias.

Language. The planned analysis was not conducted for this domain as only one study employed a far-transfer language measure.

Memory. A total of 16 studies included a far-transfer measure of memory (beyond working memory). The studies in this analysis did not include pure WM training as working memory and short-term memory share characteristics with WM. This analysis encompassed 683 intervention participants ($M_{\text{sample size}} = 42.69$, $SD = 56.59$) and 634 control participants ($M_{\text{sample size}} = 39.63$, $SD = 56.30$). There was a significant, small training effect for verbal and nonverbal memory combined, $g = 0.22$, 95% CI [0.08, 0.36], $p = .002$. The heterogeneity between studies was nonsignificant, $Q(15) = 18.84$, $p = .178$, $I^2 = 24.39\%$. Both effect sizes

for verbal and nonverbal memory were similar. For nonverbal memory, the mean effect size was small and significant ($k = 7$, $g = 0.31$, 95% CI [0.07, 0.55], $p = .011$), with nonsignificant heterogeneity between studies, $Q(6) = 7.02$, $p = .319$, $I^2 = 14.58\%$. Similarly, for verbal memory, the mean effect size was small and significant ($k = 12$, $g = 0.20$, 95% CI [0.04, 0.36], $p = .013$), with nonsignificant heterogeneity between studies, $Q(11) = 15.68$, $p = .154$, $I^2 = 29.83\%$. The funnel plot for overall memory displayed asymmetry toward the right of the mean, though Egger's test was nonsignificant ($\beta = 0.13$, $SE = 0.68$, $p = .852$). The trim and fill analysis, however, imputed three additional studies with positive effect sizes to adjust for potential publication bias, $g_{\text{adjusted}} = 0.28$, 95% CI [0.14, 0.42]. The likely impact of publication bias was considered trivial.

Processing speed. Processing speed was assessed as a far-transfer measure in 22 studies which included a total of 795 participants in intervention conditions ($M_{\text{sample size}} = 36.14$, $SD = 48.69$) and 705 participants in control conditions ($M_{\text{sample size}} = 32.05$, $SD = 48.43$). The mean effect size did not reach significance, $g = 0.18$, 95% CI [-0.01, 0.37], $p = .058$, though there was significant moderate heterogeneity between studies, $Q(21) = 58.71$, $p < .001$, $I^2 = 64.23\%$. The funnel plot showed considerable asymmetry toward the right side of the mean ($\beta = -2.46$, $SE = 0.71$, $p = .003$). Notably, after imputation of five studies with positive effect sizes, the mean effect size for processing speed outcomes adjusted for potential publication bias was small and significant, $g_{\text{adjusted}} = 0.35$, 95% CI [0.17, 0.52]. This adjusted effect size may reflect substantial impact of publication bias.

Visuospatial ability. Five studies included a measure of visuospatial ability, totaling 161 participants in an intervention group ($M_{\text{sample size}} = 32.20$, $SD = 19.66$) and 123 in a control ($M_{\text{sample size}} = 24.60$, $SD = 18.72$). The mean effect size was small and significant, $g = 0.37$, 95% CI [0.08, 0.66], $p = .012$. Heterogeneity was nonsignificant between studies, $Q(4) = 5.36$, $p = .253$, $I^2 = 25.34\%$. Egger's test indicated that the studies were distributed

symmetrically within the funnel plot ($\beta = -1.32$, $SE = 1.97$, $p = .550$). The trim and fill analysis imputed two studies with positive effect sizes of low precision to balance out the funnel plot. This adjusted training effect was significant and moderate in size, $g_{\text{adjusted}} = 0.50$, 95% CI [0.21, 0.80]. However, this adjusted effect size should be interpreted with caution as publication bias detection methods can yield unreliable estimates for analyses comprising a small number of studies (Borenstein et al., 2009).

Long-term efficacy (pre-test to follow-up).

Maintenance of trained outcomes. Eight studies measured trained outcomes at a follow-up testing occasion (see Figure 5). Between these studies, 191 participants were included in an intervention group ($M_{\text{sample size}} = 23.88$, $SD = 16.75$) and 133 in a control group ($M_{\text{sample size}} = 16.63$, $SD = 3.54$). On average, follow-up testing occurred 10 months after post-test (ranging from 4 to 18 months). The mean effect size for trained outcomes from pre-test to follow-up was significant and large, $g = 0.85$, 95% CI [0.40, 1.29], $p < .001$, with significant moderate heterogeneity between studies, $Q(7) = 23.24$, $p = .002$, $I^2 = 69.88\%$. The funnel plot did not indicate any evidence of publication bias ($\beta = 6.74$, $SE = 4.83$, $p = .213$), and no studies were imputed in the trim and fill analysis. After examining the same eight studies in an analysis examining immediate efficacy (pre-post), a slightly larger mean effect size was observed ($g = 0.89$, $p < .001$).

[INSERT FIGURE 5 ABOUT HERE]

Maintenance of near-transfer outcomes. A total of 16 studies included a near-transfer measure at follow-up, with 520 participants included in an intervention group ($M_{\text{sample size}} = 32.46$, $SD = 48.06$) and 458 in a control group ($M_{\text{sample size}} = 28.60$, $SD = 47.09$). On average, follow-up testing occurred 7 months after post-test (ranging from 3 weeks to 18 months). Figure 6 illustrates the forest plot of core-EF CCT efficacy for near-transfer outcomes from pre-test to follow-up. The mean effect size was small and significant, $g =$

0.25, 95% CI [0.06, 0.44], $p = .009$. There was significant, low heterogeneity between studies, $Q(15) = 25.27$, $p = .046$, $I^2 = 40.65\%$. Visual inspection of the funnel plot indicated asymmetry toward the right side of the mean, though Egger's test was not significant ($\beta = 0.03$, $SE = 0.73$, $p = .969$). The trim and fill analysis imputed four studies with positive effect sizes of low precision, $g_{\text{adjusted}} = 0.40$, 95% CI [0.21, 0.59]. The likely impact of publication bias was modest. When examining immediate efficacy of the same 16 studies ($g = 0.26$, $p = .017$) the mean effect size was slightly larger than that for long-term efficacy.

[INSERT FIGURE 6 ABOUT HERE]

Maintenance of far-transfer outcomes. Far-transfer tasks were measured at follow-up in 13 studies. Between these studies, 451 participants took part in an intervention ($M_{\text{sample size}} = 34.71$, $SD = 53.26$) and 412 were part of a control group ($M_{\text{sample size}} = 31.69$, $SD = 52.14$). On average, follow-up testing occurred 8 months after post-test (ranging from 1 to 18 months). The mean effect size was small and significant, $g = 0.25$, 95% CI [0.11, 0.38], $p < .001$. No significant heterogeneity was evident, $Q(11) = 5.86$, $p = .883$, $I^2 = 0.00\%$. See Figure 7 for a forest plot displaying individual and summary effect sizes for far-transfer outcomes from pre-test to follow-up. The funnel plot showed significant asymmetry toward the right of the mean ($\beta = -1.11$, $SE = 0.33$, $p = .006$) and the trim and fill analysis imputed six studies with positive effect sizes of low precision, $g_{\text{adjusted}} = 0.33$, 95% CI [0.21, 0.45]. The likely impact of bias was considered trivial. For the same 13 studies, the mean effect size for immediate efficacy ($g = 0.35$, $p < .001$), was larger than the mean effect size for long-term efficacy.

[INSERT FIGURE 7 ABOUT HERE]

Discussion

This meta-analytic investigation aimed to explore the efficacy of CCT interventions targeting at least one core-EF (cognitive flexibility, inhibition, or WM) in improving

cognitive performance amongst older adults. Immediate efficacy (pre-test to post-test) and long-term efficacy (pre-test to follow-up) were examined across 64 studies. This study also sought to elucidate any moderating factors of intervention efficacy with respect to different aspects of training conditions and study design. The findings from this meta-analysis yield important theoretical and practical implications for cognitive aging research.

Summary of Evidence

Results from the analyses examining immediate efficacy indicated that core-EF CCT groups generally demonstrated larger improvements than control groups on a variety of cognitive domains at all levels of transfer. A significant, large training effect was observed for trained outcomes and a significant, small training effect was observed for overall near-transfer outcomes. A significant, small training effect was also observed for overall far-transfer outcomes, however, when analyzed in more detail, training effects were significant for some individual cognitive domains but not others. Significant training gains were observed for the far-transfer domains of executive functioning, memory, and visuospatial ability from pre-test to post-test. The far-transfer effect on fluid intelligence was significant before but not after adjusting for publication bias. No significant effects were observed on measures of attention or processing speed, though, the training effect on processing speed might be underestimated due to publication bias. Refer to Table 2 for all effect sizes concerning immediate efficacy.

Age of mean sample and type of training moderated CCT efficacy in promoting far-transfer (overall) outcomes, indicating that larger improvements were observed for older compared to younger samples of older adults, and multidomain training compared to single-domain training. Additionally, although the formal moderator statistics were nonsignificant, adaptive training studies consistently yielded significant training effects across trained, near-transfer, and far-transfer outcomes, whereas non-adaptive training studies did not.

Analyses examining long-term efficacy revealed a significant, large mean effect size for the maintenance of trained outcomes, and significant, small mean effect sizes for the maintenance of near-transfer and far-transfer outcomes. When examining data from the same studies for immediate and long-term efficacy, the mean effect sizes for immediate efficacy were comparable to the mean effect sizes for long-term efficacy. From the analyses conducted in this study, it is evident that core-EF training promotes both immediate and long-term improvements in performance on many aspects of cognition amongst healthy older adults.

Transfer distance. As one would expect, performance was greatly improved on trained tasks for the cognitive intervention conditions compared to control conditions which did not receive repeated practice on these tasks. This supports previous findings by Karbach and Verhaeghen (2014) who also reported that EF training produced a large training effect for trained outcomes.

A small but significant training effect was obtained for overall near-transfer outcomes in the present study ($g = 0.26$). Further analysis revealed small and significant mean effect sizes for each EF domain. The size of this near-transfer effect, however, differed from that reported by Karbach and Verhaeghen (2014) who found a moderate effect size ($g = 0.50$). This disparity in effect sizes likely stems from the type of studies included in each meta-analysis. Although both meta-analyses investigated EF training studies, the current study also included multidomain training interventions and not just single-domain EF training. Thus, it may be that single-domain training yields larger training effects for near-transfer measures compared to multidomain training, as focusing on, and spending more time exercising one cognitive process may be more efficient in improving untrained tasks from that same cognitive domain. In contrast, by training a variety of different cognitive processes, performance on many cognitive domains would likely improve, though, the improvement

might be smaller than if only one cognitive process were trained due to less time spent training on each domain. In accordance with this notion, findings from this meta-analysis revealed that single-domain training studies obtained a larger effect size than multidomain studies on near-transfer outcomes, but multidomain training obtained larger effect sizes for overall far-transfer and most individual cognitive domains (see Table 2).

Lastly, small, significant effect sizes were found for overall far-transfer in the current study and by Karbach and Verhaeghen (2014). As core-EFs are required to execute other cognitive processes and perform everyday activities (Diamond, 2013), it is not surprising that improvements in EFs generalized to gains in other cognitive domains as well (Karbach & Kray, 2016). However, this conclusion should be interpreted with careful consideration as this overall far-transfer analysis included training effects from a variety of cognitive domains—some of which likely share a higher degree of underlying commonalities with EFs than others (e.g., short-term memory might be more closely linked to EFs than fluid intelligence). Thus, conclusions of far-transfer need also take into account the training effects for each individual far-transfer cognitive domain (presented in the next section).

Individual cognitive domains. The present meta-analysis observed small, significant training effect sizes on outcome measures of executive functioning, fluid intelligence, memory, and visuospatial ability, but not attention or processing speed. Findings for fluid intelligence and processing speed need to be taken with caution, however, as the statistical significance of these analyses changed after accounting for publication bias.

The findings from this study are similar to that of a recent meta-analysis conducted by Lampit et al. (2014) who examined the efficacy of single-domain and multidomain CCT interventions in healthy older adults (though interventions were not restricted to EF training). These authors reported small, significant effect sizes for overall cognition, verbal memory, nonverbal memory, WM, processing speed, and visuospatial skills. No significant training

effects were reported for attention or executive functioning.

Similar to Lampit et al. (2014), analyses for the domain of language were not performed in this study due to the limited number of studies that included a language outcome. Additionally, it is surprising that both meta-analyses did not find significant effect sizes reflecting transfer to the domain of attention, despite its association with EFs (Diamond, 2013), and despite the inclusion of numerous CCT interventions targeting attention in Lampit et al.'s (2014) meta-analysis. It is possible that CCT is not as effective in improving attention compared to other cognitive domains, or perhaps CCT improves different facets of attention that are not consistently tested by transfer tasks. For instance, due to the overlap between selective attention and inhibitory control (Diamond, 2013), core-EF training might improve performance on measures of this type of attention more than others (e.g., divided attention, focused attention). This type of analysis would currently be difficult to conduct due to the limited number of relevant studies within the literature that include outcome measures of attention. Breaking down analyses of attention into different facets would further limit the number of studies within each analysis, and consequently, any results yielded from such analyses would likely be unreliable. Further investigations are therefore needed before reaching a conclusion regarding CCT efficacy in improving attention.

As noted above, although a significant training effect was evident for the domain of fluid intelligence, this effect appeared to reflect considerable publication bias as the unbiased effect size estimate was nonsignificant. Further, studies including a passive control group produced a significant training effect for fluid intelligence outcomes, whereas studies including active control groups showed a nonsignificant, near-zero effect size (see Table 2), potentially reflecting placebo effects (Dougherty, Hamovitz, & Tidwell, 2016). Therefore, a definitive conclusion regarding improvements to fluid intelligence cannot be drawn based on this analysis and the findings for this cognitive domain should be taken with caution. Further

studies employing stronger methodology (e.g., active control groups) are warranted to provide more conclusive evidence about the efficacy of core-EF CCT in improving fluid intelligence.

Compared to Lampit et al. (2014), the present study found a significant training effect for executive functioning (tasks measuring global executive functioning or an untrained EF domain) but not for processing speed (though, the unbiased effect size estimate for processing speed was significant). Further investigation is required before a conclusion can be made regarding the efficacy of core-EF CCT in improving processing speed. Disparities in findings likely stemmed from the studies included in both meta-analyses. As the current meta-analysis examined studies targeting EFs, it is not surprising that findings reflected transfer to executive functioning—though, it should be noted that this finding was driven by significant improvements in global executive functioning and WM outcome measures; improvements were not observed for far-transfer measures of cognitive flexibility or inhibition. In comparison, Lampit et al. (2014) included numerous CCT studies targeting speed of processing, which would likely have produced an effect size estimate reflecting improvements in the domain of processing speed.

In addition, Lampit et al. (2014) did not separate outcome measures by transfer distance, thereby including a combination of trained, near-transfer, and far-transfer measures in each analysis. Including trained outcomes in their analyses might have inflated the effect sizes to some extent. Accordingly, moderate and high heterogeneity statistics were reported for many of their analyses; this between-groups variance might potentially be explained through subgroup analyses examining transfer distance. As demonstrated in the current study and by Karbach and Verhaeghen (2014), cognitive training typically yields large effect sizes for trained outcomes. Thus, it is critical for meta-analyses examining CCT efficacy to be stratified based on transfer distance—or at the very least, separate analyses should be

conducted for trained outcomes—to provide more accurate effect size estimates of intervention efficacy.

Moderators of CCT efficacy. No variables consistently moderated CCT efficacy across all levels of transfer; however, training type significantly moderated the training effect for far-transfer outcomes (overall), indicating that multidomain training studies produced significantly larger effect sizes than single-domain training studies. As both single-domain and multidomain EF training produced significant effect sizes across each transfer distance, exploratory analyses were conducted to further examine whether any specific types of training appeared to be most effective. These analyses revealed that only the multi-cognitive and video game training studies—both multidomain interventions—consistently produced significant training effects across all levels of transfer (Bonferroni corrections applied). This is not surprising as multidomain interventions exercise numerous cognitive processes and activate a more widespread neural circuitry (Guye et al., 2016), which increases the likelihood of eliciting shared activation with a broader range of transfer tasks.

Generally, slightly larger effect sizes were produced by video game training compared to multi-cognitive training (see Table 2). This may suggest that, compared to training on laboratory tasks, gamified training programs possibly engage different neural mechanisms or elicit different behavioral outcomes (e.g., higher levels of engagement or motivation), which further promotes cognitive transfer. However, it should be noted that this speculation was based on a limited number of studies, and the differences in effect sizes were rather small and not consistent across the different outcomes and cognitive domains. Additionally, the video game training group in the current study included both action video games (e.g., *Rise of Nations*) and non-action “brain training” programs (e.g., *Lumosity*). Although previous studies have found no difference in transfer effects between action and non-action video games (e.g., Boot et al., 2013; Toril et al., 2014), very few have compared video game

training to traditional cognitive training. This is an area that requires further investigation.

Age (of mean sample) was also a significant moderator for overall far-transfer outcomes, indicating that samples encompassing older participants demonstrated larger transfer effects than those comprising younger participants. This is in line with the findings from a video game training meta-analysis conducted by Toril et al. (2014) who found that samples of older participants (71 to 80 years) benefitted more from training than samples comprising younger participants (60 to 70 years). The authors suggested that, compared to their younger counterparts, older participants likely begin training interventions with poorer cognitive functions and therefore have more potential to improve, thereby exhibiting greater improvements after training. However, these findings should be interpreted with caution as these analyses were based on the *average* age of the sample in each study (not the age of individual participants). Thus, it is possible that the average age of some samples was inflated or diminished due to a small number of much older or much younger participants. Nevertheless, these findings indicate that age is an important factor to consider when examining cognitive transfer and should be explored at the level of the individual. This would offer valuable insight into the individual differences of learning and cognitive plasticity.

Further, although the formal moderator statistics were not significant, adaptive training produced significant training effects across all levels of transfer, whereas non-adaptive training did not. This supports the theoretical framework of cognitive plasticity posited by Lövdén et al. (2010) who suggest that adaptive training is essential for improving cognitive and neural efficiency. They posit that improvements in cognitive performance can only occur when there is a modest ‘mismatch’ between an individual’s functional capacity (skill level) and experienced demands (task requirements). This mismatch is provided by adaptive training which continually adjusts task difficulty based on an individual’s current skill level, thereby ensuring that the training is neither too easy nor too difficult for the

individual to complete. As an individual improves on training tasks after repeated practice, the difficulty gradually increases to maintain a certain degree of challenge.

Taken together, adaptive, multidomain CCT interventions may make up an effective cognitive training program for improving cognitive performance in healthy older adults. Further research is required to evaluate this interpretation.

Long-term efficacy. A novel contribution that this study provides for the area of cognitive training is an analysis of the long-term efficacy of CCT at each level of transfer. There was a significant, large effect of training gains (from pre-test to follow-up) for trained outcomes, and significant, small effects of training gains for near-transfer and far-transfer outcomes.

When comparing the effect sizes from the immediate efficacy (pre-test–post-test) and long-term efficacy analyses (pre-test–follow-up) using the same studies, the effect size for immediate efficacy yielded slightly larger effect sizes than that for long-term efficacy across all three levels of transfer. However, the decreases in effect sizes across time were not substantial at any transfer level: trained ($g_{\text{pre-post}} = 0.89$ versus $g_{\text{follow-up}} = 0.85$), near-transfer ($g_{\text{pre-post}} = 0.26$ versus $g_{\text{follow-up}} = 0.25$), and far-transfer ($g_{\text{pre-post}} = 0.35$ versus $g_{\text{follow-up}} = 0.25$). This suggests that improvements in cognitive performance as a function of CCT have the potential to be maintained after completing the intervention. Though, it should be noted that there was a larger decline in the long-term training gain for far-transfer than the other levels, indicating fade-out effects in some studies within this analysis. This decline is somewhat consistent with previous meta-analyses reporting negligible maintenance of WM training gains (Melby-Lervåg & Hulme, 2013; Schwaighofer et al., 2015); however, the long-term training gains were still significant in the current study. Additionally, only 16 out of the 64 different studies (25%) included in this review examined performance at a follow-up testing occasion. These findings should therefore be taken with caution as long-term efficacy from

the other 75% of studies was not considered.

Researchers conducting future CCT studies are encouraged to measure long-term efficacy by including at least one follow-up testing occasion to help fill this gap in the literature. More detailed analyses of long-term efficacy should also evaluate the effect of interval length between post-test and follow-up testing occasions. This will help inform researchers about the durability of training effects, and the need for 'booster' sessions (e.g., Ball et al., 2002) to prolong the benefits of cognitive training.

Overall, the evidence from the studies included in this meta-analytic investigation suggests that core-EF CCT interventions show promise in improving short-term and long-term cognitive performance in older adults.

Quality of Evidence

Bias and heterogeneity. The summary risk of bias for the studies included in this review was evenly distributed between high- and low-risk. Importantly, the key findings were consistent and did not differ between studies which were high- or low-risk only, or for all studies included.

The majority of studies which were evaluated as high-risk resulted from lack of outcome assessor blinding, though, it is acknowledged that this procedure may be challenging to implement for reasons such as practicality, timing, and funding. This source of bias has the potential to substantially impact the results of a CCT study as researchers may consciously or subconsciously give hints or extra encouragement to participants known to be assigned to an intervention condition. As such, these participants might perform or behave differently in the presence of the researcher compared to a blinded outcome assessor. Researchers are encouraged to take this into consideration when designing their CCT studies and methodologies. If full blinding cannot be achieved, other methods may be implemented to minimize the potential for bias and demonstrate that this source of bias was considered and

addressed. For example, a random subset of participants could be assigned to a blinded outcome assessor, or—with participants' consent—the administration of outcome assessments could be video recorded and rated by an independent viewer (or viewers) in terms of similarities and differences between administration for intervention and control participants. Overall, however, the quality of studies included in this review was sound.

Heterogeneity between studies was low and nonsignificant for most analyses. Moderate heterogeneity was observed for trained (pre-post, follow-up), and processing speed analyses. Some of this variability might be explained through subgroup analyses examining different types of training, though these were not conducted due to limited sample size. Nevertheless, the findings do suggest that core-EF training promotes short- and long-term cognitive improvements.

Significant publication bias was evident in the analysis concerning attention outcomes, suggesting that if all relevant studies had been included, the effect size would have been negative (i.e., control groups showed larger improvements than training groups). Evidence of publication bias was also apparent within the processing speed analysis indicating that the unbiased effect size estimate was significant. In contrast, the unbiased effect size estimate for fluid intelligence was nonsignificant.

Methodological limitations. Potential biases were contained as much as possible, though this study is not without limitations. There is a plethora of studies that implement CCT interventions amongst the older adult population, however, this review was restricted to studies training core-EFs. Including all types of CCT interventions was beyond the scope of this review, though it would be interesting to compare the efficacy of core-EF training to other types of training (e.g., episodic memory training, speed of processing training). This is a potential avenue for future research.

Further, it is impossible to tease apart the effects of EF training from multidomain

interventions that also train a variety of other cognitive domains. Thus, it is difficult to discern whether it is, in fact, EF training that is promoting transfer effects, or whether it is a combination of EF training and other cognitive domains. Additionally, EFs have been found to be moderately correlated (Miyake et al., 2000), which upholds the conception that the three core-EFs support one another and generally co-occur (Diamond, 2013). Thus, as core-EFs are interrelated and are essential for carrying out everyday activities, training all three (compared to a single EF) has the potential to foster gains in a wide range of higher-order cognitive functions and promote gains in skills that have real world applications. Further studies training all three EFs are therefore required to investigate this notion, in addition to a more specific analysis of data (e.g., separating multidomain studies which include only core-EFs from those which include EFs and other cognitive domains).

Although there is currently no consensus regarding the minimum number of studies required to perform a random-effects meta-analysis, the Cochrane Handbook suggests that analyses comprising fewer than 10 studies may obtain low power when conducting numerous statistical tests (Sterne, Egger, & Moher, 2011). In accordance, some analyses may have been underpowered, though, this was only the case for two out of the thirteen analyses (visuospatial ability and follow-up trained tasks). Additionally, an analysis for language was not conducted due to the limited number of available studies examining this domain. All other analyses included at least 10 studies each. On a similar note, some moderator analyses included an uneven number of studies in each subgroup (e.g., for type of training, there were many more WM, multi-cognitive, and video game studies than inhibition, and cognitive flexibility studies). It is possible that this may have had an impact on the outcome of the moderator analyses. These findings should therefore be interpreted with caution.

Future Directions and Implications

Although some types of training may not appear to be effective, the mixed outcomes

in the current literature indicate that it would be premature to cease researching the benefits of cognitive training. A constructive step forward would be to invest resources to further investigate the cognitive and neural mechanisms underlying cognitive training in order to develop a more comprehensive framework linking cognitive training, transfer, and aging. This would, in turn, provide a strong theoretical foundation for future studies to investigate, “under what conditions might this type of training be effective?” or “how might this training be modified to increase its efficacy?” Additionally, further investigation into the foundations of different cognitive constructs such as executive functions may also be required to gain a deeper understanding of cognitive functioning and development. These actions would aid in the development of effective interventions based on more defined and stronger fitting models of executive functions (and general cognitive functioning). If there is any possibility for maintaining and improving cognition in later life, then this research is a worthwhile endeavor to discover the full potential that cognitive training has to offer (Katz et al., 2018).

In particular, the cognitive domain of attention is a worthwhile candidate for further research and development. Performance in attention tasks did not improve as a function of CCT, as observed in the current meta-analysis and by Lampit et al. (2014). Thus, future studies may need to include different types of attentional outcome measures to establish whether CCT improves specific facets of attention, or alternatively, directly target attention within training programs. Other cognitive domains of interest for further research include language and visuospatial ability as studies examining these domains were limited in number. Additional research is therefore warranted in order to more accurately determine the efficacy of core-EF CCT in enhancing language and visuospatial skills amongst older adults.

Critically, the majority of outcome measures assessed in the included studies were laboratory tasks and standardized tests. Although performance on some of these measures do predict success in real-world outcomes and abilities, it is not certain that improvements on

transfer tasks will, in fact, translate into more efficient everyday functioning (Schmiedek, 2016). To this end, conclusions about whether CCT can improve general cognitive functioning would therefore be unfounded within this meta-analytic investigation. To address this issue, researchers are encouraged to incorporate real-world outcomes (e.g., driving ability, medication adherence) or more ecologically valid tasks (e.g., Timed Instrumental Activities of Daily Living; Owsley, Sloane, McGwin Jr, & Ball, 2002) in future studies. Additionally, more research is required to determine whether pure EF training promotes cognitive improvements or whether it is a combination of training core-EFs and other cognitive domains. Following the direction of Binder et al. (2016), it may also be fruitful for future studies to compare the efficacy of multidomain training and single-domain training programs.

Moreover, it may be the case that some participants do demonstrate large training and transfer gains, but these effects are masked by analyses based on group means (Karch, 2014). Exploration of individual differences such as baseline performance should also be considered in future studies to progress the understanding of whether certain individuals benefit more from training than others. Two accounts have been proposed to explain potential individual differences: the magnification account suggests that high-performing individuals will benefit most from training due to already efficient cognitive processes, whereas the compensation account suggests that low-performing individuals will benefit most as they have more room for improvement than those already performing at a high level (e.g., Lövdén, Brehmer, Li, & Lindenberger, 2012). For example, Karch, Könen, and Spengler (2017) administered a task-switching training intervention to children, young adults, and older adults. These authors observed a compensation effect, whereby individuals who performed worse as baseline (pre-test)—children and older adults in particular—showed the largest improvements at post-test (i.e., larger training and transfer gains) compared to those who had

already demonstrated efficient performance at baseline (mostly young adults). In contrast, Lövdén et al. (2012) observed a magnification effect after implementing a strategy-based memory training intervention amongst younger children, older children, younger adults, and older adults. These authors found that, compared to all other age groups, baseline memory performance was highest in younger adults who also exhibited the largest training gains. Thus, if certain individuals benefit more from specific types of training, interventions can start to be developed and targeted to suit different people.

Researchers are also encouraged to include active control groups in their study design in order to control for non-focal effects such as expectancy to improve, social contact with the researchers, and general cognitive stimulation. This would allow for more accurate comparisons between intervention and control groups and provide stronger evidence regarding the efficacy of CCT interventions.

Many studies in this field of research tend to focus on the question, “does cognitive training work?” (Katz et al., 2018). This question has been explored using different types of cognitive interventions amongst different populations, though the verdict remains inconclusive. The limited scope of this question and the mixed conclusions from numerous meta-analyses (e.g., Au et al., 2015; Karbach & Verhaeghen, 2014; Lampit et al., 2014; Melby-Lervåg & Hulme, 2016; Schwaighofer et al., 2015; Toril et al., 2014), may be fueling the ongoing debate on this topic, doing little to further this area of research and progress toward a more thorough understanding of cognitive training. Rather, more attention needs to be placed on investigating the mechanisms underlying transfer effects in order to establish the optimal training conditions that elicit consistent broad transfer, and explore whether some types of training are more beneficial for particular individuals (Jaeggi, Karbach, & Strobach, 2017). Instead of producing an endless cycle of replications, researchers need to synthesize information from previous studies and meta-analyses to improve methodology, provide

stronger empirical evidence, and advance this area of research by taking a step forward in the right direction and answering the right questions (Katz et al., 2018).

Conclusion

This meta-analytic investigation explored the immediate and long-term efficacy of core-EF CCT interventions in improving cognition amongst older adults. CCT targeting core-EFs yielded small but significant gains to untrained tasks from a variety of cognitive domains. Gains in cognitive ability appeared to be maintained at follow-up. Overall, the findings from this study suggest that using core-EF CCT to improve cognition in older adults is a worthwhile pursuit. Further high-quality studies are warranted to build upon this emerging area of research and to provide stronger empirical evidence regarding the efficacy of core-EF CCT. From here, focus needs to be placed on identifying the optimal training conditions that promote improvements in performance on real-world outcomes to help foster mental health and wellbeing in later life.

References

References marked with an asterisk indicate studies included in the quantitative review and meta-analyses.

Adrian, J., Postal, V., Moessinger, M., Rasclé, N., & Charles, A. (2011). Personality traits and executive functions related to on-road driving performance among older drivers. *Accident Analysis & Prevention*, 43, 1652-1659.

doi:10.1016/j.aap.2011.03.023

*Akimoto, Y., Nozawa, T., Kanno, A., Kambara, T., Ihara, M., Ogawa, T., . . . Kawashima, R. (2016). High-gamma power changes after cognitive intervention: Preliminary results from twenty-one senior adult subjects. *Brain and Behavior*, 6, e00427.

doi:10.1002/brb3.427

American Psychological Association (2008). Reporting standards for research in psychology:

Why do we need them? What might they be? *The American Psychologist*, 63, 839-851. doi:10.1037/0003-066X.63.9.839

*Anderson, S., White-Schwoch, T., Choi, H. J., & Kraus, N. (2013a). Training changes processing of speech cues in older adults with hearing loss. *Frontiers in Systems Neuroscience*, 7, 97-106. doi:10.3389/fnsys.2013.00097

*Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013b). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences*, 110, 4357-4362. doi:10.1073/pnas.1213555110

*Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., . . . Gazzaley, A. (2013). Video game training enhances cognitive control in older adults. *Nature*, 501, 97-101. doi:10.1038/nature12486

Au, J., Buschkuehl, M., Duncan, G.J., Jaeggi, S.M., 2016. There is no convincing evidence that working memory training is NOT effective: A reply to Melby-Lervåg and Hulme

(2015). *Psychonomic Bulletin & Review*, 23, 331-337. doi:10.3758/s13423-015-0967-

4

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuhl, M., & Jaeggi, S. M. (2015).

Improving fluid intelligence with training on working memory: A meta-analysis.

Psychonomic Bulletin & Review, 22, 366-377. doi:10.3758/s13423-014-0699-x

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., . . .

Tennstedt, S. L. (2002). Effects of cognitive training interventions with older adults:

A randomized controlled trial. *JAMA*, 288, 2271-2281. doi:10.1001/jama.288.18.2271

*Ballesteros, S., Mayas, J., Prieto, A., Toril, P., Pita, C., Ponce de León, L., . . . Waterworth,

J. A. (2015). A randomized controlled trial of brain training with non-action video

games in older adults: Results of the 3-month follow-up. *Frontiers in Aging*

Neuroscience, 7, 1-12. doi:10.3389/fnagi.2015.00045

*Ballesteros, S., Mayas, J., Prieto, A., Ruiz-Marquez, E., Toril, P., & Reales, J. M. (2017).

Effects of video game training on measures of selective attention and working

memory in older adults: Results from a randomized controlled trial. *Frontiers in*

Aging Neuroscience, 9, 354-369. doi:10.3389/fnagi.2017.00354

*Ballesteros, S., Prieto, A., Mayas, J., Toril, P., Pita, C., Ponce de León, L., . . . Waterworth,

J. (2014). Brain training with non-action video games enhances aspects of cognition

in older adults: A randomized controlled trial. *Frontiers in Aging Neuroscience*, 6, 1-

14. doi:10.3389/fnagi.2014.00277

*Basak, C., Boot, W. R., Voss, M. W., & Kramer, A. F. (2008). Can training in a real-time

strategy video game attenuate cognitive decline in older adults? *Psychology and*

Aging, 23, 765-777. doi:10.1037/a0013494

Bediou, B., Adams, D. M., Mayer, R. E., Tipton, E., Green, C. S., & Bavelier, D. (2018).

Meta-analysis of action video game impact on perceptual, attentional, and cognitive

- skills. *Psychological Bulletin*, 144, 77-110. doi:10.1037/bul0000130
- Bellander, M., Eschen, A., Lövdén, M., Martin, M., Bäckman, L., & Brehmer, Y. (2016). No evidence for improved associative memory performance following process-based associative memory training in older adults. *Frontiers in Aging Neuroscience*, 8, 326-337. doi:10.3389/fnagi.2016.00326
- Bherer, L., Erickson, K. I., & Liu-Ambrose, T. (2013). A review of the effects of physical activity and exercise on cognitive and brain functions in older adults. *Journal of Aging Research*, 2013, 1-8. doi:10.1155/2013/657508
- *Binder, J. C., Martin, M., Zolig, J., Rocke, C., Merillat, S., Eschen, A., . . . Shing, Y. L. (2016). Multi-domain training enhances attentional control. *Psychology and Aging*, 31, 390-408. doi:10.1037/pag0000081
- Boot, W. R., Champion, M., Blakely, D. P., Wright, T., Souders, D. J., & Charness, N. (2013). Video games as a means to reduce age-related cognitive decline: Attitudes, compliance, and effectiveness. *Frontiers in Psychology*, 4, 31-40. doi:10.3389/fpsyg.2013.00031
- Boot, W. R., Souders, D. J., Charness, N., Blocker, K., Roque, N., & Vitale, T. (2016). The gamification of cognitive training: Older adults' perceptions of and attitudes toward digital game-based interventions. *International Conference on Human Aspects of IT for the Aged Population*, 9754, 290-300. doi:10.1007/978-3-319-39943-0_28
- Borella, E., Carbone, E., Pastore, M., De Beni, R., & Carretti, B. (2017). Working memory training for healthy older adults: The role of individual characteristics in explaining short- and long-term gains. *Frontiers in Human Neuroscience*, 11, 99-120. doi:10.3389/fnhum.2017.00099
- *Borella, E., Carretti, B., Cantarella, A., Riboldi, F., Zavagnin, M., & De Beni, R. (2014). Benefits of training visuospatial working memory in young-old and old-old.

Developmental Psychology, 50, 714-727. doi:10.1037/a0034293

Borella, E., Carretti, B., Riboldi, F., & De Beni, R. (2010). Working memory training in older adults: Evidence of transfer and maintenance effects. *Psychology and Aging*, 25, 767-778. doi:10.1037/a0020683

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Hoboken, USA: John Wiley and Sons Ltd.

Brand, M., Recknor, E. C., Grabenhorst, F., & Bechara, A. (2007). Decisions under ambiguity and decisions under risk: Correlations with executive functions and comparisons of two different gambling tasks with implicit and explicit rules. *Journal of Clinical and Experimental Neuropsychology*, 29, 86-99. doi:10.1080/13803390500507196

Brehmer, Y., Rieckmann, A., Bellander, M., Westerberg, H., Fischer, H., & Bäckman, L. (2011). Neural correlates of training-related working-memory gains in old age. *Neuroimage*, 58, 1110-1120. doi:10.1016/j.neuroimage.2011.06.079

Brehmer, Y., Westerberg, H., & Bäckman, L. (2012). Working-memory training in younger and older adults: Training gains, transfer, and maintenance. *Frontiers in Human Neuroscience*, 6, 63-70. doi:10.3389/fnhum.2012.00063

*Brom, S. S., & Kliegel, M. (2014). Improving everyday prospective memory performance in older adults: Comparing cognitive process and strategy training. *Psychology and Aging*, 29, 744-755. doi:10.1037/a0037181

*Buitenweg, J. I., van de Ven, R. M., Prinssen, S., Murre, J. M., & Ridderinkhof, K. R. (2017). Cognitive flexibility training: A large-scale multimodal adaptive active-control intervention study in healthy older adults. *Frontiers in Human Neuroscience*, 11, 529-543. doi:10.3389/fnhum.2017.00529

*Bürki, C. N., Ludwig, C., Chicherio, C., & de Ribaupierre, A. (2014). Individual differences

in cognitive plasticity: An investigation of training curves in younger and older adults.

Psychological Research, 78, 821-835. doi:10.1007/s00426-014-0559-3

*Cantarella, A., Borella, E., Carretti, B., Kliegel, M., Mammarella, N., Fairfield, B., & De Beni, R. (2017). The influence of training task stimuli on transfer effects of working memory training in aging. *Psychologie Française*, 2, 1-15.
doi:10.1016/j.psfr.2017.04.005

Card, N. A. (2015). *Applied meta-analysis for social science research*. New York, USA: Guilford Publications.

*Chambon, C., Herrera, C., Romaiguere, P., Paban, V., & Alescio-Lautier, B. (2014). Benefits of computer-based memory and attention training in healthy older adults. *Psychology and Aging*, 29, 731-743. doi:10.1037/a0037477

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd rev ed.). New York: Taylor & Francis Inc.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science*, 14, 125-130. doi:10.1111/1467-9280.t01-1-01430

Comprehensive Meta-Analysis (Version 3.3) [Computer software] (2014). Englewood, NJ: Biostat. Available from <http://www.meta-analysis.com>

*Cuenen, A., Jongen, E. M. M., Brijs, T., Brijs, K., Houben, K., & Wets, G. (2016). Effect of a working memory training on aspects of cognitive ability and driving ability of older drivers: Merits of an adaptive training over a non-adaptive training. *Transportation Research Part F: Traffic Psychology and Behaviour*, 42, 15-27.

doi:10.1016/j.trf.2016.06.012

*Dahlin, E., Nyberg, L., Bäckman, L., & Stigsdotter Neely, A. (2008a). Plasticity of executive functioning in young and older adults: Immediate training gains, transfer, and long-term maintenance. *Psychology and Aging*, 23, 720-730.

doi:10.1037/a0014296

*Dahlin, E., Stigsdotter Neely, A., Larsson, A., Bäckman, L., & Nyberg, L. (2008b). Transfer of learning after updating training mediated by the striatum. *Science*, 320, 1510-1512.

doi:10.1126/science.1155466

Deeks, J. J., Higgins, J. P. T., & Altman, D. G. (2011). Chapter 9: Analysing data and undertaking meta-analyses. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions version 5.1.0*. (pp. 9.1-9.44). Retrieved from www.cochrane-handbook.org

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135-168.

doi:10.1146/annurev-psych-113011-143750

Dougherty, M. R., Hamovitz, T., & Tidwell, J. W. (2016). Reevaluating the effectiveness of n-back training on transfer through the Bayesian lens: Support for the null.

Psychonomic Bulletin & Review, 23, 306-316. doi:10.3758/s13423-015-0865-9

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56, 455-463.

doi:10.1111/j.0006-341X.2000.00455.x

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

Farhady, H. (1982). Measures of language proficiency from the learner's perspective. *TESOL Quarterly*, 16, 43-59. doi:10.2307/3586562

*Gaál, Z. A., & Czigler, I. (in press). Task-Switching Training and Transfer. *Journal of*

Psychophysiology. doi:10.1027/0269-8803/a000189

- *Gajewski, P. D., & Falkenstein, M. (2012). Training-induced improvement of response selection and error detection in aging assessed by task switching: Effects of cognitive, physical, and relaxation training. *Frontiers in Human Neuroscience*, 6, 130-148. doi:10.3389/fnhum.2012.00130
- *Gajewski, P. D., & Falkenstein, M. (in press). ERP and behavioral effects of physical and cognitive training on working memory in aging: A randomized controlled study. *Neural Plasticity*.
- *Goghari, V. M., & Lawlor-Savage, L. (2017). Comparison of cognitive change after working memory training and logic and planning training in healthy older adults. *Frontiers in Aging Neuroscience*, 9, 39-51. doi:10.3389/fnagi.2017.00039
- Goldstein, S., Naglieri, J. A., Princiotta, D., & Otero, T. M. (2014). Introduction: A history of executive functioning as a theoretical and clinical construct. In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 3-12). New York, NY: Springer.
- Green, C. S., & Bavelier, D. (2008). Exercising your brain: A review of human brain plasticity and training-induced learning. *Psychology and Aging*, 23, 692-701. doi:10.1037/a0014345
- *Grönholm-Nyman, P., Soveri, A., Rinne, J. O., Ek, E., Nyholm, A., Stigsdotter Neely, A., & Laine, M. (2017). Limited effects of set shifting training in healthy older adults. *Frontiers in Aging Neuroscience*, 9, 69-90. doi:10.3389/fnagi.2017.00069
- Guye, S., Röcke, C., Mérillat, S., von Bastian, C. C., & Martin, M. (2016). Adult lifespan. In T. Strobach & J. Karbach (Eds.), *Cognitive training: An overview of features and applications* (pp. 45-57). Switzerland: Springer.
- *Guye, S., & von Bastian, C. C. (2017). Working memory training in older adults: Bayesian

evidence supporting the absence of transfer. *Psychology and Aging*, 32, 732-746.

doi:10.1037/pag0000206

Haimov, I., & Shatil, E. (2013). Cognitive training improves sleep quality and cognitive function among older adults with insomnia. *PLoS One*, 8, e61390.

doi:10.1371/journal.pone.0061390

Happé, F. (2013). Fluid Intelligence. In F. F. Volkmar (Ed.), *Encyclopedia of Autism Spectrum Disorders* (pp. 1310-1311). New York, NY: Springer New York.

*Heinzel, S., Lorenz, R. C., Pelz, P., Heinz, A., Walter, H., Kathmann, N., . . . Stelzel, C.

(2016). Neural correlates of training and transfer effects in working memory in older adults. *Neuroimage*, 134, 236-249. doi:10.1016/j.neuroimage.2016.03.068

*Heinzel, S., Rimpel, J., Stelzel, C., & Rapp, M. A. (2017). Transfer effects to a multimodal dual-task after working memory training and associated neural correlates in older adults: A pilot study. *Frontiers in Human Neuroscience*, 11, 85-100.

doi:10.3389/fnhum.2017.00085

*Heinzel, S., Schulte, S., Onken, J., Duong, Q. L., Riemer, T. G., Heinz, A., . . . Rapp, M. A.

(2014). Working memory training improvements and gains in non-trained cognitive tasks in young and older adults. *Aging, Neuropsychology, and Cognition*, 21, 146-173. doi:10.1080/13825585.2013.790338

*Hering, A., Meuleman, B., Bürki, C., Borella, E., & Kliegel, M. (2017). Improving older adults' working memory: The influence of age and crystallized intelligence on training outcomes. *Journal of Cognitive Enhancement*, 1, 358-373.

doi:10.1007/s41465-017-0041-4

Higgins, J. P. T., & Deeks, J. J. (2011). Chapter 7: Selecting studies and collecting data. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions version 5.1.0*. (pp. 7.1-9.28). Retrieved from www.cochrane-

handbook.org

Higgins, J. P. T., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327, 557-560. doi:10.1136/bmj.327.7414.557

Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions version 5.1.0*. Retrieved from www.cochrane-handbook.org

Higgins, J. P. T., Altman, D. G., & Sterne, J. A. C. (2011). Chapter 8: Assessing risk of bias in included studies. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions version 5.1.0*. (pp. 8.1-8.53). Retrieved from www.cochrane-handbook.org

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., . . . Cochrane Bias Methods Group. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, 343, 5928-5937. doi:10.1136/bmj.d5928

Higgins, J. P. T., Deeks, J. J., & Altman, D. G. (2011). Chapter 16: Special topics in statistics. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions version 5.1.0*. (pp. 16.1-16.39). Retrieved from www.cochrane-handbook.org

Hogrefe, A. B., Studer-Luethi, B., Kodzhabashev, S., & Perrig, W. J. (2017). Mechanisms underlying n-back training: Response consistency during training influences training outcome. *Journal of Cognitive Enhancement*, 1, 406-418. doi:10.1007/s41465-017-0042-3

Jaeggi, S. M., Karbach, J., & Strobach, T. (2017). Editorial special topic: Enhancing brain and cognition through cognitive training. *Journal of Cognitive Enhancement*, 1, 353-357. doi:10.1007/s41465-017-0057-9

*Ji, Y., Wang, J., Chen, T., Du, X., & Zhan, Y. (2016). Plasticity of inhibitory processes and

associated far-transfer effects in older adults. *Psychology and Aging*, 31, 415-429.

doi:10.1037/pag0000102

Jonides, J. (2004). How does practice makes perfect? *Nature Neuroscience*, 7, 10-11.

doi:10.1038/nn0104-10

Karbach, J. (2014). Game-based cognitive training for the aging brain. *Frontiers in Psychology*, 5, 1100. doi:10.3389/fpsyg.2014.01100

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

*Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science*, 12, 978-990. doi:10.1111/j.1467-7687.2009.00846.x

Karbach, J., & Kray, J. (2016). Executive functions. In T. Strobach & J. Karbach (Eds.), *Cognitive training: An overview of features and applications* (pp. 93-105). Switzerland: Springer.

Karbach, J., & Verhaeghen, P. (2014). Making working memory work: A meta-analysis of executive-control and working memory training in older adults. *Psychological Science*, 25, 2027-2037. doi:10.1177/0956797614548725

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The Unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, 144, 1147-1185. doi:10.1037/bul0000160

Katz, B., Shah, P., & Meyer, D. E. (2018). How to play 20 questions with nature and lose: Reflections on 100 years of brain-training research. *Proceedings of the National Academy of Sciences*, 115, 9897-9904. doi:10.1073/pnas.1617102114

Kendig, H., McDonald, P., & Piggott, J. (2016). *Population ageing and Australia's future*.

Canberra: ANU Press.

*Kim, H., Chey, J., & Lee, S. (2017). Effects of multicomponent training of cognitive control on cognitive function and brain activation in older adults. *Neuroscience Research*. doi:10.1016/j.neures.2017.05.004

Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2002). Complex prospective memory and executive control of working memory: A process model. *Psychological Test and Assessment Modeling*, 44(2), 303-318.

Könen, T., Strobach, T., & Karbach, J. (2016). Working memory. In T. Strobach & J. Karbach (Eds.), *Cognitive training: An overview of features and applications* (pp. 59-68). Switzerland: Springer.

*Kray, J., & Feher, B. (2017). Age differences in the transfer and maintenance of practice-induced improvements in task switching: The impact of working-memory and inhibition demands. *Frontiers in Psychology*, 8, 410-428. doi:10.3389/fpsyg.2017.00410

*Küper, K., Gajewski, P. D., Friege, C., & Falkenstein, M. (2017). A randomized controlled ERP study on the effects of multi-domain cognitive training and task difficulty on task switching performance in older adults. *Frontiers in Human Neuroscience*, 11, 184-196. doi:10.3389/fnhum.2017.00184

Lampit, A., Hallock, H., & Valenzuela, M. (2014). Computerized cognitive training in cognitively healthy older adults: A systematic review and meta-analysis of effect modifiers. *PLoS Med*, 11, e1001756. doi:10.1371/journal.pmed.1001756

*Lange, S., & Süß, H. (2015). Experimental evaluation of near-and far-transfer effects of an adaptive multicomponent working memory training. *Applied Cognitive Psychology*, 29, 502-514. doi:10.1002/acp.3126

- *Li, B., Zhu, X., Hou, J., Chen, T., Wang, P., & Li, J. (2016). Combined cognitive training vs. memory strategy training in healthy older adults. *Frontiers in Psychology*, 7, 834-845. doi:10.3389/fpsyg.2016.00834
- *Li, S. C., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: Practice gain, transfer, and maintenance. *Psychology and Aging*, 23, 731-742. doi:10.1037/a0014343
- Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin*, 136, 659-676. doi:10.1037/a0020080
- Lövdén, M., Brehmer, Y., Li, S. C., & Lindenberger, U. (2012). Training-induced compensation versus magnification of individual differences in memory performance. *Frontiers in Human Neuroscience*, 6, 1-14. doi:10.3389/fnhum.2012.00141
- Lustig, C., Shah, P., Seidler, R., & Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: A review and future directions. *Neuropsychology Review*, 19, 504-522. doi:10.1007/s11065-009-9119-9
- Madden, D. J., Bennett, I. J., & Song, A. W. (2009). Cerebral white matter integrity and cognitive aging: contributions from diffusion tensor imaging. *Neuropsychology Review*, 19, 415-435. doi:10.1007/s11065-009-9113-2
- Martin, M., Kliegel, M., & McDaniel, M. A. (2003). The involvement of executive functions in prospective memory performance of adults. *International Journal of Psychology*, 38, 195-206. doi:10.1080/00207590344000123
- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, 49, 270-291. doi:10.1037/a0028228
- Melby-Lervåg, M., & Hulme, C. (2016). There is no convincing evidence that working memory training is effective: A reply to Au et al. (2014) and Karbach and

Verhaeghen (2014). *Psychonomic Bulletin & Review*, 23, 324-330.

doi:10.3758/s13423-015-0862-z

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

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21, 8-14. doi:10.1177/0963721411429458

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

*Mozolic, J. L., Long, A. B., Morgan, A. R., Rawley-Payne, M., & Laurienti, P. J. (2011). A cognitive training intervention improves modality-specific attention in a randomized controlled trial of healthy older adults. *Neurobiology of Aging*, 32, 655-668. doi:10.1016/j.neurobiolaging.2009.04.013

Noack, H., Lövdén, M., Schmiedek, F., & Lindenberger, U. (2009). Cognitive plasticity in adulthood and old age: Gauging the generality of cognitive intervention effects. *Restorative Neurology and Neuroscience*, 27, 435-453. doi:10.3233/RNN-2009-0496

*Nozawa, T., Taki, Y., Kanno, A., Akimoto, Y., Ihara, M., Yokoyama, R., . . . Takeuchi, H. (2015). Effects of different types of cognitive training on cognitive function, brain structure, and driving safety in senior daily drivers: A pilot study. *Behavioural Neurology*, 2015, 1-18. doi:10.1155/2015/525901

Oei, A. C., & Patterson, M. D. (2014). Are videogame training gains specific or general?

- Frontiers in Systems Neuroscience*, 8, 54-63. doi:10.3389/fnsys.2014.00054
- Orwin, R. G., & Vevea, J. L. (2009). Evaluating coding decisions. In C. H., L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 177-203). USA: Russell Sage Foundation.
- Otero, T. M., & Barker, L. A. (2014). The frontal lobes and executive functioning. In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 29-44). New York, NY: Springer.
- Owsley, C. (2013). Visual processing speed. *Vision Research*, 90, 52-56.
doi:10.1016/j.visres.2012.11.014
- Owsley, C., Sloane, M., McGwin Jr, G., & Ball, K. K. (2002). Timed instrumental activities of daily living tasks: Relationship to cognitive function and everyday performance assessments in older adults. *Gerontology*, 48, 254-265. doi:10.1159/000058360
- Park, D. C., & Reuter-Lorenz, P. A. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, 60, 173-196.
doi:10.1146/annurev.psych.59.103006.093656
- *Payne, B. R., & Stine-Morrow, E. A. (2017). The effects of home-based cognitive training on verbal working memory and language comprehension in older adulthood. *Frontiers in Aging Neuroscience*, 9, 256-276. doi:10.3389/fnagi.2017.00256
- *Peretz, C., Korczyn, A. D., Shatil, E., Aharonson, V., Birnboim, S., & Giladi, N. (2011). Computer-based, personalized cognitive training versus classical computer games: A randomized double-blind prospective trial of cognitive stimulation. *Neuroepidemiology*, 36, 91-99. doi:10.1159/000323950
- Princiotta, D., DeVries, M., & Goldstein, S. (2014). Executive functioning as a mediator of age-related cognitive decline in adults. In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 143-158). New York, NY: Springer.

- Reijnders, J., van Heugten, C., & van Boxtel, M. (2013). Cognitive interventions in healthy older adults and people with mild cognitive impairment: A systematic review. *Ageing Research Reviews*, 12, 263-275. doi:10.1016/j.arr.2012.07.003
- Reuter-Lorenz, P. A., Festini, S. B., & Jantz, T. K. (2016). Executive functions and neurocognitive aging. In K. W. Schaie & S. L. Willis (Eds.), *Handbook of the psychology of aging* (8th ed., pp. 245-262). San Diego, USA: Elsevier Science Publishing Co Inc.
- Reuter-Lorenz, P. A., & Park, D. C. (2014). How does it STAC up? Revisiting the scaffolding theory of aging and cognition. *Neuropsychology Review*, 24, 355-370. doi:10.1007/s11065-014-9270-9
- *Richmond, L. L., Morrison, A. B., Chein, J. M., & Olson, I. R. (2011). Working memory training and transfer in older adults. *Psychology and Aging*, 26, 813-822. doi:10.1037/a0023631
- Sala, G., Aksayli, N. D., Semir, K., Gondo, Y., & Gobet, F. (in press). Working memory training does not enhance older adults' cognitive skills: A meta-analysis. *PsyArXiv*. doi:10.31234/osf.io/5frzb
- Sala, G., Tatlidil, K. S., & Gobet, F. (2018). Video game training does not enhance cognitive ability: A comprehensive meta-analytic investigation. *Psychological Bulletin*, 144, 111-139. doi:10.1037/bul0000139
- *Salminen, T., Frensch, P., Strobach, T., & Schubert, T. (2016). Age-specific differences of dual n-back training. *Aging, Neuropsychology, and Cognition*, 23, 18-39. doi:10.1080/13825585.2015.1031723
- Salminen, T., Kühn, S., Frensch, P. A., & Schubert, T. (2016). Transfer after dual n-back training depends on striatal activation change. *Journal of Neuroscience Methods*, 36, 10198-10213. doi:10.1523/JNEUROSCI.2305-15.2016

- *Sandberg, P., & Stigsdotter Neely, A. (2016). Long-term effects of executive process training in young and old adults. *Neuropsychological Rehabilitation*, 26, 761-782. doi:10.1080/09602011.2015.1108205
- *Sandberg, P., Ronnlund, M., Nyberg, L., & Stigsdotter Neely, A. (2014). Executive process training in young and old adults. *Aging, Neuropsychology, and Cognition*, 21, 577-605. doi:10.1080/13825585.2013.839777
- Schmiedek, F. (2016). Methods and designs. In T. Strobach & J. Karbach (Eds.), *Cognitive training: An overview of features and applications* (pp. 9-18). Switzerland: Springer.
- *Schmiedek, F., Lövdén, M., & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Frontiers in Aging Neuroscience*, 2, 1-10. doi:10.3389/fnagi.2010.00027
- Schnitzspahn, K. M., Stahl, C., Zeintl, M., Kaller, C. P., & Kliegel, M. (2013). The role of shifting, updating, and inhibition in prospective memory performance in young and older adults. *Developmental Psychology*, 49, 1544-1553. doi:10.1037/a0030579
- Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist*, 50, 138-166. doi:10.1080/00461520.2015.1036274
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52, 591-611. doi:10.1093/biomet/52.3-4.591
- *Shatil, E. (2013). Does combined cognitive training and physical activity training enhance cognitive abilities more than either alone? A four-condition randomized controlled trial among healthy older adults. *Frontiers in Aging Neuroscience*, 5, 1-12. doi:10.3389/fnagi.2013.00008
- *Shatil, E., Mikulecka, J., Bellotti, F., & Bures, V. (2014). Novel television-based cognitive training improves working memory and executive function. *PLoS One*, 9, e101472.

doi:10.1371/journal.pone.0101472

Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica*, 50, 245-276.

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest*, 17, 103-186. doi:10.1177/1529100616661983

Simons, J. S., Schölvink, M. L., Gilbert, S. J., Frith, C. D., & Burgess, P. W. (2006). Differential components of prospective memory? Evidence from fMRI. *Neuropsychologia*, 44, 1388-1397. doi:10.1016/j.neuropsychologia.2006.01.005

*Simpson, T., Camfield, D., Pipingas, A., Macpherson, H., & Stough, C. (2012). Improved processing speed: online computer-based cognitive training in older adults. *Educational Gerontology*, 38, 445-458. doi:10.1080/03601277.2011.559858

*Smith, G. E., Housen, P., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., & Zelinski, E. M. (2009). A cognitive training program based on principles of brain plasticity: Results from the improvement in memory with plasticity-based adaptive cognitive training (IMPACT) study. *Journal of the American Geriatrics Society*, 57, 594-603. doi:10.1111/j.1532-5415.2008.02167.x

*Souders, D. J., Boot, W. R., Blocker, K., Vitale, T., Roque, N. A., & Charness, N. (2017). Evidence for narrow transfer after short-term cognitive training in older adults. *Frontiers in Aging Neuroscience*, 9, 1-10. doi:10.3389/fnagi.2017.00041

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

Spencer-Smith, M., & Klingberg, T. (2015). Benefits of a working memory training program for inattention in daily life: A systematic review and meta-analysis. *PloS One*, 10,

e0119522. doi:10.1371/journal.pone.0119522

*Stepankova, H., Lukavsky, J., Buschkuehl, M., Kopecek, M., Ripova, D., & Jaeggi, S. M.

(2014). The malleability of working memory and visuospatial skills: A randomized controlled study in older adults. *Developmental Psychology*, 50, 1049-1060.

doi:10.1037/a0034913

*Stern, Y., Blumen, H. M., Rich, L. W., Richards, A., Herzberg, G., & Gopher, D. (2011).

Space Fortress game training and executive control in older adults: A pilot intervention. *Aging, Neuropsychology, and Cognition*, 18, 653-677.

doi:10.1080/13825585.2011.613450

Sterne, J. A. C., Egger, M., & Moher, D. (2011). Chapter 10: Addressing reporting biases. In

J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* (5.1.0 ed.): The Cochrane Collaboration. Retrieved from

www.cochrane-handbook.org.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological*

tests: Administration, norms, and commentary (3rd revised ed.). New York, NY:

Oxford University Press Inc.

Taatgen, N. A. (2013). The nature and transfer of cognitive skills. *Psychological Review*, 120,

439-471. doi:10.1037/a0033138

Taatgen, N. A. (2016). Theoretical models of training and transfer effects. In T. Strobach & J.

Karbach (Eds.), *Cognitive training: an overview of features and applications* (pp. 19-32). Switzerland: Springer.

Tetlow, A. M., & Edwards, J. D. (2017). Systematic literature review and meta-analysis of

commercially available computerized cognitive training among older adults. *Journal of Cognitive Enhancement*, 1, 559-575. doi:10.1007/s41465-017-0051-2

Thorndike, E. L. (1906). *The principles of teaching: Based on psychology*. New York: A.G.

Seiler.

Toril, P., Reales, J. M., & Ballesteros, S. (2014). Video game training enhances cognition of older adults: A meta-analytic study. *Psychology and Aging, 29*, 706-716.

doi:10.1037/a0037507

*Toril, P., Reales, J. M., Mayas, J., & Ballesteros, S. (2016). Video game training enhances visuospatial working memory and episodic memory in older adults. *Frontiers in Human Neuroscience, 10*, 206-220. doi:10.3389/fnhum.2016.00206

*van Muijden, J., Band, G. P., & Hommel, B. (2012). Online games training aging brains: Limited transfer to cognitive control functions. *Frontiers in Human Neuroscience, 6*, 221-234. doi:10.3389/fnhum.2012.00221

*Van Vleet, T. M., DeGutis, J. M., Merzenich, M. M., Simpson, G. V., Zomet, A., & Dabit, S. (2016). Targeting alertness to improve cognition in older adults: A preliminary report of benefits in executive function and skill acquisition. *Cortex, 82*, 100-118. doi:10.1016/j.cortex.2016.05.015

Vermeij, A., Claassen, J. A., Dautzenberg, P. L., & Kessels, R. P. (2016). Transfer and maintenance effects of online working-memory training in normal ageing and mild cognitive impairment. *Neuropsychological Rehabilitation, 26*, 783-809. doi:10.1080/09602011.2015.1048694

*von Bastian, C. C., Langer, N., Jancke, L., & Oberauer, K. (2013). Effects of working memory training in young and old adults. *Memory & Cognition, 41*, 611-624. doi:10.3758/s13421-012-0280-7

*Walton, C. C., Kavanagh, A., Downey, L. A., Lomas, J., Camfield, D. A., & Stough, C. (2015). Online cognitive training in healthy older adults: A preliminary study on the effects of single versus multi-domain training. *Translational Neuroscience, 6*, 13-19. doi:10.1515/tnsci-2015-0003

- *Wang, M., Chang, C., & Su, S. (2011). What's cooking?—Cognitive training of executive function in the elderly. *Frontiers in Psychology, 2*, 228-239.
doi:10.3389/fpsyg.2011.00228
- *Wild-Wall, N., Falkenstein, M., & Gajewski, P. D. (2012). Neural correlates of changes in a visual search task due to cognitive training in seniors. *Neural Plasticity, 2012*, 1-11.
doi:10.1155/2012/529057
- *Wilkinson, A. J., & Yang, L. (2012). Plasticity of inhibition in older adults: Retest practice and transfer effects. *Psychology and Aging, 27*, 606-615. doi:10.1037/a0025926
- Wolf, D., Tüscher, O., Teipel, S., Mierau, A., Strüder, H., Drzezga, A., . . . Fellgiebel, A. (2018). Mechanisms and modulators of cognitive training gain transfer in cognitively healthy aging: Study protocol of the AgeGain study. *Trials, 19*, 337.
doi:10.1186/s13063-018-2688-2
- World Health Organization. (2017). Dementia (Fact Sheet). Retrieved from <http://www.who.int/mediacentre/factsheets/fs362/en/>
- *Xin, Z., Lai, Z. R., Li, F., & Maes, J. H. R. (2014). Near- and far-transfer effects of working memory updating training in elderly adults. *Applied Cognitive Psychology, 28*, 403-408. doi:10.1002/acp.3011
- Yuan, P., & Raz, N. (2014). Prefrontal cortex and executive functions in healthy adults: A meta-analysis of structural neuroimaging studies. *Neuroscience and Biobehavioral Reviews, 42*, 180-192. doi:10.1016/j.neubiorev.2014.02.005
- *Zajac-Lamparska, L., & Trempała, J. (2016). Effects of working memory and attentional control training and their transfer onto fluid intelligence in early and late adulthood. *Health Psychology Report, 4*, 41-53. doi:10.5114/hpr.2016.56846
- *Zelinski, E. M., Spina, L. M., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., & Smith, G. E. (2011). Improvement in memory with plasticity-based adaptive cognitive

training: Results of the 3-month follow-up. *Journal of the American Geriatrics Society*, 59, 258-265. doi:10.1111/j.1532-5415.2010.03277.x

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

Table 1

Definitions and Examples of Tasks for Each Cognitive Domain Included in Review

Domain	Definition	Examples of Tasks from Included Studies
Attention	Outcome measures tapping attention assess the ability to select specific information for further processing (Strauss, Sherman, & Spreen, 2006).	Oddball task, Useful Field of View, Attention Network Task
Fluid Intelligence	Fluid intelligence is defined as the ability to reason quickly, think abstractly, and problem-solve, independent of acquired knowledge (Happé, 2013).	Raven's Standard/Progressive Matrices, Cattell Culture Fair Test
Global EF	Global executive functioning measures the ability to control behavior and adapt to changing environmental demands (Karch & Kray, 2016).	Wisconsin Card Sorting Task, Tower of Hanoi
Language	Language proficiency refers to the degree of competency or capability in a given language (Farhady, 1982).	Nelson-Denny Reading Test
Memory	Memory is defined as the capacity to encode, store, and retrieve information or stimuli (Strauss et al., 2006).	Memory for <u>nonverbal stimuli</u> : Corsi Span (forwards), face memory Memory for <u>verbal stimuli</u> : Digit Span (forwards), Hopkins Verbal Learning Test, Rey Auditory Verbal Learning Test
Processing Speed	Processing speed refers to the time required to make a correct judgement about a stimulus (Owsley, 2013).	Digit Symbol, Pattern Comparison, Simple Reaction Time
Visuospatial Ability	Visuospatial ability requires the ability to identify and locate stimuli (Strauss et al., 2006).	Mental rotation

Table 2

Summary and Moderator Effect Sizes for Immediate Efficacy (Pre-test to Post-test) of Core-EF CCT Across All Outcomes

Groups/Subgroups	Outcomes								
	Trained	Near (overall)	Far (overall)	Attention	Executive Functioning	Fluid Intelligence	Memory	Speed	Visuospatial Ability
Summary Effect Size	1.00*** (24)	0.26*** (55)	0.22*** (57)	0.11 (24)	0.18*** (38)	0.13** (33)	0.22** (16)	0.18 (22)	0.37* (5)
Training Type									
Single-domain	1.08*** (20)	0.32*** (28)	0.13* (31)	0.06 (13)	0.16** (25)	0.11 (22)	0.04 (4)	0.17 (13)	0.12 (1)
Multidomain	0.66*** (4)	0.21*** (27)	0.29*** (26)	0.18 (11)	0.19*** (13)	0.16* (11)	0.27** (12)	0.26 (9)	0.40* (4)
Flexibility	0.77 (3)	0.33 (3)	0.09 (5)	-0.02 (3)	0.18 (4)	-0.01 (3)	-0.01 (2)	0.22 (2)	—
Inhibition	0.88** (2)	0.11 (4)	0.24 (6)	0.29 (5)	0.10 (6)	0.48 (3)	0.10 (2)	0.05 (2)	0.12 (1)
Working Memory	1.17*** (15)	0.37*** (21)	0.12 (20)	-0.09 (5)	0.20** (15)	0.08 (16)	—	0.20 (9)	—
Multi-cognitive	0.66*** (4)	0.15** (13)	0.27*** (14)	0.22 (3)	0.19** (9)	0.11 (5)	0.14 (6)	0.29 (7)	-0.16 (1)
Video Game	—	0.27*** (14)	0.30*** (12)	0.17 (8)	0.20 (4)	0.21 (6)	0.44** (6)	0.13 (2)	0.49** (3)
Adaptive									
Yes	1.12*** (18)	0.27*** (47)	0.23*** (47)	0.09 (20)	0.16*** (31)	0.13* (25)	0.27*** (13)	0.24* (17)	0.35 (4)
No	0.66*** (6)	0.13 (8)	0.20 (10)	0.30 (4)	0.32 (7)	0.16 (8)	-0.05 (3)	0.03 (5)	0.41 (1)
Location									
Home	1.08*** (6)	0.23*** (20)	0.19*** (21)	-0.04 (9)	0.17** (13)	0.08 (11)	0.15 (4)	0.24 (5)	0.23 (3)
Lab	0.97*** (18)	0.27*** (35)	0.23*** (36)	0.20* (15)	0.18** (25)	0.17*** (22)	0.27** (12)	0.15 (17)	0.66*** (2)
Mean Age									
60 to 69 years	0.98*** (18)	0.22*** (37)	0.17*** (39)	0.08 (16)	0.15* (24)	0.13* (22)	0.16 (11)	0.13 (16)	0.26 (4)
70+ years	1.06*** (6)	0.35*** (18)	0.27*** (18)	0.18 (8)	0.20** (14)	0.13 (11)	0.31** (5)	0.41* (6)	0.82** (1)
Control Type									
Active	0.87*** (11)	0.22*** (29)	0.20*** (32)	0.05 (12)	0.19*** (19)	0.04 (19)	0.24* (9)	0.21 (14)	0.54* (2)
Passive	1.11*** (13)	0.31*** (26)	0.22*** (25)	0.20* (12)	0.14 (19)	0.30*** (14)	0.19 (7)	0.15 (8)	0.16 (3)
Session Length (min)									
Up to 30	1.00*** (7)	0.26*** (21)	0.15** (21)	-0.13 (7)	0.16* (13)	0.15* (15)	0.25 (3)	0.01 (5)	0.34 (1)
Over 30	1.01*** (17)	0.26*** (34)	0.27*** (36)	0.22* (17)	0.19*** (25)	0.12 (18)	0.19** (13)	0.25* (17)	0.36 (4)
Number of Sessions									
Up to 12	1.02*** (15)	0.28*** (18)	0.20*** (24)	0.23 (8)	0.21** (19)	0.21** (17)	0.02 (4)	0.15 (15)	—
Over 12	0.97*** (9)	0.24*** (37)	0.22*** (33)	0.06 (16)	0.16** (19)	0.07 (16)	0.28*** (12)	0.24 (7)	0.37* (5)
Number of Weeks									
Up to 4 weeks	0.97*** (14)	0.33*** (21)	0.18** (25)	0.16 (9)	0.22** (19)	0.16* (19)	0.17 (4)	0.16 (11)	0.41 (1)
Over 4 Weeks	1.04*** (10)	0.22*** (34)	0.24*** (32)	0.09 (15)	0.15** (19)	0.12 (14)	0.23** (12)	0.27 (11)	0.35 (4)
Total Training (hr)									
Up to 12	1.02*** (20)	0.28*** (34)	0.20*** (37)	0.07 (13)	0.22*** (26)	0.17** (24)	0.20 (7)	0.14 (17)	—
Over 12	0.93*** (4)	0.22*** (21)	0.25*** (20)	0.16 (11)	0.14* (12)	0.80 (9)	0.19** (9)	0.28 (5)	0.37* (5)

Note. * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$, (—) represents no analysis due to limited number of studies. Number of studies included in each analysis reported in parentheses. Effect size reported as Hedge's g .

Figure 1. Flow chart summarizing study screening and selection process.

Figure 2. Forest plot displaying individual and summary effect size estimates for trained outcomes.

Figure 3. Forest plot displaying individual and summary effect size estimates for near-transfer outcomes, stratified by training type.

Figure 4. Forest plot displaying individual and summary effect size estimates for overall far-transfer outcomes, stratified by training type.

Figure 5. Forest plot displaying individual and summary effect size estimates for trained outcomes from pre-test to follow-up (long-term efficacy).

Figure 6. Forest plot displaying individual and summary effect size estimates for near-transfer outcomes from pre-test to follow-up (long-term efficacy).

Figure 7. Forest plot displaying individual and summary effect size estimates for far-transfer outcomes from pre-test to follow-up (long-term efficacy).

Appendix A

Summary Results of Sensitivity Analyses

Type of Model

Fixed: $g = 0.33$, $SE = 0.02$, 95% CI [0.28, 0.37], $p < .001$

Random: $g = 0.34$, $SE = 0.04$, 95% CI [0.27, 0.41], $p < .001$

Type of Control

Active control: $g = 0.29$, $SE = 0.05$, 95% CI [0.20, 0.38], $p < .001$

Passive control: $g = 0.41$, $SE = 0.05$, 95% CI [0.31, 0.52], $p < .001$

Mixed: $g = 0.34$, $SE = 0.04$, 95% CI [0.27, 0.41], $p < .001$

Risk of Bias

Low-risk: $g = 0.23$, $SE = 0.04$, 95% CI [0.15, 0.31], $p < .001$

High-risk: $g = 0.44$, $SE = 0.04$, 95% CI [0.37, 0.51], $p < .001$

Overall: $g = 0.34$, $SE = 0.04$, 95% CI [0.27, 0.41], $p < .001$

Pre-Post Correlation **$r = .00$**

Trained: $g = 0.81$, $SE = 0.08$, 95% CI [0.55, 0.86], $p < .001$

Near-transfer: $g = 0.18$, $SE = 0.04$, 95% CI [0.11, 0.24], $p < .001$

Far-transfer: $g = 0.17$, $SE = 0.04$, 95% CI [0.10, 0.24], $p < .001$

 $r = .60$

Trained: $g = 1.00$, $SE = 0.11$, 95% CI [0.79, 1.20], $p < .001$

Near-transfer: $g = 0.26$, $SE = 0.04$, 95% CI [0.17, 0.34], $p < .001$

Far-transfer: $g = 0.22$, $SE = 0.04$, 95% CI [0.15, 0.29], $p < .001$

 $r = .90$

Trained: $g = 1.58$, $SE = 0.17$, 95% CI [1.25, 1.91], $p < .001$

Near-transfer: $g = 0.47$, $SE = 0.08$, 95% CI [0.33, 0.61], $p < .001$

Far-transfer: $g = 0.36$, $SE = 0.06$, 95% CI [0.24, 0.47], $p < .001$

Inclusion of Borella et al. (2014) – Young-Old in the Immediate Near-Transfer Analysis

Table A1

Comparison of Summary and Moderator Effect Sizes for Near-Transfer Outcomes Including and Excluding Borella et al. (2014) – Young-Old

Groups/Subgroups	Study Included		Study Excluded	
	<i>g</i>	<i>Q</i> -statistic	<i>g</i>	<i>Q</i> -statistic
Summary effect size	0.26	$Q(54) = 68.52, p = .058$	0.23	$Q(53) = 52.02, p = .512$
Training Type				
Single-domain	0.32	$Q(1) = 1.66, p = .198$	0.26	$Q(1) = 0.42, p = .518$
Multidomain	0.21		0.21	
Flexibility training	0.33	$Q(4) = 4.25, p = .373$	0.33	$Q(4) = 3.00, p = .558$
Inhibition training	0.11		0.11	
WM training	0.37		0.31	
Multi-cog training	0.15		0.15	
Video game training	0.27		0.27	
Adaptive				
Yes	0.27	$Q(1) = 1.39, p = .239$	0.25	$Q(1) = 0.91, p = .340$
No	0.13		0.13	
Location				
Home	0.23	$Q(1) = 0.21, p = .591$	0.23	$Q(1) = 0.01, p = .908$
Lab	0.27		0.24	
Control Type				
Active	0.22	$Q(1) = 1.16, p = .282$	0.19	$Q(1) = 2.46, p = .117$
Passive	0.31		0.31	
Mean Age				
60 to 69 years	0.22	$Q(1) = 1.60, p = .206$	0.22	$Q(1) = 0.35, p = .553$
70+ years	0.35		0.27	
Session Length (min)				
Up to 30	0.26	$Q(1) = 0.00, p = .995$	0.26	$Q(1) = 0.17, p = .678$
Over 30	0.26		0.22	
Number of Sessions				
Up to 12	0.28	$Q(1) = 0.15, p = .694$	0.22	$Q(1) = 0.07, p = .792$
Over 12	0.24		0.24	
Number of Weeks				
Up to 4	0.33	$Q(1) = 1.35, p = .246$	0.27	$Q(1) = 0.40, p = .529$
Over 4	0.22		0.22	
Total Training (hr)				
Up to 12	0.28	$Q(1) = 0.58, p = .446$	0.25	$Q(1) = 0.12, p = .731$
Over 12	0.22		0.22	

Appendix B

Table B1

Summary of Sample, Intervention, and Design Characteristics for Studies Included in Both Quantitative and Qualitative Reviews

Study Characteristics			Intervention							Control Condition		Design	
Study	Sample <i>N</i>	<i>M</i> _{age} (<i>SD</i>); range	Name	Trained Domains	Session Length (min)	Sessions/ Weeks	Total Training (hrs)	Adaptive	Location	Type	Activity	Follow- Up (months)	Risk of Bias
Akimoto et al. (2016)	TG ₁ : 5 TG ₂ : 9 CG: 7 Total <i>N</i> : 21	67.92 (5.33); 60-75	In-house executive control CCT	WM, inhibition, flexibility, speed, attention	20	24s/8w	8	Yes	Lab	AC	Crossword puzzles	-	Low
Anderson et al. (2013a)	TG: 14 CG: 14 Total <i>N</i> : 28	61.86 (3.63); 55-79	Posit Science, Brain Fitness™	WM, auditory memory & processing	60	40s/8w	40	Yes	Home	AC	Educational DVD & quiz	-	Low
Anderson et al. (2013b)	TG: 35 CG: 32 Total <i>N</i> : 67	63.00 (3.70); 55-70	Posit Science, Brain Fitness™	WM, auditory memory & processing	60	40s/8w	40	Yes	Home	AC	Educational DVD & quiz	-	Low
Anguera et al. (2013)	TG: 16 AC: 15 PC: 15 Total <i>N</i> : 46	67.10 (4.20); 60-85	NeuroRacer	Multi-tasking, WM, switching, attention, speed	60	12s/4w	12	Yes	Lab	AC & PC	AC: single- task training	6	High
Ballesteros et al. (2014)	TG: 17 CG: 13 Total <i>N</i> : 30	67.92 (5.58); 57-80	Lumosity	WM, inhibition, memory, speed, visual	60	20s/10- 12w	20	Yes	Lab	PC	Converse with researcher	3— reported separately	High
Ballesteros et al. (2015)*	TG: 17 CG: 11 Total <i>N</i> : 28	-	-	-	-	-	-	-	-	-	-	3 months	-

Study Characteristics			Intervention							Control Condition		Design	
Study	Sample <i>N</i>	Age (<i>SD</i>); range	Name	Trained Domains	Session Length (min)	Sessions/ Weeks	Total Training (hrs)	Adaptive	Location	Type	Activity	Follow- Up (months)	Risk of Bias
Ballesteros et al. (2017)	TG: 30 CG: 25 Total <i>N</i> : 55	65.55 (5.20); 55-84	Lumosity	WM, inhibition, switching, speed	45	16s/10-12w	12	Yes	Lab	AC	Sim City (computer game)	-	Low
Basak et al. (2008)	TG:16 CG: 18 Total <i>N</i> : 34	69.41 (5.41)	Rise of Nations	WM, switching, STM, planning strategy	90	15s/4-5w	22.5	No	Lab	PC	No activity	1	High
Binder et al. (2016) ^a	TG _{inh} : 22 TG _{md} : 21 AC ₁ : 21 AC ₂ : 20 PC: 20 Total <i>N</i> : 104	69.49 (2.83); 64-75	In-house CCT: TG _{inh} : inhibition TG _{md} : multi-domain	TG _{inh} : inhibition TG _{md} : spatial navigation, inhibition visuomotor function	60	50s/10w	50	Yes	Home	AC & PC	AC ₁ : visuo-motor CCT AC ₂ : spatial navigation CCT	6	Low
Borella et al. (2014) ^a	TG _Y : 20 TG _O : 20 CG _Y : 20 CG _O : 20 Total <i>N</i> : 80	Y: 69.73 (2.81) 65-75 O: 79.65 (2.26); 76-84	In-house WM CCT (Matrix task)	WM	60	3s/2w	3	Yes	Lab	AC	Auto-biographical surveys	8	High

Study Characteristics			Intervention							Control Condition		Design	
Study	Sample <i>N</i>	Age (<i>SD</i>); range	Name	Trained Domains	Session Length (min)	Sessions/ Weeks	Total Training (hrs)	Adaptive	Location	Type	Activity	Follow- Up (months)	Risk of Bias
Brom & Kliegel (2014)	TG: 15 TG ₊₁ : 16 CG ₁ : 15 CG ₂ : 16 Total: 62	68.76 (5.23); 60-86	Task-switching CCT (Karch & Kray, 2009)	TG: Switching TG ₊₁ : Switching, implementation intention strategy (IMP)	25	5s/1w	2.08	No	Home	AC	CG ₁ : IMP only CG ₂ : simple intention	-	High
Buitenweg et al. (2017)	TG ₁ : 56 TG ₂ : 33 CG: 50 Total <i>N</i> : 139	67.75 (5.12); 60+	In-house multi-cog CCT	WM, attention, reasoning, logic	30	60s/12w	30	Yes	Home	AC	Mock CCT (speed, puzzles)	-	Low
Bürki et al. (2014)	TG: 22 AC: 20 PC: 23 Total <i>N</i> : 65	68.00 (5.44); 60-84	In-house WM CCT (<i>n</i> -back)	WM	30	10s/2-4w	5	Yes	Lab	AC & PC	AC: motor learning	-	High
Cantarella et al. (2017)	TG: 18 CG: 17 Total <i>N</i> : 35	68.45 (3.02); 63-73	In-house WM CCT (Matrix task)	WM	60	3s/2w	3	Yes	Lab	AC	Auto-biographical surveys	8	High
Chambon et al. (2014) ^b	TG: 15 AC: 15 PC: 15 Total <i>N</i> : 45	75.35 (10.00)	In-house memory and attention CCT	WM, attention, inhibition, dual-task, memory	60	24s/12w	24	Yes	Home	AC & PC	AC: pen-and-paper puzzles	6	High
Chiu et al. (2017)	TG: 31 CG: 31 Total <i>N</i> : 62	82.18 (5.15); 65+	RehaCom	Inhibition, set-shifting, WM	30	24s/8w	12	Yes	Lab	AC	Online newspapers, e-books, puzzles	3, 6	Low

Study Characteristics			Intervention							Control Condition		Design	
Study	Sample <i>N</i>	Age (<i>SD</i>); range	Name	Trained Domains	Session Length (min)	Sessions/ Weeks	Total Training (hrs)	Adaptive	Location	Type	Activity	Follow- Up (months)	Risk of Bias
Cuenen et al. (2016)	TG: 19 AC: 19 PC: 18 Total <i>N</i> : 56	71.21 (5.21); 60+	In-house WM CCT	WM	20	25s/5w	8.33	Yes	Home	AC & PC	AC: non- adaptive WM CCT	-	High
Dahlin et al. (2008a)	TG: 11 CG: 8 Total <i>N</i> : 19	68.32 (1.79); 65-71	In-house WM CCT	WM	45	15s/5w	11.25	Yes	Lab	PC	No activity	-	High
Dahlin et al. (2008b)	TG: 13 CG: 16 Total <i>N</i> : 29	68.31 (1.67)	In-house WM CCT	WM	45	15s/5w	11.25	Yes	Lab	PC	No activity	18	High
Gaál & Czigler (2017)	TG: 20 CG: 20 Total <i>N</i> : 40	65.70 (3.20); 60-75	In-house flexibility CCT (Informatively cued task)	Cognitive flexibility	60	8s/4w	8	Yes	Lab	PC	No activity	12	High
Gajewski & Falkenstein (2012)	TG: 32 AC _P : 35 AC _R : 34 PC: 40 Total <i>N</i> : 141	70.90 (5.20); 65+	CCT variety (Mental- Aktiv, Ahano); MAT, Sudoku	WM, attention, switching, speed, memory, reasoning, visuospatial	90	32s/16w	48	Yes	Lab	AC & PC	AC _P : Physical training AC _R : relaxation	-	Low
Gajewski & Falkenstein (in press)	TG: 32 AC _P : 35 AC _R : 34 PC: 40 Total <i>N</i> : 141	70.90 (5.20); 65+	CCT variety (Mental- Aktiv, Ahano); MAT, Sudoku	WM, attention, switching, speed, memory, reasoning, visuospatial	90	32s/16w	48	Yes	Lab	AC & PC	AC _P : Physical training AC _R : relaxation	-	Low

Study Characteristics			Intervention							Control Condition		Design	
Study	Sample <i>N</i>	<i>M_{age}</i> (<i>SD</i>); range	Name	Trained Domains	Session Length (min)	Sessions/ Weeks	Total Training (hrs)	Adaptive	Location	Type	Activity	Follow- Up (months)	Risk of Bias
Goghari & Lawlor-Savage (2017)	TG: 36 AC: 32 PC: 29 Total <i>N</i> : 97	70.48 (4.63) 65-86;	BrainGymmer	WM	30	40s/8w	20	Yes	Home	AC & PC	AC: logic & planning CCT	-	Low
Grönholm-Nyman et al. (2017)	TG: 17 CG: 16 Total <i>N</i> : 33	68.54 (7.38)	In-house set-shifting CCT	Switching	60	15s/5w	15	Yes	Lab	AC	Computer games	12	High
Guye & von Bastian (2017)	TG: 68 CG: 74 Total <i>N</i> : 142	70.35 (3.66); 65-80	In-house WM CCT	WM	30	25s/5w	12.5	Yes	Home	AC	Visual search CCT	-	Low
Heinzel et al. (2014)	TG: 15 CG: 15 Total <i>N</i> : 30	65.84 (4.25); 60-75	In-house WM CCT (<i>n</i> -back)	WM	45	12s/4w	9	Yes	Lab	PC	No activity	-	High
Heinzel et al. (2016)	TG: 15 CG: 14 Total <i>N</i> : 29	66.04 (4.35); 60-75	In-house WM CCT (<i>n</i> -back)	WM	45	12s/4w	9	Yes	Lab	PC	No activity	-	High
Heinzel et al. (2017)	TG: 18 CG: 16 Total <i>N</i> : 34	65.41 (3.32); 60-72	In-house WM CCT (<i>n</i> -back)	WM	45	12s/4w	9	Yes	Lab	PC	No activity	-	High
Hering et al. (2017)	TG: 29 AC: 29 PC: 30 Total <i>N</i> : 88	67.33 (4.81); 60-82	Computerized Categorization Working Memory Span task (Borella et al., 2010)	WM	45	4s/2w	3	Yes	Lab	AC & PC	Visual search	-	Low

Study Characteristics			Intervention							Control Condition		Design	
Study	Sample <i>N</i>	Mean (<i>SD</i>); range	Name	Trained Domains	Session Length (min)	Sessions/ Weeks	Total Training (hrs)	Adaptive	Location	Type	Activity	Follow- Up (months)	Risk of Bias
Ji et al. (2016)	TG: 18 CG: 16 Total <i>N</i> : 34	70.06 (5.53); 61-81	In-house inhibition CCT	Inhibition	60	12s/4w	12	Yes	Lab	PC	Health lectures	-	High
Karbach & Kray (2009)	TG ₁₋₄ : 42 CG: 14 Total <i>N</i> : 56	68.70 (3.00); 62-76	Task- switching (TS) CCT	TG ₁ : TS TG ₂ : TS + self- instruction TG ₃ : TS + self- instruction + variability	40	4s/4w	2.67	No	Lab	AC	Single-task training	-	High
Kim et al. (2017)	TG: 14 CG: 13 Total <i>N</i> : 27	71.44 (3.55); 64-77	In-house executive control CCT	Inhibition, WM, switching, dual-tasking	60	24s/8w	24	Yes	Lab	PC	No activity	-	High
Kray & Fehér (2017)	TG ₁ : 17 TG ₂ : 16 TG ₃ : 16 TG ₄ : 16 CG: 17 Total <i>N</i> : 82	70.80; 65-84	Task- Switching (TS) CCT (Karbach & Kray, 2009)	TG ₁ : TS+ cue, univalent TG ₂ : TS + cue, bivalent TG ₃ : TS no cue, univalent TG ₄ : TS no cue, bivalent	40	4s/4w	2.67	No	Lab	AC	Single-task training	6	High
Küper et al. (2017)	TG: 32 AC: 34 PC: 37 Total <i>N</i> : 108	70.47; 65-88	CCT variety (Mental- Aktiv, Ahano); MAT, Sudoku	WM, attention, switching, speed, memory, reasoning, visuospatial	90	32s/16w	48	Yes	Lab	AC & PC	Relaxation training	-	High

Study Characteristics			Intervention							Control Condition		Design	
Study	Sample <i>N</i>	<i>M_{age}</i> (<i>SD</i>); range	Name	Trained Domains	Session Length (min)	Sessions/ Weeks	Total Training (hrs)	Adaptive	Location	Type	Activity	Follow- Up (months)	Risk of Bias
Lange & Süß (2015)	TG: 31 AC: 31 PC: 29 Total <i>N</i> : 91	67.97 (4.04); 60-76	In-house WM CCT	WM	60	12s/6w	12	Yes	Lab	AC & PC	Computer games	-	High
Li et al. (2016)	TG: 20 AC: 19 PC: 14 Total <i>N</i> : 53	68.30 (4.40); 60-76	In-house memory and executive control CCT	WM, switching, memory	60	16s/6w	16	Yes	Lab	AC & PC	Memory strategy training	4	Low
Li et al. (2008)	TG: 21 CG: 20 Total <i>N</i> : 41	73.91 (2.83); 70-80	In-house WM CCT	WM	15	45s/11w	11.25	No	Lab	PC	No activity	3	High
Mozolic et al. (2011)	TG: 30 CG: 32 Total <i>N</i> : 62	69.40 (2.84); 65-75	In-house attention CCT	Inhibition, attention	60	8s/8w	8	Yes	Lab	PC	Educational lectures	-	Low
Nozawa et al. (2015)	TGv: 12 TGp: 11 CG: 12 Total <i>N</i> : 35	67.90 (4.87); 60-75	In-house executive control CCT	WM, inhibition, flexibility, speed, attention	20	24s/8w	8	Yes	Lab	AC	Crossword puzzles	-	Low
Payne et al. (2017)	TG: 18 CG: 22 Total <i>N</i> : 40	67.92 (4.94); 60+	In-house WM CCT (iTrain)	WM	30	15s/3w	7.5	Yes	Home	AC	Decision speed CCT	-	Low
Peretz et al. (2011)	TG: 66 CG: 55 Total <i>N</i> : 121	67.50 (7.13); 50+	CogniFit Personal Coach	WM, shifting, inhibition, attention, memory, speed, visual	30	36s/12w	18	Yes	Home	AC	Computer games	-	Low

Study Characteristics			Intervention							Control Condition		Design	
Study	Sample <i>N</i>	Mean (<i>SD</i>); range	Name	Trained Domains	Session Length (min)	Sessions/ Weeks	Total Training (hrs)	Adaptive	Location	Type	Activity	Follow- Up (months)	Risk of Bias
Richmond et al. (2011)	TG: 21 CG: 19 Total <i>N</i> : 40	66.00; 60-80	In-house WM CCT		30	20s/5w	10	Yes	Home	AC	Trivia (online quizzes)	-	Low
Salminen et al. (2016)	TG: 25 CG: 21 Total <i>N</i> : 46	64.89 (3.71); 57-77	Dual <i>n</i> -back	WM	50	14s/3-6w	11.67	Yes	Lab	PC	No activity	-	Low
Sandberg et al. (2014)	TG: 15 CG: 15 Total <i>N</i> : 30	69.27 (4.85)	In-house executive control CCT	WM, inhibition, switching	45	15s/5w	11.25	No	Lab	PC	No activity	18— reported separately	Low
Sandberg & Stigsdotter Neely (2016)*	TG: 14 CG: 10 Total <i>N</i> : 24	71.43 (5.03)	-	-	-	-	-	-	-	-	-	18 months	-
Schmiedek et al. (2010)	TG: 103 CG: 39 Total <i>N</i> : 142	71.11 (4.07); 65-81	In-house CCT (COGITO)	WM, memory, speed	60	100s/~27w	100	No	Lab	PC	No activity	-	High
Shatil (2013)	TG: 33 TG+P: 29 AC _P : 29 AC _{BC} : 31 Total <i>N</i> : 122	79.75 (5.48)	CogniFit Personal Coach TG+P: CCT + physical training	WM, shifting, inhibition, attention, memory, speed, visual	40	48s/16w	32	Yes	Lab	AC	AC _P : Physical training AC _{BC} : Book club	-	High

Study Characteristics			Intervention							Control Condition		Design	
Study	Sample <i>N</i>	Age (<i>SD</i>); range	Name	Trained Domains	Session Length (min)	Sessions/ Weeks	Total Training (hrs)	Adaptive	Location	Type	Activity	Follow- Up (months)	Risk of Bias
Shatil et al. (2014)	TG: 60 CG: 59 Total <i>N</i> : 119	68.00 (5.78); 60-87	CogniFit Personal Coach (iTV)	WM, shifting, inhibition, attention, memory, speed, visual	20	24s/8w	8	Yes	Lab	AC	Leisure activities (e.g., family trees)	-	High
Simpson et al. (2012)	TG: 17 CG: 14 Total: 31	62.29 (5.53); 53-75	My Brain Trainer	WM, inhibition, switching, attention, speed, arithmetic	20	21s/3w	7	Yes	Home	AC	Solitaire	3 weeks	Low
Smith et al. (2009)	TG: 242 CG: 245 Total <i>N</i> : 487	75.30 (6.45); 65+	Posit Science, Brain Fitness™	WM, auditory memory & processing	60	40s/8w	40	Yes	Home	AC	Educational DVD & quiz	3— reported separately	Low
Zelinski et al. (2011)*	TG: 211 CG: 204 Total <i>N</i> : 415	-	-	-	-	-	-	-	-	-	-	3 months	-
Souders et al. (2017)	TG: 30 PC: 30 Total <i>N</i> : 60	72.35 (5.20); 65+	Mind Frontiers	WM, switching, planning, memory, reasoning	45	30s/4w	22.5	Yes	Home	AC	Puzzles (e.g., sudoku, crosswords)	-	Low
Stern et al. (2011)	TG ₁ : 17 TG ₂ : 18 CG: 19 Total <i>N</i> : 54	66.30 (4.21)	TG ₁ : Space Fortress + instruction TG ₂ : Space Fortress	Attention, multi-tasking, WM, memory, inhibition, flexibility	60	36s/12w	36	No	Lab	PC	No activity	3	Low

Study Characteristics			Intervention							Control Condition		Design	
Study	Sample <i>N</i>	Mean (<i>SD</i>); range	Name	Trained Domains	Session Length (min)	Sessions/ Weeks	Total Training (hrs)	Adaptive	Location	Type	Activity	Follow- Up (months)	Risk of Bias
Stepankova et al. (2014) ^a	TG ₁ : 20 TG ₂ : 20 CG: 25 Total <i>N</i> : 65	68.10 (2.60); 65-74	<i>n</i> -back	WM	25	TG ₁ : 10s/5w TG ₂ : 20s/5w	TG ₁ : 4.17 TG ₂ : 8.33	Yes	Home	PC	No activity	-	Low
Toril et al. (2016)	TG: 19 CG: 20 Total <i>N</i> : 39	71.62 (6.72)	Lumosity	WM, inhibition, speed, memory, visuospatial	60	15s/7-8w	15	Yes	Lab	PC	Converse with researcher	3	High
Van Muijden et al. (2012)	TG: 53 CG: 19 Total <i>N</i> : 72	67.64 (3.68)	In-house executive control CCT	WM, switching, language, arithmetic	30	49s/7w	24.5	Yes	Home	AC	Document- ary & quiz	-	Low
Van Vleet et al. (2016)	TG: 11 CG: 10 Total <i>N</i> : 21	76.05 (7.85)	Tonic and Phasic Alertness Training (TAPAT)	Attention, inhibition	36	9s/3w	5.4	No	Lab	AC	Serial categorization CCT	-	High
von Bastian et al. (2013)	TG: 27 CG: 30 Total <i>N</i> : 57	68.42 (3.28); 61-77	In-house WM CCT	WM, switching	30	20s/4w	10	Yes	Home	AC	Quiz, visual search, and counting tasks	-	Low
Walton et al. (2015)	TG: 16 CG: 12 Total <i>N</i> : 28	64.18 (6.90); 55-78	My Brain Trainer	WM, inhibition, switching, attention, speed, arithmetic	20	28s/7w	9.33	No	Home	AC	Simple reaction time task	-	Low
Wang et al. (2011)	TG: 26 CG: 26 Total <i>N</i> : 52	64.20 (6.63); 56-84	In-house executive control video game	WM, switching, multi-tasking	30-60	5s/5w	5	Yes	Lab	PC	No activity	-	High

Study Characteristics			Intervention							Control Condition		Design	
Study	Sample <i>N</i>	<i>M</i> _{age} (<i>SD</i>); range	Name	Trained Domains	Session Length (min)	Sessions/ Weeks	Total Training (hrs)	Adaptive	Location	Type	Activity	Follow- Up (months)	Risk of Bias
Wild-Wall et al. (2012)	TG: 32 AC: 34 PC: 39 Total <i>N</i> : 105	70.33 (4.40); 65-88	CCT variety (Mental- Aktiv, Ahano); MAT, Sudoku	WM, attention, switching, speed, memory, reasoning, visuospatial Inhibition	90	32s/16w	48	Yes	Lab	AC & PC	Relaxation training	-	Low
Wilkinson & Yang (2012)	TG: 42 CG: 14 Total <i>N</i> : 56	71.05 (6.17); 60-84	In-house inhibition CCT (Stroop)		30	6s/2w	3	No	Lab	PC	No activity	-	High
Xin et al. (2014)	TG: 15 CG: 14 Total <i>N</i> : 29	69.53 (6.31); 60-82	In-house WM CCT	WM	30	20s/3-4w	10	Yes	Lab	AC	Computer games	-	High
Zajac- Lamparska & Trempala (2016) ^a	TG _{WM} : 30 TG _{ATT} : 30 CG: 30 Total <i>N</i> : 90	67.71 (5.57); 60-85	TG _{WM} : In- house WM CCT / TG _{ATT} : In- house attention CCT	WM / attention, inhibition	20	7s/4w	2.33	No	Home	PC	No activity	-	High
Zinke et al. (2014)	TG: 40 CG: 40 Total <i>N</i> : 80	77.20 (8.10); 65-95	In-house executive control CCT	WM, executive control	30	3s/3w	4.5	Yes	Lab	PC	No activity	9	High

Note. *Study reports results for follow-up only (these studies present follow-up data published in a separate article from an original training study meeting all eligibility criteria). Follow-up only studies are presented below the original study. ^a Study entered as two separate comparisons in the meta-analysis. ^b Study included only in review but not in meta-analysis due to insufficient data provided. TG = training group, CG = control group, AC = active control, PC = passive control, CCT = computerized cognitive training, WM = working memory.

Table B2

Summary of Overall Effect Sizes (Hedge's g) and Sample Sizes for Studies Included in the Quantitative Review

Study	Hedge's g (pretest to posttest)	Pretest–posttest sample size training (control)	Hedge's g (pretest to followup)	Pretest–follow-up sample size training (control)
Akimoto et al. (2016)	-0.23	14 (7)	—	—
Anderson et al. (2013a)	0.29	14 (14)	—	—
Anderson et al. (2013b)	0.67	35 (32)	—	—
Anguera et al. (2013)	0.30	16 (15)	—	—
Ballesteros et al. (2014)	0.15	17 (13)	Reported in Ballesteros et al. (2015)	
Ballesteros et al. (2015)	—	—	-0.20	17 (11)
Ballesteros et al. (2017)	0.00	30 (25)	—	—
Basak et al. (2008)	0.17	19 (19)	0.35	19 (19)
Binder et al. (2016) – inhibition training	0.14	22 (10)	—	—
Binder et al. (2016) – multidomain training	0.12	21 (10)	—	—
Borella et al. (2014) – old-old	0.62	20 (20)	0.57	20 (20)
Borella et al. (2014) – young-old	1.31	20 (20)	1.04	20 (20)
Brom & Kliegel (2014)	0.17	31 (31)	—	—
Buitenweg et al. (2017)	0.01	89 (50)	—	—
Bürki et al. (2014)	0.21	22 (20)	—	—
Cantarella et al. (2017)	0.36	18 (17)	0.51	18 (17)
Chiu et al. (2017)	0.17	31 (31)	0.16	31 (31)
Cuenen et al. (2016)	0.46	19 (18)	—	—
Dahlin et al. (2008a)	1.03	11 (8)	—	—
Dahlin et al. (2008b)	0.16	13 (16)	-0.04	13 (7)
Gaál & Czigler (2017)	0.56	20 (20)	0.35	20 (20)
Gajewski & Falkenstein (2012)	0.27	32 (40)	—	—
Goghari & Lawlor-Savage (2017)	-0.01	36 (32)	—	—
Grönholm-Nyman et al. (2017)	0.46	17 (16)	0.41	17 (15)
Guye & von Bastian (2017)	0.38	68 (74)	—	—
Heinzel et al. (2014)	0.94	15 (15)	—	—

Study	Hedge's <i>g</i> (pretest to posttest)	Pretest–posttest sample size training (control)	Hedge's <i>g</i> (pretest to followup)	Pretest–follow-up sample size training (control)
Heinzel et al. (2016)	0.53	16 (16)	—	—
Heinzel et al. (2017)	0.50	18 (16)	—	—
Hering et al. (2017)	0.20	29 (29)	—	—
Ji et al. (2016)	0.36	18 (16)	—	—
Karbach & Kray (2009)	0.21	42 (14)	—	—
Kim et al. (2017)	0.41	14 (13)	—	—
Kray & Fehér (2017)	0.01	65 (17)	0.11	65 (17)
Lange & Süß (2015)	0.36	31 (31)	—	—
Li et al. (2008)	0.90	12 (12)	—	—
Li et al. (2016)	0.14	20 (19)	0.02	17 (14)
Mozolic et al. (2011)	0.02	30 (32)	—	—
Nozawa et al. (2015)	-0.15	23 (12)	—	—
Payne et al. (2017)	0.50	22 (18)	—	—
Peretz et al. (2011)	0.19	66 (55)	—	—
Richmond et al. (2011)	-0.16	21 (19)	—	—
Salminen et al. (2016)	0.60	25 (21)	—	—
Sandberg et al. (2014)	-0.07	15 (15)	Reported in Sandberg & Stigsdotter Neely (2016)	
Sandberg & Stigsdotter Neely (2016)	—	—	-0.11	14 (10)
Schmiedek et al. (2010)	0.37	103 (39)	—	—
Shatil (2013)	0.45	33 (29)	—	—
Shatil et al. (2014)	0.47	60 (59)	—	—
Simpson et al. (2012)	0.05	17 (14)	0.21	17 (14)
Smith et al. (2009)	0.31	242 (245)	Reported in Zelinski et al. (2011)	
Zelinski et al. (2011)	—	—	0.33	211 (204)
Souders et al. (2017)	0.20	30 (30)	—	—
Stepankova et al. (2014) – short	0.63	20 (13)	—	—
Stepankova et al. (2014) – long	0.78	20 (12)	—	—
Stern et al. (2011)	0.14	40 (20)	0.07	35 (19)
Toril et al. (2016)	0.92	19 (20)	0.52	19 (20)
Van Muijden et al. (2012)	0.27	53 (19)	—	—

Study	Hedge's <i>g</i> (pretest to posttest)	Pretest–posttest sample size training (control)	Hedge's <i>g</i> (pretest to followup)	Pretest–follow-up sample size training (control)
Van Vleet et al. (2016)	1.44	11 (10)	—	—
von Bastian et al. (2013)	0.11	27 (30)	—	—
Walton et al. (2015)	0.27	16 (12)	—	—
Wang et al. (2011)	0.21	26 (26)	—	—
Wilkinson & Yang (2012)	0.00	42 (14)	—	—
Xin et al. (2014)	0.66	15 (14)	—	—
Zajac-Lamparska & Trempala (2016) – WM	0.67	30 (15)	—	—
Zajac-Lamparska & Trempala (2016) – inhibition	0.72	30 (15)	—	—
Zinke et al. (2014)	0.41	40 (40)	0.23	33 (18)

Appendix C

Intercorrelations of the Regression Coefficient Estimates Between Each Moderator Variable

Table C1

A Correlation Matrix of Regression Coefficient Estimates Between Each Moderator Variable

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Inhibition training	—												
2. Multi-cog training	.60	—											
3. Video game training	.60	.77	—										
4. WM training	.65	.77	.76	—									
5. Adaptive (yes)	-.13	-.28	-.26	-.38	—								
6. Location (lab)	.02	-.11	.06	-.05	.04	—							
7. Control type (PC)	-.36	-.13	-.15	-.26	.12	-.03	—						
8. Risk of bias (low)	-.13	-.16	-.05	-.09	-.17	.46	.22	—					
9. Mean age (of sample)	-.14	-.10	-.03	-.07	-.03	-.03	.15	.25	—				
10. Session length (min)	.02	.05	-.07	.13	-.36	-.35	.04	-.18	-.04	—			
11. No. sessions	-.04	-.09	-.11	-.06	-.23	.16	.23	-.02	-.06	.67	—		
12. No. weeks	.01	.05	-.19	.08	-.04	-.55	-.14	-.20	-.04	.20	-.28	—	
13. Total training (hrs)	.03	-.05	.14	-.03	.31	.23	-.18	.12	.04	-.82	-.79	-.29	—