Cognitive Training Does Not Enhance General Cognition

Giovanni Sala\textsuperscript{1} and Fernand Gobet\textsuperscript{2}

\textsuperscript{1}Osaka University
\textsuperscript{2}University of Liverpool

Author’s Note

Corresponding author: Fernand Gobet, Department of Psychological Sciences, University of Liverpool, UK. E-mail: fgobet@liverpool.ac.uk

Giovanni Sala is a JSPS International Research Fellow.

Keywords: cognitive ability; cognitive training; intelligence; meta-analysis; music; neural plasticity; video game; working memory

Abstract

Due to potential theoretical and societal implications, cognitive training has been one of the most influential topics in psychology and neuroscience. The assumption behind cognitive training is that one’s general cognitive ability can be enhanced by practicing cognitive tasks or intellectually demanding activities. The hundreds of studies published so far have provided mixed findings and systematic reviews have reached inconsistent conclusions. To resolve these discrepancies, we carried out several meta-analytic reviews. The results are highly consistent across all the reviewed domains: minimal effect on domain-general cognitive skills. Crucially, the observed between-study variability is accounted for by design quality and statistical artefacts. The cognitive-training program of research has showed no appreciable benefits, and other more plausible practices to enhance cognitive performance should be pursued.
Cognitive Training: A Controversy?

General cognitive ability (GCA; see Glossary), or, more prosaically, intelligence, is a well-established scientific construct in psychology [1, 2] (Box 1). GCA is positively associated with a large number of socially relevant outcomes such as educational and professional achievement [3, 4], income [5] and even longevity [6]. Obviously, intelligent people are a fundamental resource for society, especially if they can preserve their skills throughout their whole life.

Given the stakes at play, the last two decades have seen an impressive effort towards designing and implementing cognitive-training programs for enhancing GCA and slowing down its decline in the elderly. So far, the results have been inconsistent. Researchers are yet to reach an agreement in what has been called the “cognitive-training controversy.” This article aims to show that this so-called controversy can be easily resolved through the appropriate extraction, analysis, and interpretation of the empirical data. The apparently contradictory results are, in fact, highly consistent in all the reviewed domains: cognitive-training programs, while they do improve the skills that are being trained, do not enhance GCA or any of its components.

Theories of Cognitive Training

Substantial research into education and the psychology of expertise has established that skill acquisition is largely based on perceptual and conceptual domain-specific knowledge [7, 8]. Due to this specificity, the generalization of such knowledge across different domains, also called “far transfer”, rarely occurs [9]. Moreover, the more specialized the skill, the less the overlap between skills, and the more difficult the transfer will be [10].

If domain-specific skills cannot be transferred, how can a specific activity impact on GCA or any domain-general cognitive skill? Depending on the type of training, several explanations have been proposed. For example, working memory (WM) training has been claimed to boost fluid intelligence because these two cognitive constructs have a shared
capacity constraint [11]. Action video games have been suggested to improve one’s probabilistic inference and, in turn, visual-attentional skills [12]. Chess and music require a broad range of cognitive skills such as focused attention, reasoning, and WM. Therefore, practicing these activities might positively influence GCA [13]. Finally, brain-training research provides several explanations ranging from generic appeals to neuroplasticity to more elaborated hypotheses about the effects of sensory-discrimination training on information processing. (A detailed analysis of all these theoretical models is beyond the scope of this opinion article; for reviews, see [14, 15].)

There is a common denominator across these explanations: the idea that the enhancement of domain-general cognitive mechanisms is a by-product of domain-specific training [16]. In line with the research on skill acquisition and expertise, engaging in cognitive-training programs improves performance on the trained task and similar tasks. However, these activities are also believed to boost GCA or, at least some of its fundamental components such as fluid intelligence, memory, and processing speed (Box 1). Once improved, cognitive skills foster professional and academic domain-specific abilities that depend on them. In this program of research, neural plasticity represents the crucial mediator of this process [17]. Cognitive-training regimens are thought to produce functional and anatomical changes in the neural system that, in turn, explain the improvements in cognitive function. In accordance with this hypothesis, the brain of experts such as professional musicians, chess masters, and assiduous video-game players exhibits specific functional and anatomical neural patterns.

In the present article, we focus on the empirical evidence regarding the effects of five types of cognitive-training programs: WM training, video-game training, music instruction, chess instruction, and brain training. To date, these are the most studied and debated cognitive-training programs. We also briefly review other less-researched cognitive-training programs such as executive-functions training, cognitive-flexibility training, mindfulness/meditation training, exergame training, and spatial training. Finally, it must be noted that this article does not deal with non-cognitive interventions aimed at
optimizing cognitive function, such as lifestyle modifications [e.g. 18], health and fitness activities [e.g. 19], and drugs [e.g. 20].

**Correlational and Cross-Sectional Studies**

Before reviewing the experimental data, we briefly discuss the evidence based on correlational and cross-sectional designs on playing video-games, music, and chess. While such studies do not allow causal inferences to be drawn, they might be suggestive of a possible link between cognitive ability and cognitive training.

Researchers strongly disagree on the role played by GCA in video-game skill [15, 21-23]. While many small experiments have found that video-game players outperform non-players in a broad range of cognitive tests [e.g. 24, 25], some large studies have failed to replicate these results [26]. It is even doubtful whether, within the population of video-game players, video-game skill correlates with cognitive skills [23].

Our meta-analysis shows that the inconsistency of the results is only apparent [27]. To begin with, the correlation between video-game skill and cognitive skills is moderated by the way video-game skill is measured. Many studies use the number of hours per week of video-game practice as a proxy for video-game skill. However, how much one performs an activity is not necessarily a good measure of how good one is at this activity [7], and this applies to video-game playing as well. Performance assessed by video-game scores is a more reliable measure. In fact, our meta-analysis shows that, although no correlation is observed between frequency of video-game practice and any cognitive skill, there is a moderate correlation between video-game performance and some cognitive skills (e.g., spatial ability) in particular types of video games (e.g., non-action). However, there is no overall correlation between video-game skill and GCA. (For ease of exposition, this article will use “near-zero” to refer to effects sizes larger than -0.05 or less than 0.05, “small” to refer to effect sizes between 0.05 and 0.20, and “moderate” to refer to effect sizes between 0.20 and 0.50.)
Regarding cross-sectional evidence (i.e., players vs. non-players), the results are, once again, highly consistent despite disagreement among researchers. All the large meta-analyses on the topic concur that video-game players outperform non-players in many different cognitive tasks [27-29]. Where the few discrepancies occur, it is about the size of the difference between players and non-players. Our meta-analysis shows that, once corrected for publication bias (Box 2), the estimates should be relatively small [27].

With respect to musical skill, cross-sectional and correlational studies have consistently shown a clear relationship with superior GCA. Musicians often outperform non-musicians not only in cognitive tasks related to their domain of expertise (e.g., recall of notes and chord discrimination [30]), but also in a broad range of domain-general cognitive skills such as memory, fluid intelligence, and processing speed [31-34]. Thus, it is natural to postulate a possible causal link between engaging in music and superior GCA.

Finally, studies on chess with a correlational or cross-sectional design provide strong evidence for a positive relationship between playing chess and cognitive skills. Chess masters consistently show an advantage over novices on recall and detection tasks with domain-specific material [8, 35]. In addition, in line with what has been found in music, chess skill predicts superior domain-general cognitive ability, as revealed by two meta-analyses. The first [36] showed that chess skill, often measured by the Elo rating, positively correlates with a broad range of measures of cognitive skills. The second [37] found that chess players exhibit overall superior cognitive skills when compared to non-chess players. Notably, this pattern of results remains even when the participants’ educational level is controlled for.

Under the Meta-Analytic Lens: Experimental Studies

In true experiments on cognitive training, an experimental group carries out specific tasks (e.g., computer games practising working memory) and its performance on pre-tests and post-tests is compared with that of (ideally) both a passive and an active control group. When participants are randomly allocated to the groups, this design makes it possible to
draw strong inferences about causality – it is the specific intervention, and not unspecific effects such as placebo effects, that leads to an improvement in post-tests.

The experimental evidence regarding the influence of cognitive-training programs on cognitive function has been inconsistent. While some authors have reported data upholding the idea that such programs enhance cognition, some others have claimed the opposite. To resolve these discrepancies, independent laboratories have carried out dozens of systematic reviews and, most importantly, meta-analyses (Box 2). Regrettably (and ironically), these reviews and meta-analyses have sometimes provided opposite views about the actual effects of cognitive-training programs. In this section, we summarize these findings and argue that the observed inconsistencies (when any) are mainly due to mistakes in the interpretation of the results, placebo effects, and systematic biases in the statistical modelling of the data.

**Working-Memory Training**

WM is the cognitive system used to store and manipulate the information needed to perform cognitive tasks [38]. WM capacity – the amount of information WM can handle – is correlated with measures of fluid and general intelligence, cognitive control, and mathematical and reading abilities [39, 40]. Due to its fundamental role in cognitive function, it has been claimed that increasing WM capacity through training can foster several domain-general cognitive skills and, hence, GCA. To date, WM training is certainly the most studied and discussed type of cognitive-training program.

The most influential and comprehensive meta-analytic review in the field [41] is sceptical about the alleged benefits of WM training. While WM training seems to improve performance on memory tasks, no appreciable effect is observed in measures of verbal and non-verbal intelligence. The lack of generalized benefits is particularly evident when the treated groups are compared to active control groups in order to rule out potential placebo effects.

This pattern of results has been replicated in two other meta-analyses examining the effects of WM training on healthy adults [42] and typically developing children and young
adolescents [43], respectively. Notably, these meta-analyses have found near-zero effects not only on tests of intelligence, but also on measures of cognitive control and academic skills. Another smaller meta-analysis has reported similar outcomes in the population of children and adolescents with learning disabilities [44]. Finally, another meta-analysis [45] has found that transcranial direct-current stimulation (tDCS) does little to boost the effects of WM training on performance in measures of GCA or memory.

Other authors show more optimism about the cognitive benefits of WM training, at least with regard to specific training regimens and populations. Another meta-analysis reported that practicing n-back tasks – a particular sub-category of WM tasks – exerts a small, yet significant, effect on fluid intelligence [46]. However, a re-analysis of the original data shows that the training effects are close to zero in studies including active control groups [47]. (For a detailed discussion, see [47-49]). The heterogeneity (Box 2) observed in the field is thus explained by the type of control group implemented in the studies. Therefore, there is again no evidence that the benefits of this particular cognitive-training regimen go beyond placebo effects. Lastly, the effects of WM training on older adults’ GCA have been the subject of a lively debate [48, 50]. A recent large investigation suggests that, once again, the effects of WM memory on older adults’ cognitive performance are small and limited to memory tasks [51].

**Video-Game Training**

Since their worldwide diffusion in the eighties, video games have represented one of the most popular leisure activities among young and adult populations. Due to their societal relevance, the impact of playing commercial video games on human behaviour is a topic of major interest. Like other cognitive-training activities, playing video games is cognitively demanding. Puzzle games such as Tetris require spatial ability and processing speed. Action video games such as Call of Duty involve quick and accurate visual-attentional processing. Finally, strategy games such as Rise of Nations demand planning. Thus, it seems reasonable to propose that playing such games would improve cognitive function.
In spite of a large number of experimental studies, the alleged positive effects of video-game training on cognitive skills are much debated. Since the publication of a seminal study in 2003 [12], several experiments [e.g. 52] suggesting that playing action video-games enhances cognitive skills such as cognitive control and visual attention have been carried out. However, the initial positive results have rarely been replicated [28]. Mixed findings have also been reported about non-action video-game training [e.g. 53].

Several meta-analyses have concluded in the past five years that video-game training exerts a moderate effect on diverse cognitive skills [29, 54, 55]. However, we have showed [27] that the observed overall positive effects found in these meta-analyses are probably unreliable due to flaws in the modelling approach, such as the use of suboptimal formulas to calculate effect sizes and sampling error variances, the lack of an appropriate sensitivity analysis (e.g., examination of study-design quality as a potential moderator), and inadequate publication bias analysis. In fact, our own meta-analytic investigations suggest that the true effect of video-game training on GCA is close to zero, if not null. This negative conclusion is valid regardless of the type of cognitive tests, population age, and video-game genre. Finally, another recent meta-analysis [28] has claimed that action video-game training exerts small to medium effects on cognitive skills such as spatial cognition and attention. However, this meta-analysis reports a highly asymmetrical distribution of the observed effects. As correctly acknowledged by the authors, this condition suggests that the overall effect is inflated by publication bias.

**Music Instruction**

Despite substantial inconsistency in the results reported in the primary studies, several reviews of the literature have cautiously suggested that music does exert a positive impact on GCA and cognitive skills [13, 56]. The experimental inconsistencies in the empirical evidence are, once again, only apparent and disappear when the data are correctly analysed. Contrary to the previous meta-analyses, our meta-analysis [30] estimated only a small overall effect of music instruction on children’s cognitive skills and academic achievement. Critically, our analyses show that the observed variability found in the
primary studies is mostly due to the type of control group (active or passive). When the music-trained group is compared to an active control group, the effects are close to zero or null. Thus, just like WM training, the degree of true heterogeneity (Box 2) observed in the field is accounted for by the type of control group. Therefore, the benefits of music training do not go beyond placebo effects.

Corroborating our findings, a recent study [57] has shown that music-trained twins have the same IQ as their non-music-trained co-twins. In the same vein, another investigation [33] has found that, while musicians’ GCA correlates with music aptitude, it does not correlate with the amount of music training. Put together, these findings constitute substantial evidence that music does not enhance GCA or any other domain-general cognitive skill. Rather, intelligent people are more likely to engage and succeed in the field of music.

**Chess Instruction**

Like music, the experimental studies on the cognitive effects of playing chess does not bring compelling evidence. Our meta-analysis [58] has shown that chess-training interventions exert a small to medium effect on cognitive and academic skills. However, the overall design quality of the primary studies is poor. To the best of our knowledge, only four studies have compared chess with alternative activities, finding no significant treatment effects [59-62]. Given the lack of an active control group in most experimental studies, it is not possible to rule out the possibility that the observed effects are in fact placebo effects.

**Brain Training**

The term “brain-training programs” denotes computer games specifically designed to improve cognitive skills. In this context, we thus refer to brain-training programs as a subset of cognitive-training programs. Examples are the games developed by companies such as *CogMed*, *Lumosity*, and *BrainHQ*. Brain-training programs usually include a series of activities aimed to boost several core cognitive skills such as processing speed, WM, and
reasoning. Enhancing such cognitive mechanisms is supposed to enhance GCA and, hence, academic and professional performance.

Whether brain training works has been the topic of a heated debate, which is vividly reflected by two open letters about the cognitive benefits of commercial cognitive-training programs. The first one – issued by the Stanford Center on Longevity and the Max Planck Institute for Human Development – “object[ed] to the claim that brain games offer consumers a scientifically grounded avenue to reduce or reverse cognitive decline when there is no compelling scientific evidence to date that they do” [63]. The second one – published on the Cognitive Training Data website – argues that “there is, in fact, a large and growing body of such evidence” [64].

To the best of our knowledge, no comprehensive meta-analysis has been carried out on this topic so far. Meta-analyses including brain-training interventions usually contain studies about other cognitive-training programs as well [65]. Mixing studies makes it impossible to isolate the specific effects of brain-training programs. Also, most of these meta-analyses do not implement appropriate methods for correcting for publication bias. The reported overall effect sizes are often small to medium, are sometimes highly inconsistent (i.e., high degree of true heterogeneity; Box 2), and often do not distinguish between near- (i.e., domain-specific) and far-transfer (i.e., domain-general) effects.

A recent comprehensive systematic review [15] offers the best discussion of the state of the art in this field. The authors highlight that, whilst brain-training regimens positively impact on performance on the trained tasks and, to a lesser extent, on similar tasks, there is no evidence of generalized effects on unrelated tasks. Thus, brain-training programs do not enhance either GCA or any broad cognitive construct (Box 3).

Another key point raised by that review concerns the relation between the observed effects and the design quality of the experiment. Limitations such as the use of passive control groups without active control groups, small samples, and selectivity in reporting the results often contribute to artificially inflating the overall effect of the treatment. Conversely, the effects decrease with better-designed studies. This pattern of results – that
is, the better the design, the smaller the size of the effect – militates in favour of the null hypothesis according to which brain-training programs do not exert any appreciable effects on cognition.

**Other Cognitive-Training Programs**

Other cognitive-training regimens have been recently examined. Examples include cognitive-flexibility training [66], task-switching training [67, 68], spatial training [69, 70], learning relational reasoning [71], exergames [72-74], meditation/mindfulness [75, 76], and multimodal cognitive training [77]. In accordance with the findings presented above, none of these types of training appears to exert appreciable effects on overall cognitive function. It is worth mentioning that, given the small number of experimental studies carried out so far, the impact on these programs on cognition requires further investigation.

**Cognitive Training: The Broader Picture**

Meta-analytic methods provide precious information about the overall effect of interventions, between-study variability, moderating variables, and publication bias. They have proved to be excellent tools for resolving open controversies in social and life sciences [78]. However, our scepticism about the alleged cognitive benefits of cognitive-training programs also stems from other broader considerations.

To begin with, education has been found to have only a small impact on GCA [2, 79, 80]. In fact, education exerts larger effects on those tests assessing skills taught in school rather than IQ tests [81]. Moreover, the observed slight improvements in IQ associated with additional years of schooling seem to be limited to specific subtests [80, 82] or mediated by the acquisition of other skills such as reading and mathematics [81, 83, 84]. If years of cognitively demanding school activities only marginally enhance GCA or its core components (e.g., working memory capacity, and processing speed), it is hard to see why a few hours of cognitive training should. Second, GCA is a substantially heritable factor [85], which means that, unlike domain-specific skills, it is hardly malleable to environmental factors such as education and training. Finally, as seen earlier, research into the psychology of expertise has consistently shown that generalized (i.e., far) transfer is rare because skill
acquisition is based on domain-specific conceptual and perceptual information [7, 10, 86-88]. Thus, while prior levels of GCA influence the speed and quality of skill acquisition, domain-specific training does not foster GCA. In other words, the benefits of a particular training do not go beyond the trained tasks and, at best, similar tasks. Put together, the insights from these disciplines strongly suggest that the idea of enhancing GCA through training is scientifically implausible.

**Concluding Remarks and Future Directions**

Evidence from different laboratories supports the idea that practicing cognitively demanding activities has little, if any, impact on GCA or its components. The lack of generalized effects seems to manifest itself regardless of the type of training, the specific cognitive skills assessed, the age of participants, or the presence of training-induced neural changes (Box 4). In our opinion, the impossibility of training GCA should be regarded, to use a metaphor from Physics, as an “elementary particle” in the “standard model” of human cognition. Furthermore, the absence of generalized cognitive benefits should be considered a fundamental litmus test for theories in cognitive sciences. Theories whose predictions (or corollaries) are in line with this principle (e.g., theories of expert performance such as chunking and template theories [89, 90]) should be preferred over theories supporting the hypothesis that training can affect domain-general cognitive skills [91]. Similarly, theories that emphasize the difficulty of far transfer [9, 92, 93] should be preferred over theories that are more optimistic about far transfer [94, 95]. An important practical implication is that school instruction and professional training should concentrate on domain-specific material and avoid domain-general cognitive training.

The impossibility of enhancing GCA by training does not imply that human cognition is not malleable to training (see Outstanding Questions). Rather, it must be acknowledged that the benefits associated with training are limited to the trained tasks and, sometimes, similar tasks. Our conviction is that cognitive-training programs should not be utterly abandoned, as long as the claims and expectations about the benefits do not go beyond what has been empirically verified. For example, brain-training programs
implementing arithmetic games do improve the participants’ ability to perform simple calculations [96]. Also, given the similarities between the chess board and the Cartesian graph, chess instruction may help children to learn basic geometrical concepts. Such improvements may not be considered cognitive enhancement, but that does not mean that they are not useful. In any case, improving the quality of experimental designs will be an essential requirement for future studies in order to estimate the extent to which these practices promote transfer of skills from the laboratory to real-life contexts.

Furthermore, other avenues for improving performance in cognitively demanding tasks should be pursued. Learning strategies have been claimed to help people to learn more quickly and cope with cognitive decline. Examples include learning situations that are desirably difficult [97], retrieval-based learning [98], mnemonics [99], and elaboration strategies [100]. Once again, further research is needed to verify whether skills acquired via this type of training transfer to real-life situations [101].

One additional step forward is the design of cognitive-training programs aimed at improving domain-specific performance. As mentioned above, experts such as professional musicians and chess masters exhibit largely superior cognitive performance compared to novices when the task involves domain-specific material (e.g., recall and detection of chess pieces or music notes). Such skills are almost certainly a by-product of practice in these fields. However, we do not know whether practicing cognitive tasks with domain-specific material (e.g., recalling chess positions) contributes to boosting domain-specific cognitive performance (e.g., playing chess well). We believe that domain-specific cognitive-training methods represent a new opportunity towards the amelioration of educational and professional performance. Only time will tell us whether this hypothesis is empirically correct.
Glossary

**Cognitive skills**: in this context, this phrase refers to latent factors produced from cognitive tests that strongly correlate to each other. These factors constitute Stratum II of the Cattel-Horn-Carroll model (Box 1).

**Correlational**: a correlational study investigates the (linear) relationship between two or more variables (e.g., chess skill and intelligence) within a particular population (e.g., chess players).

**Cross-sectional**: a cross-sectional study compares two or more populations (e.g., chess players and non-chess players) on one or more variables (e.g., cognitive skills).

**Domain-general**: refers to mental abilities that are used to solve complex tasks regardless of their content.

**Domain-specific**: refers to mental abilities that are engaged only when material related to a particular field is involved.

**Experimental**: refers to true experiments, that is, studies implementing an intervention (e.g., music instruction) aimed to exert an effect on one or more variables (e.g., academic skills).

**Far transfer**: the generalization of acquired skills across domains loosely related to each other (e.g., studying mathematics to improve in Latin).

**General cognitive ability (GCA)**: also referred to as intelligence, Spearman’s $g$, and general mental ability, depending on the sources. This term refers to the latent factor emerging from all the tests of mental ability. It represents Stratum III in the Cattel-Horn-Carroll model (Box 1).

**Latent factor**: refers to a variable that is mathematically derived from observed variables. For example, scores in a large set of cognitive tests may be strongly correlated to each other and thus reflect the same underlying (i.e., latent) cognitive construct (e.g., GCA). The most common statistical technique to estimate latent factors is *factor analysis*. 
Neural plasticity: refers to the capability of the neural system to adapt to environmental pressures.
Box 1. The Cattel-Horn-Carroll Model

A well-known finding in psychology is that all tests of mental ability positively correlate with each other. This fact has led researchers to propose that all these tests measure a common factor: GCA. In psychometrics, GCA represents the portion of the between-individual variance that is common across all tests of cognitive ability [102]. Thus, GCA is a theoretical construct expressed in the form of a **latent factor**.

Correlations between cognitive tests are particularly strong when they refer to the same construct (e.g., processing speed). As a consequence, researchers have developed a hierarchical model – the Cattel-Horn-Carroll (CHC) model – that includes an intermediate level between GCA and tests of cognitive ability. This level consists of a series of latent factors referring to broad cognitive constructs (cognitive skills in the main text). The CHC model is thus designed on three levels (or strata). Stratum I refers to observed variables performance on cognitive tests. Stratum II consists of broad latent factors (i.e., cognitive skills) derived from cognitive tests (Stratum I) that highly correlate with each other. Finally, Stratum III (GCA) is the latent factor that subsumes all the common variance across latent factors in Stratum II [103].

Examples of cognitive skills in Stratum II are fluid reasoning (Gf), short-term memory (Gsm), processing speed (Gs), comprehension knowledge (Gc), and visual processing (Gv). Gf is the capability of solving new problems and adapting to novel situations. Gsm is the ability to retain and recall information over a short period of time. Gs refers to the ability to perform elementary cognitive tasks when high intellectual efficiency (i.e., speed and accuracy) is needed. Gc reflects language, knowledge and skills acquired through experience. Finally, Gv denotes the ability to generate and manipulate visual images.

The CHC model does not make any specific predictions about the possibility of enhancing GCA through training. Nevertheless, the CHC model can provide a useful theoretical framework to understand the claims about the presumed benefits of cognitive training. According to cognitive-training theories, engaging in Stratum I activities boosts one or more cognitive skills (Stratum II) or even GCA (Stratum III). Fostered cognitive skills
increase GCA (Stratum III) which, in turn, exerts a positive influence on all the measures of Strata II and I. As seen in the main text, this hypothesis has never received robust empirical support, despite its appeal [e.g., 77].
Box 2. Meta-Analytic Modelling

Meta-analysis is a set of statistical techniques for integrating research findings across studies on a particular topic [104, 105]. Its primary objective is to evaluate (a) whether an effect (e.g., the relationship between two variables) is statistically significant and (b) how big and consistent the effect is. To date, meta-analysis is arguably the most effective tool for resolving disputes in quantitative empirical research.

In its simplest form, meta-analysis is just a sample-size-weighted mean. The so-called overall effect size is calculated by averaging all the effect sizes (e.g., standardized mean differences between groups) extracted from the primary studies. Effect sizes are weighted on precision (i.e., inverse of the variance), which is primarily (sometimes solely) a function of sample size. This way, the bigger the sample, the bigger the weight of the effect in the analysis.

Another fundamental information offered by meta-analysis is the degree of true heterogeneity ($I^2$). $I^2$ is a measure of the between-study variability in the population of the effect sizes that is not due to random error. While a low $I^2$ is good evidence of the consistency of the effect across the primary studies, a high $I^2$ suggests that the effect is moderated by some variables (e.g., type of control group). Accounting for true heterogeneity – when it exists – is essential to provide reliable and interpretable results.

Finally, meta-analysis can estimate the amount of publication bias, that is, the inflation of the overall effect size due to the systematic suppression of statistically non-significant ($p > .05$) results from a particular literature. This problem has been unanimously recognized as one of the worst threats to credibility in scientific research [106]. Thus, estimating an overall effect corrected for publication bias is crucial.

Techniques for detecting publication bias are often based on the degree of asymmetry of the distribution of the effect sizes around the meta-analytic mean [107], and the relationship between the variance and the size of the effect [108]. Ideally, unbiased overall effects are associated with symmetrically distributed effect sizes that are not directly related to their
variances. Other techniques focus on the distribution of statistically significant ($p < .05$) effect sizes [109] or test the robustness of the results by postulating a probability for the non-significant effects to be published [110, 111]. Technical details apart, it is worth noting that these methods not only correct for the suppression of smaller-than-average effect sizes, but, to some extent, also for statistical artefacts due to questionable research practices (e.g., $p$-hacking).
Box 3. Task Performance and Latent Factors

Standardized cognitive tests are used to assess one’s cognitive skills. The performance on a cognitive test is considered a reliable proxy for more complex cognitive constructs such as memory or fluid intelligence. For example, we can assume that our performance in the $n$-back task assesses our WM capacity.

This equivalence is harder to establish in studies implementing an intervention than in cross-sectional and correlational studies. In fact, it is not simple to understand whether pre-post-test differences represent cognitive enhancement or just the improvement in carrying out a cognitive task. Better performance may be due to the similarities between the trained task and the outcome measures rather than improved cognitive function.

This is the main objection formulated against the alleged benefits of WM training on general memory skills [112]. WM training has been found to exert a moderate yet reliable positive effect on several memory tests. However, extended training in a particular memory task may help trainees to develop strategies to solve similar tasks. Similarly, playing *Tetris* is likely to boost the ability to carry out some mental-rotation tasks rather than general spatial ability. This problem is even more evident with brain-training interventions, which often consist of cognitive tasks that have been turned into computer games. It is thus obvious that these activities help to increase one’s performance on the original cognitive tasks. However, this does not represent any reliable evidence of cognitive enhancement.

To address this issue, experiments must include multivariate measures of cognitive skills that are not too similar to the trained task(s). In fact, such a set of cognitive measures, rather than performance in a single cognitive task, is needed to investigate the impact of cognitive-training programs on latent factors representing cognitive skills. Designs based on latent factors have rarely been implemented in cognitive-training studies. When they have, no evidence of training-related benefits for GCA or particular cognitive skills has been documented [e.g., 23, 77].
Box 4. Neural Patterns and Cognitive Skills

As seen, neural plasticity has been proposed to be one of the links between cognitive-training programs and cognitive enhancement. This idea is supported by the fact that performance in tests of intelligence is associated with particular functional (e.g., brain activation during a complex task) and structural (e.g., volume of grey matter) neural patterns [113]. Experts often exhibit structural and functional neural patterns specific to their field. This is the case of musicians, chess players, and mnemonists, just to mention a few examples [114, 115]. However, correlational and cross-sectional evidence does not tell us anything about the direction of causality. It is not possible to establish whether these neural patterns are due to training or existed before. So, what do experimental studies show us?

Interestingly, training-induced functional and sometimes structural changes have been observed after music-based interventions [116, 117], WM training [118, 119], and video-game training [120]. Regarding chess, particular structural and functional neural patterns have been observed in experts [121-123]. (To the best of our knowledge, no experimental study has examined the effect of chess training on structural/functional patterns.)

With the exception of working-memory-related tasks, where experts use long-term memory brain regions typically not used by novices [124], the occurrence of particular neural correlates in both experts and participants undergoing cognitive-training regimens is quite consistent. Yet, the impact of cognitive-training programs on GCA and cognitive skills is, as seen, substantially null. These results suggest a decoupling of observed neural changes from domain-general cognitive skills. It is likely that the neural changes observed after cognitive-training interventions reflect the improved ability to perform the trained tasks and, sometimes, similar tasks.

A possibility is that training-induced domain-specific neural patterns underlie those mechanisms necessary to store, retrieve, and manipulate domain-specific information to carry out complex cognitive tasks. For example, extended training in chess and music may result in localized neural changes [e.g., 114, 115] that sustain domain-specific tasks such as
retrieving chess and music configurations of pieces and notes from long-term memory and use this information to find the correct move or play a Mozart Sonata. By contrast, GCA and domain-general cognitive skills are probably expressed by more holistic neural processes, such as the dynamic reorganization of brain networks [102].
Acknowledgements

We thank Walter Boot, Zach Hambrick, and Dan Simons for providing insightful information about cognitive-training programs and the structure of intelligence.

References

2. Detterman, D.K. (2016) Education and Intelligence: Pity the Poor Teacher because Student Characteristics are more Significant than Teachers or Schools. The Spanish Journal of Psychology 19, E93.
16. Taatgen, N.A. (2016) Theoretical models of training and transfer effects. In Cognitive...


Aging Clinical and Experimental Research 29, 387.
82. Ritchie, S.J. et al. (2013) Education is associated with higher later life IQ scores, but not with faster cognitive processing speed. Psychology and Aging 28, 515-21.


Outstanding questions Box

Is generalized transfer a function of intelligence? Are more intelligent people more able to see the abstract structure of a task and thus be able to use it to solve other tasks? If this is the case, the putative cases of transfer from one domain to another are more the reflection of pre-existing cognitive differences than the effect of practicing a given task on cognitive skills.

Has the field of cognitive training be naïve in thinking that cognitive enhancement can be obtained from practice in one domain only? Possibly, training in several domains is necessary for extracting common abstract structures. If this is the case, increasing intelligence requires acquiring a minimum level of expertise in several domains, which is obviously more time consuming and more challenging motivationally than practicing only one domain.

Does cognitive training have positive effects with specific populations such as patients with brain-damage, Alzheimer’s and other dementias? The idea is that, although it exerts no positive effects with healthy populations, cognitive training may benefit some particular populations whose cognitive function has been impaired.

Is it possible to improve cognition with drugs? General cognitive ability is associated with specific genotypes. It has been proposed that drug treatments may be employed to modify gene expression to promote cognitive enhancement.

Cognitive training does not enhance overall cognitive ability or core cognitive mechanisms. However, is it possible that engaging in cognitively demanding activities slows down cognitive decline in the elderly? This claim refers to the so-called “use it or lose it” hypothesis.
Trends Box

- General cognitive ability (GCA) has been consistently found to correlate with performance in cognitive tasks and complex activities such as playing music, board games, and video games.
- In the last two decades, researchers have thus extensively investigated the effects of engaging in cognitive-training programs and intellectually demanding activities on GCA. The results have been mixed.
- Several independent researchers have noticed that the between-study variability can be accounted for by the quality of the experimental design and statistical artifacts. Those studies including large samples and active control groups often report no training-related effects.
- These findings show that practicing cognitive-training programs or intellectually demanding activities do not enhance GCA or any cognitive skill. At best, such interventions boost one’s performance in tasks similar to the trained task.