

Once Upon a Time There Was Far Transfer

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by *Giovanni Sala*.

August 2017

Contents

| | |
|-----------------------------------------------------------------------------------------|----|
| List of Figures | 9 |
| List of Tables | 11 |
| Abstract | 13 |
| Acknowledgements | 15 |
| Rationale for Submitting the Thesis in an Alternative Format | 17 |
| Chapter 1: Introduction | 19 |
| 1. Preview of The Dissertation | 21 |
| Chapter 2: Transfer, Expertise, and Cognitive Training | 25 |
| 1. The Curse of Specificity: The Difficulty of Far Transfer | 25 |
| 2. Training Domain-General Cognitive Abilities: A Way to Get Around Far Transfer? | 27 |
| 2.1 Mixed Effects and The Problem of Design Quality | 28 |
| 3. Different Types of Cognitive Training | 30 |
| 3.1 Practicing Cognitive Tasks | 30 |
| 3.2 Engaging in Cognitively Demanding Activities | 35 |
| Chapter 3: Meta-Analytic Techniques | 43 |
| 1. Effect Sizes | 43 |
| 2. Fixed and Random Effect Models | 45 |
| 2.1 Assessing Heterogeneity | 45 |
| 2.2 Moderator Analysis | 46 |
| 3. Publication Bias | 46 |
| 3.1 Trim-and-Fill | 47 |
| 3.2 PET-PEESE | 47 |
| 3.3 Begg and Mazumdar's (1994) Rank Correlation Test | 47 |
| 3.4 P-Curve | 48 |
| 3.5 Egger's Regression Test | 48 |
| 3.6 Selection Models | 48 |
| 4. Statistical Dependence of Effect Sizes | 49 |
| 5. Techniques for Detecting Outliers | 49 |
| 5.1 Winsorizing | 49 |
| 5.2 Influential Case Analysis | 50 |
| Chapter 4: Meta-Analysis of Working Memory Training | 51 |
| Rationale for the Meta-Analysis in Chapter 4 | 51 |
| 1. Introduction | 52 |
| 1.1 Working Memory Training | 52 |
| 1.2 Working Memory Training in Children | 54 |

| | |
|-------------------------------------------------------------------|-----|
| 1.3 The Present Meta-Analysis | 56 |
| 2. Method | 57 |
| 2.1 Literature Search | 57 |
| 2.2 Inclusion/Exclusion Criteria | 57 |
| 2.3 Moderators | 60 |
| 2.4 Effect Size | 80 |
| 2.5 Statistical Dependence of the Samples | 80 |
| 3. Results | 81 |
| 3.1 Near-Transfer Effects | 81 |
| 3.2 Far-Transfer Effects | 84 |
| 3.3 Follow-Up Effects | 89 |
| 4. Discussion | 90 |
| 4.1 Theoretical and Practical Implications | 91 |
| 4.2 Limitations of the Present Meta-Analysis | 93 |
| 4.3 Conclusions | 94 |
| Chapter 5: Meta-Analysis of Chess Training | 95 |
| Rationale for the Meta-Analysis in Chapter 5 | 95 |
| 1. Introduction | 96 |
| 1.1 Difficulty of Transfer | 97 |
| 1.2 The issue of Transfer in Chess Research | 98 |
| 1.3 Chess in School | 99 |
| 2. Scope, Aims, and Hypotheses of the present Meta-Analysis | 100 |
| 3. Method | 102 |
| 3.1 Literature Search | 102 |
| 3.2 Inclusion/Exclusion Criteria | 103 |
| 3.3 Effect Size | 109 |
| 4. Results | 110 |
| 4.1 Moderator Analyses | 111 |
| 4.2 Additional Meta-Analytic Models | 112 |
| 5. Discussion | 112 |
| 5.1 Substantive Results | 113 |
| 5.2 Methodological Moderators | 115 |
| 5.3 Limitations of this Study | 115 |
| 5.4 Conclusions and Recommendations for Future Research | 116 |
| Chapter 6: Meta-Analysis of Music Training | 119 |
| Rationale for the Meta-Analysis in Chapter 6 | 119 |

| | |
|------------------------------------------------------------------------------------------|-----|
| 1. Introduction..... | 120 |
| 1.1 The Question of Transfer of Skills..... | 120 |
| 1.2 Why Should Music Skill Transfer to non-Music Skills? | 121 |
| 1.3 Does Music Skill Transfer to non-Music Skills? A Look at the Empirical Evidence..... | 122 |
| 1.4 Aims of the Present Meta-Analysis | 124 |
| 2. Method | 126 |
| 2.1 Literature Search | 126 |
| 2.2 Inclusion/Exclusion Criteria | 126 |
| 2.3 Moderators | 127 |
| 2.4 Effect Size..... | 141 |
| 2.5 Statistical Dependence of the Samples | 142 |
| 3. Results..... | 142 |
| 3.1 Meta-Regression Analysis | 143 |
| 3.2 Additional Meta-Analytic Models | 143 |
| 3.3 Publication Bias Analysis | 144 |
| 3.4 Sensitivity Analysis..... | 145 |
| 4. Discussion..... | 146 |
| 4.1 Substantive Results | 147 |
| 4.2 Methodological Results..... | 148 |
| 4.3 Conclusions and Recommendations for Future Research..... | 149 |
| Chapter 7: Meta-Analysis of Video-Game Training..... | 151 |
| Rationale for the Meta-Analysis in Chapter 7 | 151 |
| 1. Introduction..... | 152 |
| 2. The Meta-Analytic Evidence | 154 |
| 3. The Present Meta-Analytic Investigation | 156 |
| 4. General Method | 156 |
| 4.1 Literature Search | 156 |
| 4.2 Inclusion Criteria | 157 |
| 4.3 Outcome Measures..... | 158 |
| 4.4 Statistical Dependence of the Samples | 159 |
| 4.5 Calculations of the Overall Meta-Analytic Means..... | 160 |
| 4.6 Publication Bias Analysis | 160 |
| 4.7 Influential Cases Analysis..... | 161 |
| 5. Meta-Analysis 1: Meta-Analysis of Correlational Data Among Video Game Players..... | 161 |
| 5.1 Method..... | 162 |
| 5.2 Results..... | 163 |

| | |
|-----------------------------------------------------------------------|-----|
| 5.3 Discussion | 177 |
| 6. Meta-Analysis 2: Meta-Analysis of Quasi-Experimental Data..... | 178 |
| 6.1 Method | 178 |
| 6.2 Results..... | 180 |
| 6.3 Discussion..... | 189 |
| 7. Meta-Analysis 3: Meta-Analysis of Experimental Training Data | 189 |
| 7.1 Method | 189 |
| 7.2 Results..... | 191 |
| 7.3 Discussion..... | 206 |
| 8. General Discussion | 207 |
| 8.1 Theoretical and Practical Implications..... | 208 |
| 8.2 Recommendations for Future Research | 210 |
| 8.3 Conclusion | 211 |
| Chapter 8: Meta-Analysis of Exergame Training..... | 213 |
| Rationale for the Meta-Analysis in Chapter 8 | 213 |
| 1. Introduction..... | 214 |
| 2. Method | 217 |
| 2.1 Included Studies..... | 217 |
| 2.2 Effect Sizes | 217 |
| 2.3 Meta-Analytic Models | 218 |
| 2.4 Publication Bias Analyses..... | 219 |
| 3. Results..... | 219 |
| 3.1 Model 1 (K = 16) | 219 |
| 3.2 Model 2 (K = 75) | 220 |
| 3.3 Sensitivity Analysis..... | 221 |
| 4. Discussion..... | 222 |
| 4.1 Inflation of the Effect Sizes | 223 |
| 4.2 Heterogeneity | 223 |
| 4.3 Publication Bias | 224 |
| 4.4 Conclusions and Recommendations for Future Research..... | 224 |
| Chapter 9: Overall Discussion | 227 |
| 1. Summary of the Results | 227 |
| 2. Theoretical and Practical Implications..... | 230 |
| 2.1 Main Implications | 230 |
| 2.2 Training and Neural Plasticity | 231 |
| 2.3 Education and Skill Acquisition | 232 |

| | |
|----------------------------------------------------------------------|-----|
| 3. Methodological Considerations | 233 |
| 4. Generalizability of the Results..... | 235 |
| 5. Conclusions and Future Research Directions | 236 |
| References..... | 239 |
| Appendix A. Studies Included in the Meta-Analysis in Chapter 4..... | 271 |
| Appendix B. Studies Included in the Meta-Analysis in Chapter 5 | 275 |
| Appendix C. Studies Included in the Meta-Analysis in Chapter 6 | 279 |
| Appendix D. Studies Included in Chapter 7 (Meta-Analysis 1) | 285 |
| Appendix E. Studies Included in Chapter 7 (Meta-Analysis 2)..... | 293 |
| Appendix F. Studies Included in Chapter 7 (Meta-Analysis 3)..... | 305 |
| Appendix G. Studies Included in the Meta-Analysis in Chapter 8 | 313 |

List of Figures

Figure 1. Flow diagram of the studies included in the meta-analytic review.

Figure 2. Forest plot of the near-transfer model. Hedges's g s (circles) and 95% CIs (lines) are shown for all the effects entered into the meta-analysis. The diamond at the bottom indicates the meta-analytically weighted mean \bar{g} . When studies had multiple samples, the table reports the result of each sample (S1, S2, etc.) separately. Similarly, when studies used multiple outcome measures, the table reports the result of each measure (M1, M2, etc.) separately.

Figure 3. Contour-enhanced funnel plot of standard errors and effect sizes (Hedges's g s) in the near-transfer meta-analysis. The black circles represent the effect sizes included in the meta-analysis. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

Figure 4. p -curve analysis. The blue (continuous) line shows that most of the significant p -values are smaller than .025, suggesting evidential value.

Figure 5. Forest plot of the far-transfer model. Hedges's g s (circles) and 95% CIs (lines) are shown for all the effects entered into the meta-analysis. The diamond at the bottom indicates the meta-analytically weighted mean \bar{g} . When studies had multiple samples, the table reports the result of each sample (S1, S2, etc.) separately. Similarly, when studies used multiple outcome measures, the table reports the result of each measure (M1, M2, etc.) separately.

Figure 6. Contour-enhanced funnel plot of standard errors and effect sizes (g s) in the far-transfer meta-analysis. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

Figure 7. p -curve analysis. The blue (continuous) line shows that most of the significant p -values are smaller than .025, suggesting evidential value.

Figure 8. Flow diagram of the studies considered and ultimately included in the meta-analysis.

Figure 9. Funnel plot of standard errors and effect sizes (g). The diamond at bottom represents the meta-analytically weighted mean Hedges's \bar{g} .

Figure 10. p -curve analysis of the studies reporting significantly positive results. The blue (continuous) line shows that most of the significant p -values are smaller than .01.

Figure 11. Flow diagram of the studies included in the meta-analyses.

Figure 12. Contour-enhanced funnel plot of standard errors and effect sizes (Fisher's Z s) in the meta-analysis of the correlational data. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

Figure 13. Contour-enhanced funnel plot of standard errors and effect sizes (g s) in the meta-analysis of the quasi-experimental data. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

Figure 14. Contour-enhanced funnel plot of standard errors and effect sizes (g s) in the meta-analysis of the experimental data. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

Figure 15. Contour-enhanced funnel plot of standard errors and effect sizes (Hedges's g s) in Model 1. The black circles represent the effect sizes included in the model. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

Figure 16. Contour-enhanced funnel plot of standard errors and effect sizes (Hedges's g s) in Model 2. The black circles represent the effect sizes included in the model. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

List of Tables

- Table 1. Studies and moderators of the 30 near-transfer effect sizes included in the meta-analysis
- Table 2. Studies and moderators of the 74 far-transfer effect sizes included in the meta-analysis
- Table 3. Studies and moderators of the 6 near-transfer follow-up effect sizes included in the meta-analysis
- Table 4. Studies and moderators of the 24 near-transfer follow-up effect sizes included in the meta-analysis
- Table 5. Summary of the 24 studies included in the meta-analysis
- Table 6. Studies, dependent variables, and moderators of the 118 effect sizes included in the meta-analysis
- Table 7. Overall effect sizes, confidence intervals, ks, and p-values in each outcome measure
- Table 8. Overall effect sizes, confidence intervals, ks, and p-values in each outcome measure
- Table 9. Meta-analytic and publication bias results of the main model sorted by outcome measure (Meta-analysis 1)
- Table 10. Meta-analytic and publication bias results sorted by outcome measure, with video game skill measured by frequency of video game playing (Meta-analysis 1)
- Table 11. Meta-analytic and publication bias results sorted by outcome measure, with video game skill measured by video game scores (Meta-analysis 1)
- Table 12. Meta-analytic and publication bias results sorted by outcome measure, for the correlations referring to action video games (Meta-analysis 1)
- Table 13. Meta-analytic and publication bias results sorted by outcome measure, for the correlations referring to non-action video games (Meta-analysis 1)
- Table 14. Meta-analytic and publication bias results sorted by outcome measure, for the

correlations referring to mixed video games (Meta-analysis 1)

Table 15. Meta-analytic and publication bias results of the main model sorted by outcome measure (Meta-analysis 2)

Table 16. Meta-analytic and publication bias results sorted by outcome measure, for the effect sizes referring to action video games (Meta-analysis 2)

Table 17. Meta-analytic and publication bias results sorted by outcome measure, for the effect sizes referring to mixed video games (Meta-analysis 2)

Table 18. Meta-analytic and publication bias results of the main model sorted by outcome measure (Meta-analysis 3)

Table 19. Meta-analytic and publication bias results sorted by outcome measure, for the effect sizes referring to action video game players vs. non-action video game players (Meta-analysis 3)

Table 20. Meta-analytic and publication bias results sorted by outcome measure, for the effect sizes referring to action video game players vs. non-video game players (Meta-analysis 3)

Table 21. Meta-analytic and publication bias results sorted by outcome measure, for the effect sizes referring to non-action video game players vs. non-video game players (Meta-analysis 3)

Table 22. Meta-analytic and publication bias results sorted by outcome measure, for the effect sizes referring to adult video game players (Meta-analysis 3)

Table 23. Meta-analytic and publication bias results of the effect sizes referring to the old video game players sorted by outcome measure (Meta-analysis 3)

Abstract

The last two decades have seen the rise of cognitive-training research. Strong claims have been made. Roaring refutations have been published. Then again counter-evidence supporting the effectiveness of cognitive training has been produced. Definite conclusions are far from being drawn.

Undoubtedly, due to the potential theoretical and practical implications, the idea of enhancing cognitive function and, hence, a broad range of other real-life skills by training is appealing. However, this idea is at variance with substantial research into the psychology of expertise showing that performance in specific tasks relies massively on perceptual information. In fact, such information is hardly transferable across different domains.

To solve these discrepancies, I ran a series of meta-analytic models to examine the effects of several types of cognitive training (i.e., chess, music, working memory, video-game, and exergame training) on cognitive and academic skills in different types of populations. None of the five types of cognitive training exerted any meaningful effect on any non-trained skill.

While confirming the previous findings of the research on expertise, these results convincingly reject the cognitive-training hypothesis. The lack of generalization across different domains of skills acquired by training appears to be a constant in human cognition. The program of research of cognitive training has failed. Transfer of skills across loosely related domains remains a chimera.

Keywords: cognitive training; expertise; meta-analysis; transfer.

Acknowledgements

I am grateful to all the colleagues, friends, and relatives who have supported me in the last three years. They too deserve some credit for my research.

I also thank my supervisor, Professor Fernand Gobet (IM). He is a tough guy.

Rationale for Submitting the Thesis in an Alternative Format

This dissertation has been submitted in the alternative paper format. This format consists of research chapters that are formatted to be suitable for submission to peer-reviewed scientific journals. The policy of the University of Liverpool on submitting theses in this format requires to: (a) call each paper a chapter; (b) include an introductory section for each paper explaining how the study links to the previous and following ones; (c) to re-format the papers according to the general guidelines (e.g., all the references listed together); and (d) the thesis must include a general introduction and a general conclusion/discussion integrating and discussing the results presented in the papers. All these requirements have been satisfied.

Five papers have been included in this dissertation. The first study (Chapter 4) has been published in *Developmental Psychology*. The second study (Chapter 5) has been published in *Educational Research Review*. The third study (Chapter 6) has also been published in *Educational Research Review*. The fourth study (Chapter 7) has been submitted and is currently under review at *Psychological Bulletin*. The fifth study (Chapter 8) has been submitted and is currently under review at *Neuroscience and Biobehavioral Reviews*.

I am the first author of all the papers. I have been responsible for the design of the studies, the extraction and coding of the data, running the analyses, and writing and revising the papers. All the papers are co-authored with my supervisor for my Ph.D. program, Prof. Fernand Gobet. He has supported me in all the phases of research.

I have submitted the dissertation in this format to offer the reader, when possible, peer-reviewed material. I think that including such material is – beyond my professionalism and effort – a further guarantee of quality for this dissertation.

Chapter 1: Introduction

Transfer of learning is something all of us experience in our daily life. Knowledge of Samsung smartphones transfers to iPhones. Driving one's car generalizes to other models of cars. Knowing how to cook spaghetti Bolognese is useful for cooking chicken pasta. All these are examples of near transfer, that is, the generalization of a set of skills across two (or more) domains tightly related to each other. However, another type of transfer has attracted the attention of researchers for over a century: *far transfer*. Far transfer occurs when a set of skills generalizes across two (or more) domains that are only loosely related to each other (e.g., mathematics and Latin).

In a seminal article, Thorndike and Woodworth (1901) proposed their common elements theory according to which transfer is a function of the extent to which two domains share common features. The theory predicts that, while near transfer takes place often, far transfer is much less common. This point has been echoed by extensive research into the psychology of expertise and skill acquisition. For example, research on chess players has established that expert performance relies, to a large extent, on perceptual information such as the knowledge of tens of thousands of chunks (i.e., meaningful configurations of chess pieces; Chase & Simon, 1973; Sala & Gobet, 2017a). Due to its high specificity, such information is hardly transferable to other fields, as predicted by chunking theory (Chase & Simon, 1973) and template theory (i.e., an extension of chunking theory; Gobet, 2016; Gobet & Simon, 1996). However, research on expertise has also provided convincing evidence that experts – such as chess masters and professional musicians – possess, on average, superior overall cognitive ability. Importantly, domain-general cognitive abilities (e.g., intelligence, processing speed, and working memory) are reliable predictors of success for outcomes such as academic achievement (Deary, Strand, Smith, & Fernandes, 2007) and job proficiency (Hunter & Hunter, 1984).

At this point, I can see the reader waving their hands: this evidence establishes correlation, but can we conclude that there is a causal relationship? Does training in cognitively demanding activities *make* people smarter? Is it possible to train domain-general cognitive abilities in one domain and hence obtain benefits in a vast number of areas? In other words; does far transfer occur?

The answers to these questions have profound theoretical and practical implications. In fact, establishing whether and under what conditions far transfer occurs would represent a major contribution to our comprehension of how humans acquire and use knowledge. Also, understanding whether and to what extent cognitive ability is malleable to training would have huge societal implications. Consider the academic advantages of fostering cognitive ability in youth or the benefits of slowing down cognitive decline in adulthood for the global economy and public health. Increasing human cognition is thus one of the most influential and potentially impacting scientific enterprises in cognitive science.

Due to the above potential implications, hundreds (if not thousands) of studies have investigated the possible far-transfer effects of several types of cognitive training in the last two decades. Examples include working memory training, executive function training, spatial training, chess instruction, music training, video-game training, exergaming, and brain training. The research into the effects of cognitive training has provided mixed results, and no agreement among researchers has been reached. A perspicuous example of this divergence of opinions is provided by two open letters about the benefits of commercial brain-training programs. The first letter, issued by the Stanford Center on Longevity and the Max Planck Institute for Human Development, has expressed serious doubts about the ability of brain games to enhance overall cognitive ability (“A Consensus on the Brain Training Industry from the Scientific Community,” 2014). The impact of such games seems to be task-specific, and the effects transfer to similar tasks at best. In other words, people certainly improve their

performance in the games they practice (and in similar games), but these benefits may not transfer to real-life tasks. The second one, posted on the Cognitive Training Data website (www.cognitivetrainingdata.org) and signed by a group of 133 researchers, has claimed that specific cognitive-training regimens can benefit overall cognitive function.

1. Preview of The Dissertation

Due to the substantial disagreement between studies and researchers, this dissertation is aimed at solving the discrepancies in the field by performing several statistical reviews of the literature. I will run meta-analyses on some of the key domains in the field of cognitive training, namely working memory training, chess instruction, music training, video-game training, and exergames. To achieve this goal, I will use a broad range of meta-analytic techniques. Meta-analysis comprises a set of statistical procedures for merging, correcting, and modelling the results from all the studies concerning a specific topic. Meta-analysis can thus estimate the actual size of the effect of a treatment far more reliably than the single experiment. Moreover, meta-analysis allows one to calculate the degree of between-study variability and test whether such variability is explained by some moderating variables, missing studies, and outliers. Put simply, meta-analysis provides the necessary statistical tools to account for the contradicting findings in the field of cognitive training.

Chapter 2 presents the theories of cognitive training and the accounts of the research into the psychology of expertise about transferability of skills. Expert performance appears to rely, to a large extent, on perceptual information (e.g., chunks). Crucially, such information is believed to be hardly transferable across different domains. For instance, there is no evident reason why memorizing the configuration of pieces in a particular chess opening helps one to learn how to play music or solve a math problem. According to theories of expertise (e.g., template theory), the benefits of training do not go beyond the trained tasks. In other words, substantial research into the psychology of expertise suggests that far transfer does not occur.

To get around the problem of domain-specificity of the training, several researchers have proposed that training domain-general cognitive skills is the most direct way to improve in a broad range of skills (Strobach & Karbach, 2016; Taatgen, 2016). The idea is simple. Cognitive skills such as working memory, focused attention, and fluid intelligence are necessary to carry out a wide set of different tasks in many domains. Enhancing these skills would necessarily lead to improving individuals' academic and professional general performance. This general hypothesis and its particular variants will be discussed with regard to the most common cognitive-training programs.

Chapter 3 introduces the basic concepts of meta-analysis (e.g., effect sizes, publication bias, and detection of outliers) and all the meta-analytic techniques used in the following meta-analyses. Understanding the rationale behind these techniques is essential. In fact, the meta-analyses included in the present dissertation often provide significantly different results compared to previous meta-analytic investigations in the field of cognitive training. These discrepancies are mainly due to the use of more precise methods for the calculation of the effect sizes and more advanced sensitivity analyses.

Chapters 4 to 8 present the results of a series of meta-analyses evaluating the effects of five types of cognitive-training programs on cognitive ability. All the relevant experimental (i.e., treatment) studies will be inserted into a series of meta-analytic models. A systematic search strategy and a set of statistical analyses will be adopted to estimate the effect sizes and test the robustness of the results (for details, see Chapter 3).

Chapter 4 reports the findings of a meta-analysis about the effects of working memory training on academic achievement, cognitive skills, and performance on working memory tasks in typically developing children (Sala & Gobet, 2017b). Chapter 5 is a meta-analysis regarding the impact of chess instruction on academic disciplines – such as mathematics and

literacy – and several cognitive abilities (e.g., focused attention) in children (Sala & Gobet, 2016). Chapter 6 studies the effects of music training on children and young adolescents' cognitive and academic skills (Sala & Gobet, 2017c). Chapter 7 presents a broad meta-analytical investigation about the impact of the practice of action and non-action video game on children, adults, and older adults' cognitive skills (Sala, Tatlidil, & Gobet, submitted-a). In Chapter 8, the effects of exergames – i.e., video games requiring both cognitive and physical engagement – on participants' cognitive skills will be assessed by a meta-analysis including all the relevant randomized control trials (Sala, Tatlidil, & Gobet, submitted-b).

Finally, Chapter 9 discusses the theoretical and practical implications of the findings. As previously mentioned, establishing whether domain-general cognitive skills can be enhanced and, hence, transferred to a broad range of domain-specific skills would have profound consequences. A positive result would pave the way for a plethora of practical applications in fields such as education, the professions, and cognitive rehabilitation. On the other hand, a negative result would provide further corroboration for classical theories of expertise and skill acquisition. Most importantly, a negative result would suggest that the lack of generalization of skills acquired by training is a constant in human cognition.

Chapter 2: Transfer, Expertise, and Cognitive Training

The question of the alleged benefits of cognitive training is strictly linked to the issue of transfer of learning. Transfer of learning occurs when a set of skills acquired in one domain generalizes to other domains (e.g., Barnett & Ceci, 2002). It is customary to distinguish between *near transfer* – i.e., the transfer taking place across two domains tightly related to each other – and *far transfer*, where the source domain and the target domain are only loosely related to each other. In a seminal article, Thorndike and Woodworth (1901) proposed that transfer of learning is a function of the extent to which two domains share common features. Thorndike and Woodworth's (1901) *common elements theory* thus predicts that, while near transfer is fairly common, far transfer is infrequent at best. As a direct consequence, the effects of cognitive training are expected to be limited to the trained task and other similar tasks.

1. The Curse of Specificity: The Difficulty of Far Transfer

Thorndike and Woodworth's (1901) common elements theory has received robust corroboration from research on the psychology of expertise. For example, the research on expert chess players has shown that expert performance relies, to a large extent, on domain-specific perceptual information – such as chunks, that is, perceptual and meaningful configurations of elements – acquired in years of training, as proposed by the *chunking theory* and *template theory* (Chase & Simon, 1973; Gobet & Simon, 1996; Sala & Gobet, 2017a).

As proposed by Chase and Simon (1973), expertise in chess is acquired by learning, through practice and study, a large number of chunks, which are units of both perception and meaning; in chess, chunks consist of constellations of pieces occurring often together in masters' games. Experts' superiority with meaningful material (game positions in chess) is

explained by their ability to rapidly identify patterns present on the board, and retrieve chunks from their long-term memory (LTM).

Template theory (Gobet & Simon, 1996) is an extension of chunking theory. According to the theory, chunks that are frequently used in a specific domain can evolve into more complex data structures called *templates*. Templates comprise two parts. The *core* consists of stable information and is comparable to a chunk. The *slots* consist of variable information, and their role is to encode information that occurs regularly but with some difference. For example, let a castle-like configuration in a chess position be the template. The core of such a template would be the position of the King and the Rook (stable information), while the slots would encode the positions of the *f*, *g*, and *h* Pawns (variable information). Crucially, due to its computational formulation, the template theory has been implemented in a cognitive architecture (CHREST; Gobet, 2016), and its predictions have been tested in both computer simulations and human participants (Gobet & Simon, 2000; Gobet & Waters, 2003).

Beyond chess, the chunking mechanism and, hence, perceptual information have been found to play an essential role in the acquisition of expertise in a wide range of fields, such as music (Knecht, 2003; Sloboda, 1976), programming (Adelson, 1981; Guerin & Matthews, 1990), and sports (Allard, Graham, & Paarsalu, 1980; Allard & Starkes, 1980; Abernethy, Neal, & Konig, 1994; Williams, Davids, Burwitz, & Williams, 1993). As predicted by chunking theory (Chase & Simon, 1973) and template theory (Gobet & Simon, 1996), perceptual information is scarcely transferable to other fields, or even across subspecialties in the same fields (e.g., Bilalić, McLeod, & Gobet, 2009; Rikers, Schmidt, & Boshuizen, 2002), because of its high specificity (Ericsson & Charness, 1994; Gobet, 2016).

2. Training Domain-General Cognitive Abilities: A Way to Get Around Far Transfer?

While domain-specific training rarely transfers across domains, some researchers have argued that training domain-general cognitive abilities – rather than domain-specific skills – can positively affect performance in a wide variety of fields that rely on those cognitive abilities. This idea can be considered a modern and more sophisticated version of *formal discipline theory* (James, 1890). According to formal discipline theory, cognition consists of a set of domain-general abilities (e.g., reasoning, memory, and concentration) that are thought to be malleable to training.

One theoretical foundation of the cognitive-training hypothesis is neural plasticity, that is, the ability of the neural system to adapt and modify under the pressure of the environment (Strobach & Karbach, 2016). Cognitive training is thought to lead to changes in the neural system, which, in turn, are supposed to account for the improvements on cognitive tests (Johnson, Munakata, & Gilmore, 2002; Karbach & Schubert, 2013). Another element in favour of the putative broad effects of cognitive training is that domain-general cognitive abilities correlate with performance in a wide variety of domain-specific skills. For example, fluid intelligence predicts academic achievement (Deary et al., 2007; Rohde & Thompson, 2007) and general intelligence is positively associated with job proficiency (Hunter & Hunter, 1984; Hunter, Schmidt, & Le, 2006). Thus, it is plausible to suggest that fostering overall cognitive ability by training affects people's academic and professional lives positively.

According to Taatgen (2016), there are two ways to train domain-general cognitive abilities: (a) deliberately training the particular skill(s) by practicing cognitive tasks (e.g., *n*-back in working memory training); or (b) engaging in cognitively demanding activities (e.g., playing chess in order to train spatial working memory and planning). While in the former case the improvement of general cognitive abilities is a direct consequence of training these

abilities, in the latter case it is the by-product of learning domain-specific skills. Either way, the enhancement of domain-general cognitive abilities is supposed to improve one's performance in activities requiring these cognitive abilities. Like the theories of expertise, cognitive-training theories acknowledge the fundamental role of domain-specific information in skill acquisition and expert performance. However, enhancing overall cognitive function is thought to facilitate and accelerate the acquisition of domain-specific skills in a broad range of areas. In other words, cognitive-training may make people smarter, and smarter people learn faster and more easily.

2.1 Mixed Effects and The Problem of Design Quality

Both methods have been extensively investigated. Research into working memory (WM) is a perfect example of the direct training of a particular cognitive ability. A classical result in cognitive psychology is that WM capacity strongly correlates with fluid intelligence (Kane, Hambrick, & Conway, 2005). Searching for a possible causal relationship, Jaeggi, Buschkuhl, Jonides, and Perrig (2008) tested the effects of WM training on a test of fluid intelligence (Raven's Progressive Matrices) in a sample of healthy adults. The treated participants showed a significant improvement compared to the control group. Following this experiment, the research has been extended to the effects of WM training on other cognitive abilities (e.g., cognitive control and spatial cognition) and academic achievement (e.g., mathematics, literacy). Despite the initial promising results, other studies have challenged the idea that WM training fosters a broad range of cognitive abilities (for a review, see Shipstead, Redick, & Engle, 2012). The topic is still lively debated, and no definitive conclusion has been reached.

When the focus shifts to the potential far-transfer effects of engaging in cognitively demanding activities, the story remains essentially unaltered. For example, the research on chess training has reported mixed results. While some authors express optimism about the

capability of chess to enhance cognitive abilities and academic achievement (e.g., Aciego, García, & Betancort, 2012; Taatgen, 2016), others seem more sceptical (e.g., Gobet & Campitelli, 2006; Sala, Foley, & Gobet, 2017). The same applies to the field of music training (for a review, see Miendlarzewska & Trost, 2013), video game training (e.g., Green et al., 2017; Redick, Unsworth, Kane, & Hambrick, 2017), brain-training (e.g., Anguera et al., 2013; Simons et al., 2016), and exergames (e.g., Mirelman et al., 2016; Stanmore, Stubbs, Vancampfort, de Bruin, & Firth, 2017).

The quality of the design may be a major source of such between-study variability. In the present dissertation, I examine two design-related features: (a) random (or non-random) allocation of the participants to the study groups and (b) type of control group (active or passive).

Randomization is essential to control for every potential difference at baseline. For example, non-random allocation may lead to differences between experimental and control groups in pre-test scores. In turn, such differences often produce statistical artefacts such as positive effects due to regression to the mean at post-test. In other words, the lack of randomization can inflate the effect size.

The type of control used to assess the effect of a particular treatment is important too. Researchers agree that passive control groups (i.e., no-contact or business-as-usual control groups) are not sufficient to establish the true impact of treatments (e.g., Moreau, Kirk, & Waldie, 2016). In fact, passive control groups do not control for non-specific factors such as placebo effects because simply belonging to a treatment group often affects behaviour. Conversely, the use of control groups receiving an alternative treatment (i.e., active control groups) contributes to rule out potential placebo effects. It is thus reasonable to expect that

comparing experimental groups to active control groups provides, on average, smaller effect sizes than passive control groups.

The quality of the study design has a crucial role in determining the size of the observed effect of an experimental intervention. Thus, the meta-analyses presented in this dissertation examine (when possible) whether type allocation and type of control group are moderating variables in the meta-analytical models.

3. Different Types of Cognitive Training

I now introduce the most common and studied types of cognitive training. The current state of the art and most relevant theories will be briefly discussed. For the detailed reviews, see Chapters 4 to 8.

As already mentioned, there are two possible ways to enhance cognition. The first method is to practice cognitive tasks such as n -back tasks in working memory training, mental rotation tasks in spatial training, and brain training games. The second method is to engage in intellectually demanding activities such chess, music, and video-games to train domain-general cognitive abilities.

3.1 Practicing Cognitive Tasks

3.1.1 Working Memory Training

Working memory is the cognitive system used to store and manipulate the information necessary to carry out cognitive tasks (Baddeley, 1992). A classical result in cognitive psychology is that fluid intelligence correlates with measures of working memory capacity (Engle, Tuholski, Laughlin, & Conway, 1999). Moreover, working memory capacity is also associated with measures of cognitive control such as the Stroop task (Kane & Engle, 2003), the go/no-go task (Redick, Calvo, Gay, & Engle, 2011), and the dichotic-listening task (Conway, Cowan, & Bunting, 2001).

WM capacity is also related to academic skills (e.g., Conway & Engle, 1996; Peng, Namkung, Barnes, & Sun, 2016), and plays a fundamental role in cognitive development. Children with reading difficulties (Swanson, 2006), mathematical disorders (Passolunghi, 2006), attention deficit/hyperactivity disorder (ADHD; Klingberg et al., 2005), and language impairment (Archibald & Gathercole, 2006) often suffer from deficits in working memory capacity.

Several researchers have thus proposed that increasing working memory capacity by training can enhance fluid intelligence (Jaeggi, Buschkuhl, Jonides, & Perrig, 2008) and boost cognitive control (Chein & Morrison 2010; for details, see 1. Introduction in Chapter 4). In turn, such an improvement is thought to transfer to other subject areas such as academic achievement and professional performance.

These hypotheses have been tested extensively. A vast body of research has been produced to determine whether working memory training can enhance fluid intelligence and, more generally, overall cognitive ability. Despite such impressive amount of experimental evidence, no definite conclusion has been reached. Also, the many meta-analyses and systematic reviews that have dealt with the topic have provided opposite results. While some of these reviews support the idea that WM training is a valuable tool for increasing fluid and enhancing overall cognitive function (Au et al., 2015; Au, Buschkuhl, Duncan, & Jaeggi, 2016; Klingberg, 2010; Morrison & Chein, 2011), others seem far more pessimistic (Dougherty, Hamovits, & Tidwell, 2016; Melby-Lervåg & Hulme, 2013, 2016; Melby-Lervåg, Redick, & Hulme, 2016; Schwaighofer, Fischer, & Buhner, 2015; Soveri, Antfolk, Karlsson, Salo, & Laine, 2017).

3.1.2 Brain Training

Brain training usually refers to those programs that convert cognitive tasks into computerized games (e.g., Lumosity®). The aim of brain-training programs is to enhance overall cognitive ability by practicing cognitive tasks. The basic assumption is that the improvements in the trained tasks generalize to real-life skills such as academic and professional attainment. In addition, this transfer is thought to be facilitated by the gamification of cognitive tasks that may encourage one's engagement in training such tasks (Anguera et al., 2013).

The research into the effects of brain training has mainly focused on adults and older adults in both clinical (e.g., schizophrenia, Alzheimer's disease, and brain trauma) and non-clinical populations (healthy participants). Despite the claims of the companies involved in the business, research has provided only modest evidence for the alleged cognitive benefits of brain-training programs. For example, in an influential study by Anguera et al. (2013), a small group of older adults played *NeuroRacer*, a multitasking brain-training program, and were compared to an active control group (single-task condition of the program) and a passive control group (no-contact). The multitasking group significantly outperformed the active control group in only three out of 11 cognitive tests. This outcome suggests that the treatment exerted little or no effect on the participants' overall cognitive function.

A recent systematic review of the literature (Simons et al., 2016) has provided further support to the hypothesis according to which brain-training programs do not provide any real benefit. The review points out that many brain-training studies lack proper controls, include very small samples (e.g., $N < 20$ per group), and do not report all the results of the outcome measures. Therefore, no definite conclusion can be drawn about this type of cognitive training until more powerful and better-designed studies are carried out.

3.2.2 Spatial Training

Another, relatively understudied, type of intervention to enhance cognitive ability is spatial training. Spatial training includes activities such as 2D and 3D mental rotation, spatial reasoning and visualizations (Sorby, 2011). Unlike working memory training and brain training, this type of cognitive-training intervention is often intended to enhance mathematical ability rather than overall cognitive function. However, given the difficulty of far transfer to take place, why should spatial training increase mathematical ability?

Problem solving in mathematics and STEM disciplines largely relies on spatial ability (Stieff & Uttal, 2015). Mechanical physics and engineering deal with movement and interaction between elements in a geometrical space. Mathematicians work with functions represented in 2D and 3D space. More generally, several branches of mathematics – necessary to master disciplines such as physics and engineering – require the manipulation of spatial relationships (e.g. geometry, calculus, topology).

The tight relation between spatial ability and mathematical ability has been established empirically. These two separate constructs are highly correlated to each other (Mix et al., 2016). Spatial abilities – such as mental rotation ability (Mix et al., 2016; Wai, Lubinski, & Benbow, 2009) – are thus strong predictors of achievement in mathematics, in children (Lauer & Lourenco, 2016), undergraduate and doctorate students (Wai et al., 2009). Thus, several researchers have suggested that training spatial ability causes improvement in mathematics achievement.

Before asking whether spatial training leads to improving mathematical skills such as arithmetic or geometry, one has to verify whether spatial ability can be trained. A meta-analysis carried out by Uttal et al. (2013) suggests that this is the case. Spatial training appears to transfer both to the trained tasks and other spatial tasks not directly trained.

Crucially, from a practical point of view, spatial ability seems to be malleable enough to be significantly boosted by a short-term training (Uttal et al., 2013).

The evidence supporting the effectiveness of spatial training at improving performance on spatial tasks appears to be quite solid. Regrettably, it is not possible to reach the same conclusion for non-spatial tasks. The research on spatial training to improve STEM achievement has provided promising results, but the number of studies is still relatively limited.

In Hsi, Linn, and Bell (1997), a group of undergraduates improved their attainment in an engineering course after attending a voluntary spatial training (3D orthographic projections). However, the fact that the sample was self-selected casts serious doubts upon the reliability of the outcome. More recently, Sorby (2009) reported that a group of undergraduates in engineering with low spatial ability improved their course grades after spatial training (Sorby, 2011), whereas a control group with no training did not show any amelioration. These positive findings were replicated two years later (Sorby, Casey, Veurink, & Dulaney, 2013). Less clear were the results in Miller and Halpern's (2013) study. They did find a moderate positive effect after delivering spatial training, but only in items related to Newtonian mechanics. No benefits occurred in other courses.

The studies mentioned above dealt with university students. Cheng and Mix (2014) focused on the effects of short-term (40 minutes) spatial training on children's basic arithmetical ability. The training consisted of 40 minutes of mental rotation and mental translation exercises suitable for children (Ehrlich, Levine, & Goldin-Meadow, 2006). The treatment group showed a small improvement (approximately $d = 0.20$) in the test of arithmetic, limited to one particular type of items (missing-term problems). A study by Hawes, Moss, Caswell, and Poliszczuk (2015) found no significant effects of mental rotation

training on a group of primary school children's arithmetical ability. Xu and LeFevre (2016) reported no transfer from spatial training to a number line task in a sample of kindergarten children. Finally, Sala, Bolognese, & Gobet (2017) tested the effects of one-hour mental rotation training on a sample of first-, second-, and third-grade children's arithmetical abilities. No significant effect was found overall.

In sum, the number of studies in this field is still too small to draw definite conclusions (or to do a meta-analysis). To date, the evidence suggests that spatial training can provide benefits for some specific spatial-related disciplines (e.g., mechanics) rather than overall cognitive function.

3.2 Engaging in Cognitively Demanding Activities

3.2.1 Chess Training

Students' poor achievement in mathematics has been the subject of debate both in the United States (Hanushek, Peterson & Woessmann, 2012; Richland, Stigler, & Holyoak, 2012) and in Europe (Grek, 2009). Researchers and policy makers have investigated alternative methods and activities with the purpose of improving the effectiveness of mathematics teaching. One such activity is play. The rationale is that, because children are highly motivated to play, they could learn important concepts in mathematics (and other curricular domains) without realizing it, through implicit learning (Brousseau, 1997); they could also acquire general cognitive skills such as concentration and intelligence, which would positively affect their school results generally.

Several authors have argued that chess is an ideal game for educational purposes (Bart, 2014; Jerrim, Macmillan, Micklewright, Sawtell, & Wiggins, 2016; Kazemi, Yektayar, & Abad, 2012). Chess offers an optimal trade-off between complexity and simplicity, and the balance between tactics and strategy is ideal. It combines numerical, spatial, temporal and combinatorial aspects. In addition, unlike games such as *awalé* and *Go*, the diversity of pieces

helps maintain attention – an important consideration with younger children. Altogether, these characteristics of chess may foster attention, problem solving, and self-monitoring of thinking (i.e., meta-cognition). Finally, there is some overlap between chess and mathematics (e.g., basic arithmetic with the value of the pieces, geometry of the board, and piece movements), which is an obvious advantage when using chess to foster mathematical skills. Thus, like working memory training and brain training, playing chess is meant to enhance domain-general cognitive abilities. In turn, these improvements (when any) are thought to foster children’s academic achievement in general and mathematical ability in particular.

In recent years, considerable efforts have been made to validate these ideas empirically. Not only has chess instruction been included in the school curriculum in several countries, but several educational projects and studies involving chess are currently ongoing or have recently ended in Germany, Italy, Spain, Turkey, the United Kingdom, and the United States. Even the European Parliament has expressed its interest and positive opinion on teaching chess in schools as an educational tool (Binev, Attard-Montalto, Deva, Mauro, & Takkula, 2011). If successful, using chess in school for fostering academic achievement would shed considerable light on the question of skill acquisition and transfer.

One psychological mechanism has been regularly proposed for explaining the putative effects of chess instruction: being a cognitively demanding activity, chess improves pupils’ domain-general cognitive abilities (e.g., intelligence, attention, and reasoning), abilities that then transfer to other domains, and therefore benefits a wide set of non-chess-related skills (e.g., Bart, 2014). The idea is intuitive and attractive. This view of chess as a cognitive enhancer has been mentioned in popular newspapers in the United Kingdom (e.g., Garner, 2012) and was the key theoretical assumption of a recent large experimental study that took place in the United Kingdom (Jerrim et al., 2016).

These explanations, albeit lacking detail, are plausible and provide the basis for the hypothesis that chess instruction strengthens cognitive abilities that are positively correlated to achievements in mathematics. Unfortunately, only a few studies have investigated the effects of chess on both cognitive abilities and academic outcomes. The results so far have been disappointing (Sala & Gobet, 2017d; Sala, Gobet, Trincherro, & Ventura, 2016; Scholz et al., 2008). In brief, the causal mechanisms remain substantially untested.

With regard to correlational evidence, a recent meta-analysis (Sala et al., 2017) reported that chess players outperformed non-chess players in several cognitive skills (e.g., planning, numerical ability, and reasoning). The difference between the two groups was approximately half a standard deviation. Another meta-analysis (Burgoyne et al., 2016) found positive correlations between chess skill and cognitive abilities such as fluid intelligence, processing speed, short-term and working memory (WM) memory, and comprehension knowledge.

However, the positive relationship between chess skill and cognitive ability does not necessarily imply that chess instruction enhances cognitive ability. An alternative explanation is that individuals with better cognitive ability are more likely to excel and engage in the game of chess. To establish causality, one needs to turn attention to studies where instruction is under experimental control. This will be the aim of Chapter 5.

3.2.2 Music Training

The idea that learning how to play an instrument improves one's cognitive abilities and academic achievement is extremely popular. Music ability is often associated with talent and superior cognitive skills. Blogs and newspapers often report enthusiastically on the benefits of music for the intellect (e.g., Costandi, 2016; Jaušovec & Pahor, 2017). Even the popular TV series *The Simpsons* has echoed this common belief by defining musical instruments as “the way to encourage a gifted child.”

However, how is music training supposed to provide such diverse benefits? Learning how to play a musical instrument engages executive functions such as cognitive control and working memory (Bialystok & Depape, 2009). Also, music training requires focused attention and learning complex visual patterns. Schellenberg (2004, 2006) has thus proposed that the most likely explanation for the presumed broad set of benefits provided by music training is that it enhances individuals' overall cognitive function and general intelligence. These cognitive skills are major predictors of academic achievement (e.g., Deary et al., 2007), and it might be the case that some domain-specific abilities acquired by music training generalize to other non-music skills.

One further theoretical foundation of the hypothesis according to which music training exerts a positive influence on overall cognitive ability is neural plasticity. In fact, musicians do exhibit specific anatomical and functional neural patterns. An increased density of grey matter in musicians has been observed in areas involved in cognitive skills such as auditory localization (right Heschl's gyrus; Bermudez, Lerch, Evans, & Zatorre, 2009) and language production (Broca's area; Sluming et al., 2002). With regard to functional differences, expert musicians seem to show, for example, an enhanced bilateral activation of the Rolandic operculum (for a review, see Neumann, Lotze, & Eickhoff, 2016). Probably, this activation reflects superior ability in the processing of auditory information (Koelsch, Fritz, von Cramon, Müller, & Friederici, 2006).

The hypothesis that music training induces significant anatomical and functional changes in the brain which, in turn, lead to increased cognitive function, seems plausible. Also, the improvements in cognitive ability are claimed to be both domain-specific – such as superior memory for music-related material (Sala & Gobet, 2017a) – and domain-general (e.g., fluid intelligence; Schellenberg, 2004, 2006).

Like in chess, a link between superior cognitive ability and music skill does exist. In a study by Ruthsatz, Detterman, Griscom, and Cirullo (2008), a group of professional musicians outperformed a group of novices in a standardized measure of fluid intelligence (Raven's Progressive Matrices). Also, Lee, Lu, and Ko (2007) found a correlation between music skill and working memory. Finally, Schellenberg (2006) reported positive, yet moderate, correlations between engagement in musical activities and IQ in a group of children and undergraduates. Critically, this positive relationship remained even after controlling for parental income and education. This finding was replicated in a more recent study concerning 7- and 8-year-old children (Schellenberg & Mankarious, 2012).

Other correlational studies have shown that music ability is associated with academic skills as well. Anvari, Trainor, Woodside, and Levy (2002) found that music perception skills correlated with reading abilities in preschool children. Similarly, Forgeard et al. (2008) reported that music discrimination skill correlated with phonological processing ability in a group of dyslexic and typically-developing children. With regard to mathematical ability, Cheek and Smith (1999) found that students who had received private music lessons achieved better results in the mathematics portion of the Iowa Test of Basic Skills. In line with the latter three studies, Wetter, Koerner, and Schwaninger (2009) reported a positive relationship between engagement in musical activities and overall academic attainment.

3.2.3 Video-Game Training and Exergames

Along with working memory training, video-game training is the most studied, influential, and debated type of cognitive training. It is customary to distinguish between two categories of video games: action video games and non-action video games. Since the publication of Green and Bavelier's (2003) seminal article, action video games have attracted the attention of many researchers in the field. The practice of action video games such as *Unreal Tournament 2004* and *Call of Duty 2* has been claimed to improve a variety of

perceptual and attentional tasks. This improvement seems to occur in both quasi-experimental studies – when regular video-game players are compared to non-players – and experimental studies, when non-players are trained with action video games and compared to control groups of non-action video game players (e.g., Bejjanki et al., 2014).

The “learning to learn” theory (Green, Gorman, & Pouget, & Bavelier, 2016) is the most influential attempt to explain such results. According to this theory, the practice with action video games leads to an improvement in probabilistic inference. It is proposed that playing action video games makes people better at using and processing information. Then, this ability can be transferred to other tasks (e.g., go/no-go and enumeration tasks). This theory thus postulates the existence of a general learning system that can be trained by the practice of action video games. Training this system allows one to extract and elaborate relevant information from the environment more efficiently and, hence, learn to perform a task more quickly. Put simply, action video games are claimed to improve the computational ability of the brain in general.

Non-action video game training, albeit relatively understudied, has been claimed to provide some cognitive benefits as well. For example, Okagaki and Frensch (1994) reported that playing Tetris improved the spatial abilities in a group of older adolescents. Also, Basak, Boot, Voss, and Kramer (2008) found positive effects of the practice of a real-time strategy video game (*Rise of Nations*) on measures of short-term memory and spatial ability in a group of older adults.

However, several studies have challenged the idea that video-game training can positively impact on domain-general cognitive ability with regard to both action and non-action video games. For example, Terlecki, Newcombe, and Little (2008) found no effect of playing *Tetris* on mental rotation. Similarly, Minear et al.’s (2016) study failed to show

any significant improvement in several measures of memory, spatial ability, and fluid intelligence in individuals practicing a real-time strategy video game (*Starcraft: Brood War*). The lack of replication of the initial positive results applies to action video-game training as well. No significant effect of action video game training was found in cognitive tests such as span and *n*-back tasks, enumeration, and perceptual tasks (e.g., Boot, Kramer, Simons, Fabiani, Gratton, & 2008; van Ravenzwaaij, Boekel, Forstmann, Ratcliff, & Wagenmakers, 2014). Given the inconsistent results in the literature, Oei and Patterson (2013, 2014, 2015) have offered an explanation alternative to the “learning to learn” theory. Action video game training may foster only those skills necessary to engage in particular video games. This hypothesis is consistent with Thorndike and Woodworth’s (1901) common elements theory.

Several meta-analyses have addressed the question of the effects of video-game training on cognitive abilities (Powers & Brooks, 2014; Powers, Brooks, Aldrich, Palladino, & Alfieri, 2013; Toril, Reales, & Ballesteros, 2014; Wang et al., 2016). All these meta-analyses report positive overall effect sizes suggesting that video-game training (both action and non-action) has some impact on cognitive function. However, these meta-analyses suffer from several major methodological weaknesses that do not allow us to draw any reliable conclusion (for details, see 2. The Meta-Analytical Evidence in Chapter 7).

A further source of scepticism comes from several recent cross-sectional and correlational studies. For example, Gobet et al. (2014) found no differences between a group of action video game players and a group of non-players in a flanker task and a change detection task. Similar results were obtained in other investigations (e.g., Castel, Pratt, & Drummond, 2005; Irons, Remington, & McLean, 2011; Murphy & Spencer, 2009). Finally, Unsworth et al. (2015) found near-zero correlations between video game experience and several measures of processing speed, WM capacity, and fluid reasoning in a large sample of adults. Green et al. (2017) have questioned Unsworth and colleagues’ findings (see also

Redick, et al., 2017). The debate is still ongoing, and the question of the alleged cognitive benefits of video-game training is yet to be solved.

3.2.3.1 Exergames

Exergames are probably the most recent type of cognitive-training programs that has been undergone experimental research. Exergames are video games combining cognitive and physical training. The rationale behind such games is to exploit the benefits of physical exercise (Fabel et al., 2009; Firth et al., 2016; Kempermann et al., 2010), cognitive exercise, and trainees' engagement stemming from the gamification of the tasks (Stine-Morrow et al., 2014). Examples of such training regimens are interactive dancing, "cyber-cycling," and walking on a treadmill in a virtual environment.

A recent meta-analysis (Stanmore et al., 2017) has examined the impact of exergames on cognitive function and found positive effects. However, this meta-analysis suffers from severe flaws that have probably biased the results. This topic is covered in detail in Chapter 8.

Chapter 3: Meta-Analytic Techniques

This chapter summarizes all the techniques used in the meta-analyses (Chapters 4 to 8). These techniques include formulas to calculate effect sizes, types of meta-analytical, moderator and publication bias analyses, and other methods to correct effect sizes. As mentioned in Chapter 1, it is necessary to know how these techniques work to understand the findings presented in this dissertation. For an extensive discussion of these formulas and meta-analytical techniques and models, see Schmidt and Hunter (2015).

1. Effect Sizes

The extraction of effect sizes is necessary to compare data from different studies and tests. Thus, the correct calculation of effect sizes is fundamental to avoid biased results. For the correct calculation of effect sizes in studies with an *only-post-test design*, the standardized means difference (Cohen's d) was calculated with the following formula:

$$d = (M_e - M_c) / SD_{pooled} \quad (1)$$

where SD_{pooled} is the pooled standard deviation and M_e and M_c are the means of the experimental group and the control group, respectively.

For the studies with a *repeated-measure design*, the standardized means difference was calculated with the following formula:

$$d = (M_{g-e} - M_{g-c}) / SD_{pooled-pre} \quad (2)$$

where $SD_{pooled-pre}$ is the pooled standard deviation of the two pre-test standard deviations, and M_{g-e} and M_{g-c} are the gain of the experimental group and the control group, respectively.

For the studies with an *ANCOVA design*, the standardized means difference was calculated with the following formula:

$$d = (M_{adj-e} - M_{adj-c}) / SD_{pooled-pre} \quad (3)$$

where SD_{pooled} is the pooled standard deviation of the two standard deviations of the unadjusted means, and $M_{adj-e} - M_{adj-c}$ are the adjusted means of the experimental group and the control group, respectively.

When means and standard deviations were not available, t -statistics referring to pre-post improvements within groups were converted to d s and then subtracted to calculate the standardized mean difference between the experimental and control groups. Alternatively, the statistics referring to between-group differences at pre- and post-tests were converted to d s and then subtracted. The conversion formula was:

$$d = t \times \sqrt{(N_e + N_c) / (N_e \times N_c)} \quad (4)$$

where N_e and N_c are the total sample size of the experimental group and control group, respectively.

The standard error of Cohen's d s was calculated with the following formula:

$$st. \text{ err.} = \sqrt{\frac{N}{N_e \times N_c} + \frac{d^2}{N \times 2}} \quad (5)$$

where N , N_e , and N_c are the total sample size of the study, experimental group, and control group, respectively.

When correcting for the upward bias, Cohen's d s were converted into Hedges's g by using the following formula:

$$g = d * \left(1 - \left(\frac{3}{4 \times N - 9}\right)\right) \quad (6)$$

where N is the sample size of the study. The same correction was applied to standard errors.

2. Fixed and Random Effect Models

Fixed-effect meta-analytic models assume that all the included studies share a common effect size. In other words, the true effect is believed to be the same in all the studies. Consequently, the difference between the observed effects is due to random error in fixed-effect meta-analyses.

This assumption is not always met. The included studies share a set of common features (the inclusion criteria). However, there is generally no reason to assume that the true effect is the same across all the studies. More realistically, some factors (e.g., populations' age, duration of interventions, different settings, etc.) may exert an influence on the effect sizes. In this case, the overall effect size does not represent a single true effect. Rather, the overall effect size is the product of several true effects.

Given that assuming only one true effect is a severe constraint, random-effect models allow the potential occurrence of a distribution of true effect sizes (Borenstein, Hedges, Higgins, & Rothstein, 2009). More specifically, every effect size is the combination of its true effect and within-study error. The true effect is, in turn, determined by the overall effect size and between-study error. Due to their superior flexibility, random-effect models were used in all the meta-analyses of the present dissertation.

2.1 Assessing Heterogeneity

As just mentioned, the observed overall effect size is sometimes the mean of a series of true effects rather than *the* true effect. It is thus imperative to evaluate whether between-study variability is due only to random error or some moderating factor. Moreover, it is necessary to estimate the *ratio* of between-study variability explained by random error and moderating factors.

To address this issue, meta-analysts use the test of heterogeneity and report the values of the I^2 statistic (for details, see Schmidt & Hunter, 2015). The I^2 statistic refers to the percentage of between-study variance due to true heterogeneity and not to random error (Higgins, Thompson, Deeks, & Altman, 2003). The higher the value of the I^2 statistic, the higher the percentage of between-study variance due to true heterogeneity. When I^2 is zero, between-study error is zero. Consequently, in this case, random-effect models and fixed-effect models produce the same results.

2.2 Moderator Analysis

In the presence of true heterogeneity, moderator analysis (or meta-regression) is run to investigate the potential role of several study-related factors, that is, the moderators, in determining the size of the effects. This technique is the meta-analytic homologous of linear multiple-regression analysis. In fact, while in primary studies the unit of analysis is usually the subject, in moderator analysis the unit is the effect size.

Like independent variables in a regression model, moderators are chosen by the researcher to test specific hypotheses and control for potential confounding effects. Put simply, the choice of what moderators should be included in the meta-regression model should always be theory-driven.

3. Publication Bias

Publication bias occurs when studies with small samples and small effect sizes are systematically suppressed from the literature. Thus, in the presence of publication bias, overall effect sizes tend to be greater than the true effects. There are numerous techniques to detect publication bias and estimate a corrected overall effect size.

3.1 Trim-and-Fill

Trim-and-fill analysis (Duval & Tweedie, 2000) estimates the symmetry of a funnel plot representing the relation between effect size and standard error. In the presence of publication bias, effect sizes are missing from the bottom left part of the funnel plot (small effect sizes with high standard error; e.g. Figure 16). That is, when standard error is high, larger-than-average effects sizes (those on the bottom right) are more likely to be published than smaller-than-average effect sizes (those on the bottom left). The trim-and-fill analysis estimates the number of missing studies from the funnel plot and imputes the missing effect sizes based on the observed data's asymmetry to create a more symmetrical funnel plot and calculate a corrected overall effect size.

3.2 PET-PEESE

PET estimator is the intercept of a weighted linear regression where the dependent variable is the effect size, the independent variable is the standard error, and the weight is the inverse of the standard error squared. PEESE estimator is obtained by replacing the standard error with the standard error squared as the independent variable. If PET suggests the presence of a real effect (i.e., intercept different from zero), PEESE estimator must be considered as the corrected overall effect size (Stanley & Doucouliagos, 2014).

3.3 Begg and Mazumdar's (1994) Rank Correlation Test

If publication bias occurs, this test assumes that there will be an inverse correlation between standard error (which is driven primarily by sample size) and effect size. The rank order correlation (Kendall's tau) between the treatment effect and the standard error tells us whether publication bias occurs. However, this test does not provide a corrected overall effect size.

3.4 P-Curve

P-curve tests the presence of publication bias by analysing the distribution of *only* statistically significant p -values (i.e., $ps < .05$) associated with the effect sizes (Simonsohn, Nelson, & Simmons, 2014). The key assumption of this method is that real effects tend to be highly significant ($p < .01$). Thus, if the p -values distribution is flat or left-skewed (i.e., no difference or greater number of large p -values than small p -values), then publication bias is likely. By contrast, if the distribution is right-skewed (i.e., more small p -values than large p -values), then publication bias is unlikely.

3.5 Egger's Regression Test

In this method (Egger, Smith, Schneider, & Minder, 1997), the inverse of the standard error of the effect size (i.e., precision) is used as an independent variable in a regression to predict the “standardized effect” – i.e., the effect size divided by its standard error. If the intercept of this regression is zero, then there is no publication bias. By contrast, a positive value for the intercept indicates the presence of publication bias because small- N studies are associated with larger effect sizes.

3.6 Selection Models

Vevea and Woods's (2005) selection model analysis estimates four adjusted values by pre-weighted functions of p -values' distributions. These distributions represent different patterns of possible publication bias. If all (or most of) the four adjusted values are shown not to differ significantly from the overall effect size, then it can be reliably concluded that the results are not affected by publication bias (Schmidt & Hunter, 2015). Notably, this analysis stays reliable even when the number of effect sizes is modest.

4. Statistical Dependence of Effect Sizes

In the meta-analyses presented in this dissertation, the effect sizes were calculated for each dependent variable reported in the studies. For each independent sample, those effect sizes referring to the same type of measure (e.g., reaction times) and extracted from the same test (e.g., different subscales) were meta-analytically merged into one effect size. This procedure was used to calculate more reliable estimates and reduce the number of statistically dependent effect sizes in the model (Schmidt & Hunter, 2015).

For those effect sizes that were statistically dependent and referred to different constructs or were extracted from different tests, Cheung and Chan's (2004) correction for statistically dependent samples was applied. This method decreases the weight of dependent samples in the analysis by calculating an adjusted (i.e., smaller) N in each meta-analytic model.

The violation of the assumption of statistical independence does not necessarily cause a systematic bias in the estimation of overall meta-analytic means. However, the violation of the assumption of statistical independence is associated with an underestimation of sampling error inflating the variability between studies (Schmidt & Hunter, 2015), with possible consequent biases in moderator analysis.

Therefore, Cheung and Chan's (2004) method allows one to build more powerful models without losing any information from the primary studies, biasing the meta-analytic mean, or artificially inflating the degree of heterogeneity.

5. Techniques for Detecting Outliers

5.1 Winsorizing

Winsorizing (Lipsey & Wilson, 2001; Tukey, 1962) is the reduction of extreme values to reduce the effect of potential outliers on overall results. The definition of extreme values

is, to some extent, subjective and depends on the distribution of the effect sizes. Therefore, the major limitation of this procedure is its arbitrariness.

5.2 Influential Case Analysis

Viechtbauer and Cheung's (2010) analysis of influential cases estimates whether some effect sizes have a significantly greater impact on the overall effect size compared to the other effect sizes in a model. Such impact can be due to the size of the effect or its weight (i.e., large sample size). The main advantage of this technique is that influential cases are detected via a series of estimated parameters rather than the meta-analyst's subjective judgment.

Chapter 4: Meta-Analysis of Working Memory Training

Rationale for the Meta-Analysis in Chapter 4

Chapter 4 reports a meta-analysis on the effect of working memory training on typically developing children's cognitive abilities and academic achievement. As mentioned in Chapter 2, there is substantial disagreement among researchers about the actual benefits of working memory training. Typically developing children are the ideal population to test the potential of working memory training as a cognitive enhancer. In fact, a child's brain tends to be more malleable to training than an adult's one. Assuming that the benefits of cognitive training are mediated by neural plasticity, the occurrence of far-transfer effects should be more likely in children than adults. Therefore, a null result would represent robust evidence against the alleged cognitive benefits of working memory training.

The studies included in this meta-analysis are listed in Appendix A.

1. Introduction

Transfer of learning occurs when a set of skills acquired in a particular domain generalizes to other domains. The occurrence of transfer is either a tacit assumption or a deliberate objective of most educational interventions: any learned skills are meant to be applied beyond the learning context (Perkins & Salomon, 1994). For example, one's ability in analytic geometry is supposed to generalize to calculus.

According to Thorndike and Woodworth's (1901) common element theory, transfer is a function of the extent to which two tasks share common features and cognitive elements. In accordance with this hypothesis, while *near-transfer* – i.e., the transfer of skills between strictly related domains (e.g., analytic geometry and calculus) – takes place frequently, *far-transfer* – i.e., the transfer occurring between source and target domains weakly related to each other (e.g., Latin and mathematics) – has rarely been observed (Donovan, Bransford, & Pellegrino, 1999). Examples of failed far-transfer include teaching the computer language LOGO to improve children's reasoning skills (De Corte & Verschaffel, 1986; Gurtner, Gex, Gobet, Nunez, & Restchitzki, 1990) and, as reported in a recent meta-analysis (Sala & Gobet, 2016), teaching chess to improve children's cognitive and academic skills.

The training investigated in those studies was highly specific (learning a programming language and chess, respectively). However, it is possible that boosting a domain-general cognitive mechanism is an effective way to improve other cognitive and real-life skills, such as academic achievement. This assumption is the key principle underlying the research on WM training.

1.1 Working Memory Training

WM is the cognitive system used to store and manipulate the information necessary to carry out cognitive tasks (Baddeley, 1992). Measures of WM capacity, such as the number of

items WM can store and the ability to keep information in active memory during interfering tasks, correlate positively with fluid intelligence (Engle, Tuholski, Laughlin, & Conway, 1999) and measures of cognitive control such as the Stroop task (Kane & Engle, 2003), the go/no-go task (Redick, Calvo, Gay, & Engle, 2011), and the dichotic-listening task (Conway, Cowan, & Bunting, 2001). In addition, WM capacity is related to academic skills such as reading comprehension (Conway & Engle, 1996) and mathematical ability (Peng, Namkung, Barnes, & Sun, 2016). WM also seems to play a fundamental role in cognitive development. Deficits in WM capacity in children are associated with reading difficulties (Swanson, 2006), mathematical disorders (Passolunghi, 2006), attention deficit/hyperactivity disorder (ADHD; Klingberg et al., 2005), and language impairment (Archibald & Gathercole, 2006).

Several hypotheses have linked WM to intelligence and academic achievement. It has been proposed that WM and fluid intelligence share a common capacity constraint (Halford, Cowan, & Andrews, 2007). The amount of information (e.g., the number of items) that can be handled in WM is limited. Consequently, the number of interrelationships among elements that can be held and manipulated by WM in a reasoning task (e.g., Raven's progressive matrices) is bounded. If such limits are alleviated by training, then an improvement in fluid intelligence might occur (Au et al., 2015; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008). Crucially, such an improvement is supposed to generalize to subject areas such as mathematics or literacy, because fluid intelligence is a key predictor of academic achievement (Deary, Strand, Smith, & Fernandes, 2007; Rohde & Thompson, 2007). Another related hypothesis concerns the role of attentional control processes in both working memory and fluid intelligence (Gray, Chabris, & Braver, 2003). Chein and Morrison (2010), for example, have suggested that WM training induces positive effects on measures of cognitive control (e.g., Go/no-go, Stroop task), which, in turn, boosts performance in other tasks outside the domain of WM. Finally, it has been hypothesized that WM training is especially

beneficial for individuals with low WM capacity (e.g., children with ADHD or other learning disabilities). The idea is simple. If one's learning difficulties stem from reduced WM capacity, then training that specific skill might help to improve academic performance. The common assumption underlying these three hypotheses is that WM training boosts domain-general mechanisms (WM capacity, cognitive control, and attention), and hence enhances many other cognitive and academic skills.

However, in spite of a vast amount of research, no definite conclusion on the putative effectiveness of WM training at boosting cognitive skills and academic achievement has been reached yet. There is substantial agreement about the existence of near-transfer effects due to WM training – such as improvements in measures of verbal and non-verbal WM and short-term memory. However, while several reviews of the available experimental evidence have upheld the idea that WM training is a valuable cognitive enhancement tool (Au et al., 2015; Au, Buschkuhl, Duncan, & Jaeggi, 2016; Klingberg, 2010; Morrison & Chein, 2011), others have challenged the hypothesis according to which WM training effects substantially transfer to other cognitive skills outside the domain of WM (Dougherty, Hamovits, & Tidwell, 2016; Melby-Lervåg & Hulme, 2013, 2016; Melby-Lervåg, Redick, & Hulme, 2016; Redick, Shipstead, Wiemers, Melby-Lervåg, & Hulme, 2015; Schwaighofer, Fischer, & Buhner, 2015; Shipstead, Redick, & Engle, 2010, 2012).

1.2 Working Memory Training in Children

Children represent an important population on which to test the ability of WM training to boost cognitive and academic skills. During childhood, cognitive ability and academic skills are still at the beginning of their development, and, thus, cognitive training is likely to be more efficient than in adulthood. In agreement with this idea, research into expertise has clearly established that the likelihood of far-transfer is inversely related to the level of expertise in a discipline, which needs several years to acquire (Ericsson & Charness,

1994; Gobet, 2016). That is, WM training is more likely to improve, for example, a child's basic arithmetic abilities than an undergraduate student's skill in solving differential equations. In fact, while the skill to develop is quite general and based to some extent on cognitive ability in the former case, it depends to a large extent on domain-specific knowledge in the latter case. Thus, from a theoretical point of view, children are an ideal population to test the occurrence of transfer.

Several recent reviews have addressed the issue of the putative benefits of WM training in children, without reaching any agreement. According to Klingberg (2010), WM training can be used as an effective remediating intervention. By contrast, Rapport, Orban, Kofler, and Friedman's (2013) meta-analysis reported little or no evidence of amelioration in academic achievement in children with ADHD after WM training. In line with Rapport et al.'s (2013) results, Redick et al.'s (2015) review showed that WM training did not provide any benefit to academic performance in children with ADHD (e.g., Chacko et al., 2014) and poor WM (e.g., Ang, Lee, Cheam, Poon, & Koh, 2015), or in typical developing children (e.g., Rode, Robson, Purviance, Geary, & Mayr, 2014).

Evaluating the effects of WM training on children with no learning disability has substantial practical and theoretical implications. If a brief training can improve overall cognitive ability and academic achievement, the impact of such an intervention on educational practices and policies would be profound. Any positive effect of WM training would provide an advantage for a vast cohort of individuals, not just for a relatively small sub-sample (children with ADHD or children with poor WM). However, it is yet to be established whether increasing WM capacity in typically developing (TD) children with no WM impairment can enhance academic achievement and cognitive abilities outside the domain of WM. The aim of the present study is to quantitatively evaluate the available evidence via meta-analysis.

1.3 The Present Meta-Analysis

The present meta-analysis focuses on the putative effectiveness of WM training at enhancing cognitive and academic skills in TD children. While several previous meta-analyses (e.g., Melby-Lervåg & Hulme, 2013; Melby-Lervåg et al., 2016; Schwaighofer et al., 2015) included studies dealing with the putative benefits of WM training in TD children, no meta-analysis has yet been specifically devoted to this issue.¹

The main purpose of this meta-analysis is to estimate the overall effect sizes obtained with WM training with respect to near-transfer (i.e., WM-related outcomes) and far-transfer (i.e., outcomes outside the domain of WM). Also, we aimed to test the possible effects of several moderators, with particular attention to far-transfer measures (e.g., fluid intelligence, cognitive control, and academic achievement measures). Therefore, the meta-analysis followed five steps. First, to estimate the presence or absence of near-transfer and far-transfer at the end of the intervention, we calculated the overall standardized difference between WM training groups and control groups on (a) near-transfer measures (e.g., visuospatial working memory, short-term memory) and (b) measures related to abilities outside the domain of WM (e.g., fluid intelligence, cognitive control, mathematics).

Second, we carried out a moderator analysis. As noted in previous meta-analyses (e.g., Melby-Lervåg & Hulme, 2013; Schwaighofer et al., 2015), two methodological features may be a major source of variability between intervention studies—random assignment to groups and the presence of an active control group to control for potential confounding

¹ Weicker, Villringer, and Thöne-Otto's (2016) meta-analysis reported several overall effect sizes regarding the effect of WM training on TD children's cognitive abilities such as fluid intelligence and processing speed. However, the total sample included only nine studies.

effects (e.g., differences at baseline level between experimental and control groups, Hawthorne effect). The absence of these features may result in an inflation of the positive effects of the training due to confounds such as differences at baseline level, self-selection of the treated sample, and placebos. Therefore, we evaluated the potential moderating effects of the type of control group (active or passive control group) and the presence of randomization for the assignment to the groups. We also investigated the potential moderating effects of the age of the participants and the total duration of the training. Third, we focused on the far-transfer effects and investigated whether WM training is more (or less) successful in boosting particular academic/cognitive skills. Fourth, we performed publication bias analyses. Finally, we calculated the follow-up overall effect sizes for near- and far-transfer measures.

2. Method

2.1 Literature Search

In accordance with the PRISMA statement (Moher, Liberati, Tetzlaff, & Altman, 2009), a systematic search strategy was used to find the pertinent studies. Using several combinations of the terms “working memory,” “training,” “cognitive,” “intervention,” and “children”, we searched Scopus, ERIC, Psyc-Info, ProQuest Dissertation & Theses, and Google Scholar databases to identify all the potentially relevant studies. Also, earlier narrative reviews were examined, reference lists were scanned, and we e-mailed scholars in the field ($n = 13$) requesting unpublished studies and inaccessible data.

2.2 Inclusion/Exclusion Criteria

The studies were included according to the following six criteria:

1. The design of the study included an intervention aimed to train working memory skills (e.g., verbal working memory, visuospatial working memory); correlational and ex-post facto studies were excluded;

2. The study presented a comparison between a treated group and at least one control group;
3. During the study, a measure of academic or cognitive skill other than working memory was collected; importantly, to assess a genuine near-transfer effect, all the measures of performance in the trained WM intervention task were excluded;
4. The participants in the study were aged three to sixteen;
5. The participants in the study were TD children without any specific learning disability (e.g., ADHD) or borderline cognitive ability (e.g., low IQ, poor working memory capacity);²
6. The data presented in the study (or provided by the author) were sufficient to calculate an effect size.

To identify studies meeting these criteria, we searched for relevant published and unpublished articles through April 1, 2016. We found 25 studies, conducted from 2007 to 2016, that met all the inclusion criteria. These studies included 26 independent samples and 104 effect sizes (30 for WM-related measures, see Table 1; 74 for non-WM-related measures, see Table 2), with a total of 1,601 participants. Finally, a subsample of the included studies ($n = 6$) reported follow-up effects. A total of 30 follow-up effect sizes were computed (6 for WM-related measures, see Table 3; 24 for non-WM-related measures, see Table 4), with a total of 249 participants.³ The entire procedure is summarized in Figure 1.

² In Shavelson, Yuan, Alonzo, Klingberg, and Andersson (2008), eight participants (out of 37) had ADHD or learning difficulties. Since separate results were not available, we calculated the effect sizes considering the whole sample of 37 participants.

³ In Soderqvist and Bergman-Nutley (2015), no post-test assessment was administered immediately after the training, but only 24 months later. Thus, we included the effect sizes

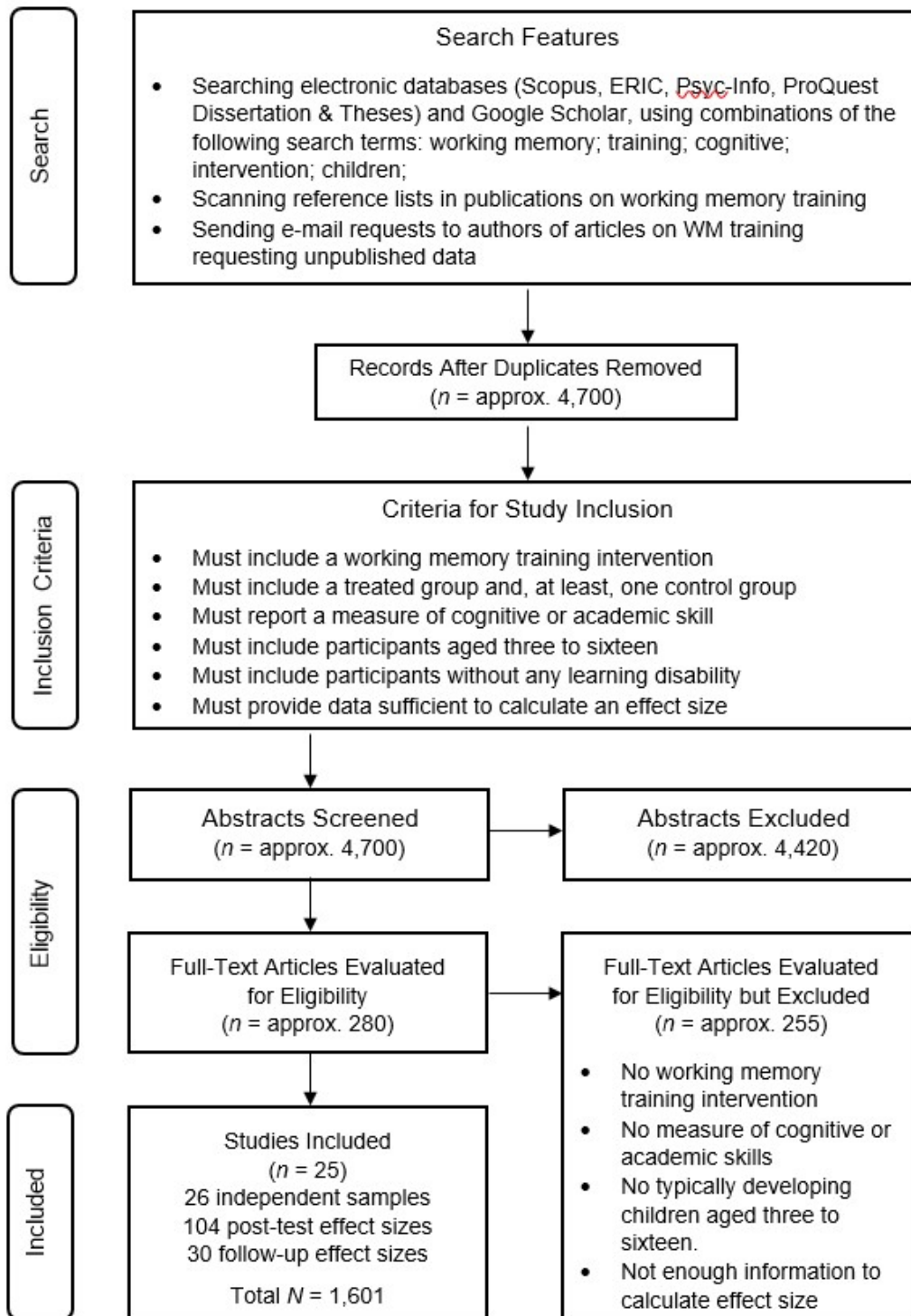


Figure 1. Flow diagram of the studies included in the meta-analytic review.

extracted from this study in both the main models and the follow-up models.

2.3 Moderators

We selected five potential moderators:

1. Random allocation (dichotomous variable): Whether the participants were randomly allocated to the groups;
2. Type of control group (active or passive; dichotomous variable): Whether the WM training-treated group was compared to another activity;
3. Duration of training (continuous variable): The total time of training in hours;
4. Age (continuous variable): The mean age (in years) of the participants; when the mean age was not provided ($n = 3$) we used either the median age ($n = 1$) or an age estimation based on the school grade ($n = 2$; e.g., third graders = 9-year-olds);
5. Domain (categorical variable): This variable, which was inserted only in the far-transfer model, includes literacy/word decoding, mathematics, science, fluid intelligence, crystallized intelligence, and cognitive control.⁴

The two authors coded each effect size for moderator variables independently. There was no disagreement with respect to Random allocation, Type of control group, and Age.

Regarding the moderator Duration of training, 87% agreement was obtained. For the moderator Domain, the Cohen's kappa was $\kappa = .95$. The authors resolved every discrepancy.

⁴ These broad categories were built by aggregating different outcomes related to a particular domain (e.g., go/no-go task and Stroop task under the category of cognitive control). For all the details about the reviewed studies, see Tables S1.1 to S1.4 in the Supplemental material available online.

Table 1

Studies and moderators of the 30 near-transfer effect sizes included in the meta-analysis

| Study | Age | Duration of training | Random allocation | Type of control group |
|-----------------------------------------------|------------|-----------------------------|--------------------------|------------------------------|
| Bergman-Nutley et al. (2011) - M1 | 4.27 | 6.25 | Yes | Active |
| Bergman-Nutley et al. (2011) - M2 | 4.27 | 6.25 | Yes | Active |
| Henry, Messer, & Nash (2014) | 7.00 | 3.00 | Yes | Active |
| Karbach, Strobach, & Schubert (2015) | 8.30 | 9.33 | Yes | Active |
| Kroesbergen, Noordende, & Kolkman (2014) - M1 | 5.87 | 4.00 | Yes | Passive |
| Kroesbergen, Noordende, & Kolkman (2014) - M2 | 5.87 | 4.00 | Yes | Passive |
| Kuhn & Holling (2014) - S1 | 9.00 | 5.00 | Yes | Passive |
| Kuhn & Holling (2014) - S2 | 9.00 | 5.00 | Yes | Active |
| Kun (2007) - S1 - M1 | 12.84 | 8.00 | Yes | Active |
| Kun (2007) - S1 - M2 | 12.84 | 8.00 | Yes | Active |
| Kun (2007) - S2 - M1 | 13.52 | 14.58 | Yes | Active |
| Kun (2007) - S2 - M2 | 13.52 | 14.58 | Yes | Active |
| Kun (2007) - S2 - M3 | 13.52 | 14.58 | Yes | Active |

| | | | | |
|-------------------------------------------------------------|-------|-------|-----|---------|
| Lee (2014) | 9.00 | 3.00 | Yes | Active |
| Lindsay (2012) | 5.49 | 3.00 | Yes | Active |
| Passolunghi & Costa (2016) - S1 - M1 | 5.44 | 10.00 | Yes | Active |
| Passolunghi & Costa (2016) - S1 - M2 | 5.44 | 10.00 | Yes | Active |
| Passolunghi & Costa (2016) - S2 - M1 | 5.42 | 10.00 | Yes | Passive |
| Passolunghi & Costa (2016) - S2 - M2 | 5.42 | 10.00 | Yes | Passive |
| Pugin et al. (2015) - M1 | 13.00 | 8.05 | No | Passive |
| Pugin et al. (2015) - M2 | 13.00 | 8.05 | No | Passive |
| Rode, Robson, Purviance, Geary, & Mayr (2014) | 9.00 | 7.14 | Yes | Passive |
| Shavelson et al. (2008) - M1 | 13.50 | 14.58 | Yes | Active |
| Shavelson et al. (2008) - M2 | 13.50 | 14.58 | Yes | Active |
| St Clair- Thompson, Stevens, Huth, & Bolder (2010) | 6.83 | 6.00 | No | Passive |
| Studer-Luethi, Bauer, & Perrig (2016) - S1 | 8.25 | 4.50 | Yes | Active |
| Studer-Luethi, | 8.25 | 4.50 | Yes | Passive |

Bauer, & Perrig

(2016) - S2

Thorell,

Lindqvist,

Bergman, Bohlin,

& Klingberg

(2008) - S1

4.67

6.25

No

Active

Thorell,

Lindqvist,

Bergman, Bohlin,

& Klingberg

(2008) - S2

4.67

6.25

No

Passive

Witt (2011)

9.68

7.50

No

Passive

Table 2

Studies and moderators of the 74 far-transfer effect sizes included in the meta-analysis

| Study | Age | Duration of training | Random allocation | Type of control group | Domain |
|-------------------------------------------------|------------|-----------------------------|--------------------------|------------------------------|--------------------|
| Bergman-Nutley et al. (2011) | 4.27 | 6.25 | Yes | Active | Fluid intelligence |
| Henry, Messer, & Nash (2014) - M1 | 7.00 | 3.00 | Yes | Active | Literacy/WD |
| Henry, Messer, & Nash (2014) - M2 | 7.00 | 3.00 | Yes | Active | Mathematics |
| Horvat (2014) | not given | not given | No | Passive | Fluid intelligence |
| Jaeggi, Buschkuehl, Jonides, & Shah (2011) - M1 | 8.98 | 5.00 | No | Active | Fluid intelligence |
| Jaeggi, Buschkuehl, Jonides, & Shah (2011) - M2 | 8.98 | 5.00 | No | Active | Fluid intelligence |
| Karbach, Strobach, & Schubert (2015) - M1 | 8.30 | 9.33 | Yes | Active | Literacy/WD |
| Karbach, Strobach, & Schubert (2015) - M2 | 8.30 | 9.33 | Yes | Active | Mathematics |
| Karbach, Strobach, & | 8.30 | 9.33 | Yes | Active | Cognitive control |

Schubert (2015) - M3

| | | | | | |
|------------------------------------------------------|-------|-------|-----|---------|--------------------|
| Karbach, Strobach, & Schubert (2015) - M4 | 8.30 | 9.33 | Yes | Active | Cognitive control |
| Kroensbergen, Noordende, & Kolkman (2014) - M1 | 5.87 | 4.00 | Yes | Passive | Cognitive control |
| Kroensbergen, Noordende, & Kolkman (2014) - M2 | 5.87 | 4.00 | Yes | Passive | Mathematics |
| Kuhn & Holling (2014) - S1 | 9.00 | 5.00 | Yes | Passive | Mathematics |
| Kuhn & Holling (2014) - S2 | 9.00 | 5.00 | Yes | Active | Mathematics |
| Kun (2007) - S1 - M1 | 12.84 | 8.00 | Yes | Active | Fluid intelligence |
| Kun (2007) - S1 - M2 | 12.84 | 8.00 | Yes | Active | Science |
| Kun (2007) - S2 - M2 | 13.52 | 14.58 | Yes | Active | Science |
| Lee (2014) - M1 | 9.00 | 3.00 | Yes | Active | Literacy/WD |
| Lee (2014) - M2 | 9.00 | 3.00 | Yes | Active | Literacy/WD |
| Lindsay (2012) - M1 | 5.49 | 3.00 | Yes | Active | Literacy/WD |
| Lindsay (2012) - M2 | 5.49 | 3.00 | Yes | Active | Literacy/WD |
| Loosli, Buschkuehl, | 9.50 | 2.00 | No | Passive | Fluid intelligence |

Perrig, & Jaeggi

(2012) - M1

Loosli, Buschkuehl,

Perrig, & Jaeggi

(2012) - M2

9.50

2.00

No

Passive

Literacy/WD

Mansur-Alves &

Flores-Mendoza

(2015) - M1

11.17

13.33

Yes

Passive

Fluid intelligence

Mansur-Alves &

Flores-Mendoza

(2015) - M2

11.17

13.33

Yes

Passive

Fluid intelligence

Mansur-Alves, Flores-

Mendoza, & Tierra-

Criollo (2013) - M1

9.19

10.00

Yes

Active

Fluid intelligence

Mansur-Alves, Flores-

Mendoza, & Tierra-

Criollo (2013) - M2

9.19

10.00

Yes

Active

Fluid intelligence

Mansur-Alves, Flores-

Mendoza, & Tierra-

Criollo (2013) - M3

9.19

10.00

Yes

Active

Crystallized
intelligence

Mansur-Alves, Flores-

9.19

10.00

Yes

Active

Literacy/WD

| | | | | | | |
|------------------------------------------------------------|-------|-------|-----|---------|--------------------|--|
| Mendoza, & Tierra-Criollo (2013) - M4 | | | | | | |
| Mansur-Alves, Flores-Mendoza, & Tierra-Criollo (2013) - M5 | 9.19 | 10.00 | Yes | Active | Mathematics | |
| Mansur-Alves, Flores-Mendoza, & Tierra-Criollo (2013) - M6 | 9.19 | 10.00 | Yes | Active | Literacy/WD | |
| Nevo & Breznitz (2014) - M1 | 8.50 | 4.80 | Yes | Active | Literacy/WD | |
| Nevo & Breznitz (2014) - M2 | 8.50 | 4.80 | Yes | Active | Literacy/WD | |
| Passolunghi & Costa (2016) - S1 | 5.44 | 10.00 | Yes | Active | Mathematics | |
| Passolunghi & Costa (2016) - S2 | 5.42 | 10.00 | Yes | Passive | Mathematics | |
| Pugin et al. (2015) - M1 | 13.00 | 8.05 | No | Passive | Fluid intelligence | |
| Pugin et al. (2015) - M2 | 13.00 | 8.05 | No | Passive | Cognitive control | |
| Pugin et al. (2015) - | 13.00 | 8.05 | No | Passive | Cognitive control | |

M3

| | | | | | |
|----------------------------------------------------------|-------|-----------|-----|---------|--------------------|
| Pugin et al. (2015) - M4 | 13.00 | 8.05 | No | Passive | Cognitive control |
| Rode, Robson, Purviance, Geary, & Mayr (2014) - M1 | 9.00 | 7.14 | Yes | Passive | Mathematics |
| Rode, Robson, Purviance, Geary, & Mayr (2014) - M2 | 9.00 | 7.14 | Yes | Passive | Mathematics |
| Rode, Robson, Purviance, Geary, & Mayr (2014) - M3 | 9.00 | 7.14 | Yes | Passive | Literacy/WD |
| Rode, Robson, Purviance, Geary, & Mayr (2014) - M4 | 9.00 | 7.14 | Yes | Passive | Literacy/WD |
| Shavelson et al. (2008) | 13.50 | 14.58 | Yes | Active | Fluid intelligence |
| Söderqvist & Bergman-Nutley (2015) - M1 | 9.85 | not given | No | Passive | Literacy/WD |
| Söderqvist & | 9.85 | not given | No | Passive | Mathematics |

Bergman-Nutley
(2015) - M2

| | | | | | |
|--------------------------------------------------------------|------|------|-----|---------|--------------|
| St Clair-Thompson, Stevens, Huth, & Bolder (2010) - M1 | 6.83 | 6.00 | No | Passive | Literacy/WD |
| St Clair-Thompson, Stevens, Huth, & Bolder (2010) - M2 | 6.83 | 6.00 | No | Passive | Mathematics |
| St Clair-Thompson, Stevens, Huth, & Bolder (2010) - M3 | 6.83 | 6.00 | No | Passive | Mathematics |
| St Clair-Thompson, Stevens, Huth, & Bolder (2010) - M4 | 6.83 | 6.00 | No | Passive | Mathematics |
| Studer-Luethi, Bauer, & Perrig (2016) - S1- M1 | 8.25 | 4.50 | Yes | Active | Literacy/WD |
| Studer-Luethi, Bauer, & Perrig (2016) - S1- M2 | 8.25 | 4.50 | Yes | Active | Mathematics |
| Studer-Luethi, Bauer, | 8.25 | 4.50 | Yes | Active | Crystallized |

| | | | | | | |
|------------------------------------------------------|------|------|-----|---------|--|------------------------------|
| & Perrig (2016) - S1- M3 | | | | | | intelligence |
| Studer-Luethi, Bauer, & Perrig (2016) - S1- M4 | 8.25 | 4.50 | Yes | Active | | Fluid intelligence |
| Studer-Luethi, Bauer, & Perrig (2016) - S1- M5 | 8.25 | 4.50 | Yes | Active | | Cognitive control |
| Studer-Luethi, Bauer, & Perrig (2016) - S2- M1 | 8.25 | 4.50 | Yes | Passive | | Literacy/WD |
| Studer-Luethi, Bauer, & Perrig (2016) - S2- M2 | 8.25 | 4.50 | Yes | Passive | | Mathematics |
| Studer-Luethi, Bauer, & Perrig (2016) - S2- M3 | 8.25 | 4.50 | Yes | Passive | | Crystallized intelligence |
| Studer-Luethi, Bauer, & Perrig (2016) - S2- M4 | 8.25 | 4.50 | Yes | Passive | | Fluid intelligence |
| Studer-Luethi, Bauer, | 8.25 | 4.50 | Yes | Passive | | Cognitive control |

& Perrig (2016) - S2-
M5

| | | | | | |
|----------------------------------------------------------------------------|------|------|----|---------|--------------------|
| Thorell, Lindqvist, Bergman, Bohlin, & Klingberg (2008) - S1 - M1 | 4.67 | 6.25 | No | Active | Cognitive control |
| Thorell, Lindqvist, Bergman, Bohlin, & Klingberg (2008) - S1 - M2 | 4.67 | 6.25 | No | Active | Cognitive control |
| Thorell, Lindqvist, Bergman, Bohlin, & Klingberg (2008) - S1 - M3 | 4.67 | 6.25 | No | Active | Fluid intelligence |
| Thorell, Lindqvist, Bergman, Bohlin, & Klingberg (2008) - S1 - M4 | 4.67 | 6.25 | No | Active | Cognitive control |
| Thorell, Lindqvist, Bergman, Bohlin, & Klingberg (2008) - S2 | 4.67 | 6.25 | No | Passive | Cognitive control |

- M1

| | | | | | | |
|--------------------------------------------------------------------|-------|------|-----|---------|--------------------|--|
| Thorell, Lindqvist, Bergman, Bohlin, & Klingberg (2008) - S2 | | | | | | |
| - M2 | 4.67 | 6.25 | No | Passive | Cognitive control | |
| Thorell, Lindqvist, Bergman, Bohlin, & Klingberg (2008) - S2 | | | | | | |
| - M3 | 4.67 | 6.25 | No | Passive | Fluid intelligence | |
| Thorell, Lindqvist, Bergman, Bohlin, & Klingberg (2008) - S2 | | | | | | |
| - M4 | 4.67 | 6.25 | No | Passive | Cognitive control | |
| Wang, Zhou, & Shah (2014) - S1 | 10.50 | 6.67 | Yes | Active | Fluid intelligence | |
| Wang, Zhou, & Shah (2014) - S2 | 10.50 | 6.67 | Yes | Active | Fluid intelligence | |
| Wang, Zhou, & Shah (2014) - S3 | 10.50 | 6.67 | Yes | Active | Fluid intelligence | |
| Wang, Zhou, & Shah (2014) - S4 | 10.50 | 6.67 | Yes | Active | Fluid intelligence | |

| | | | | | |
|--------------------------------|------|-----------|-----|---------|--------------------|
| Witt (2011) | 9.68 | 7.50 | No | Passive | Mathematics |
| Zhao, Wang, Liu, & Zhou (2011) | 9.76 | not given | Yes | Passive | Fluid intelligence |

Table 3

Studies and moderators of the 6 near-transfer follow-up effect sizes included in the meta-analysis

| Study | Age | Duration of training | Random allocation | Type of control group |
|--------------------------------------------|------------|-----------------------------|--------------------------|------------------------------|
| Henry, Messer, & Nash (2014) | 7.00 | 3.00 | Yes | Active |
| Karbach, Strobach, & Schubert (2015) | 8.30 | 9.33 | Yes | Active |
| Pugin et al. (2015) - M1 | 13.00 | 8.05 | No | Passive |
| Pugin et al. (2015) - M2 | 13.00 | 8.05 | No | Passive |
| Studer-Luethi, Bauer, & Perrig (2016) - S1 | 8.25 | 4.50 | Yes | Active |
| Studer-Luethi, Bauer, & Perrig (2016) - S2 | 8.25 | 4.50 | Yes | Passive |

Table 4

Studies and moderators of the 24 near-transfer follow-up effect sizes included in the meta-analysis

| Study | Age | Duration of training | Random allocation | Type of control group | Domain |
|-------------------------------------------------|------------|-----------------------------|--------------------------|------------------------------|--------------------|
| Henry, Messer, & Nash (2014) - M1 | 7.00 | 3.00 | Yes | Active | Literacy/WD |
| Henry, Messer, & Nash (2014) - M2 | 7.00 | 3.00 | Yes | Active | Mathematics |
| Jaeggi, Buschkuehl, Jonides, & Shah (2011) - M1 | 8.98 | 5.00 | No | Active | Fluid intelligence |
| Jaeggi, Buschkuehl, Jonides, & Shah (2011) - M2 | 8.98 | 5.00 | No | Active | Fluid intelligence |
| Karbach, Strobach, & Schubert (2015) - M1 | 8.30 | 9.33 | Yes | Active | Literacy/WD |
| Karbach, Strobach, & Schubert (2015) - M2 | 8.30 | 9.33 | Yes | Active | Mathematics |
| Karbach, Strobach, & Schubert (2015) - M3 | 8.30 | 9.33 | Yes | Active | Cognitive control |
| Karbach, Strobach, & Schubert (2015) - M4 | 8.30 | 9.33 | Yes | Active | Cognitive control |

| | | | | | | |
|------------------------------------------------|-------|-----------|-----|---------|--------------------|--|
| Schubert (2015) - M4 | | | | | | |
| Pugin et al. (2015) - | | | | | | |
| M1 | 13.00 | 10.00 | No | Passive | Fluid intelligence | |
| Pugin et al. (2015) - | | | | | | |
| M2 | 13.00 | 10.00 | No | Passive | Cognitive control | |
| Pugin et al. (2015) - | | | | | | |
| M3 | 13.00 | 8.05 | No | Passive | Cognitive control | |
| Pugin et al. (2015) - | | | | | | |
| M4 | 13.00 | 8.05 | No | Passive | Cognitive control | |
| Söderqvist & Bergman-Nutley (2015) - M1 | | | | | | |
| | 9.85 | not given | No | Passive | Literacy/WD | |
| Söderqvist & Bergman-Nutley (2015) - M2 | | | | | | |
| | 9.85 | not given | No | Passive | Mathematics | |
| Studer-Luethi, Bauer, & Perrig (2016) - S1- | | | | | | |
| M1 | 8.25 | 4.50 | Yes | Active | Literacy/WD | |
| Studer-Luethi, Bauer, & Perrig (2016) - S1- | | | | | | |
| M2 | 8.25 | 4.50 | Yes | Active | Mathematics | |

| | | | | | |
|------------------------------------------------------|------|------|-----|---------|------------------------------|
| Studer-Luethi, Bauer, & Perrig (2016) - S1- M3 | 8.25 | 4.50 | Yes | Active | Crystallized intelligence |
| Studer-Luethi, Bauer, & Perrig (2016) - S1- M4 | 8.25 | 4.50 | Yes | Active | Fluid intelligence |
| Studer-Luethi, Bauer, & Perrig (2016) - S1- M5 | 8.25 | 4.50 | Yes | Active | Cognitive control |
| Studer-Luethi, Bauer, & Perrig (2016) - S2- M1 | 8.25 | 4.50 | Yes | Passive | Literacy/WD |
| Studer-Luethi, Bauer, & Perrig (2016) - S2- M2 | 8.25 | 4.50 | Yes | Passive | Mathematics |
| Studer-Luethi, Bauer, & Perrig (2016) - S2- M3 | 8.25 | 4.50 | Yes | Passive | Crystallized intelligence |
| Studer-Luethi, Bauer, & Perrig (2016) - S2- M4 | 8.25 | 4.50 | Yes | Passive | Fluid intelligence |

Studer-Luethi, Bauer,
& Perrig (2016) - S2-

M5

8.25

4.50

Yes

Passive

Cognitive control

2.4 Effect Size

The standardized means difference (Cohen's d) was calculated with the following formula:

$$d = (M_{g-e} - M_{g-c}) / SD_{pooled-pre} \quad (1)$$

where $SD_{pooled-pre}$ is the pooled standard deviation of the two pre-test standard deviations, and M_{g-e} and M_{g-c} are the gain of the experimental group and the control group, respectively (Schmidt & Hunter, 2015).⁵ The follow-up effect sizes were calculated by using the standardized difference between the follow-up and the pre-test measures.

Finally, the Comprehensive Meta-Analysis (Version 3.0; Biostat, Englewood, NJ) software package was used for correcting the effect sizes for upward bias (Hedges's g ; Hedges & Olkin, 1985), computing the overall effect sizes (\bar{g} s), and conducting statistical analyses.

2.5 Statistical Dependence of the Samples

The effect sizes were calculated for each relevant measure reported in the studies (Schmidt & Hunter, 2015). When several subscales of a test were used to measure the same construct (e.g. block recall and digit recall as measures of working memory), the measures were averaged, following Schmidt and Hunter's (2015) recommendation. Also, when the study presented a comparison between the treatment group and two control groups (passive and active), two effect sizes – one for each comparison with experimental and control groups – were calculated. As this procedure violates the principle of statistical independence of the samples, Cheung and Chan's (2004) method was applied to all the meta-analytic models. This method reduces the weight of dependent samples in the analysis by estimating an

⁵ When only the t -statistics were available, the t -values were converted into Cohen's d s (Lee, 2014; Witt, 2011).

adjusted (i.e., smaller) N (for a list of the adjusted N s, see Tables 2.1 to 2.13 in the Supplemental material available online; <http://psycnet.apa.org/record/2017-05288-001>). Since the method of Cheung and Chan (2004) cannot be used for partially dependent samples,⁶ we ran our analyses as if the comparisons between experimental samples and two different control groups were statistically independent. As shown by Bijmolt and Pieters (2001) and Tracz, Elmore, and Pohlmann (1992), the violation of statistical independence has little or no effect on means, standard deviations, and confidence intervals. Thus, the entire procedure is a reliable way to deal with the statistical dependence of part of the samples.

3. Results

3.1 Near-Transfer Effects

The random-effects meta-analytic overall effect size was $\bar{g} = 0.46$, 95% CI [0.35; 0.57], $k = 30$, $p < .001$. The forest plot is shown in Figure 2. The degree of heterogeneity between effect sizes was close to zero, $I^2 = 7.94$.⁷

⁶ In addition, in three studies, a few participants did not take part in all the tests (i.e., attrition). In these cases, we used the mean number of participants as the number to be adjusted.

⁷ The I^2 statistic refers to the percentage of between-study variance due to true heterogeneity and not to random error (Higgins, Thompson, Deeks, & Altman, 2003).

Study name

Hedges's g and 95% CI

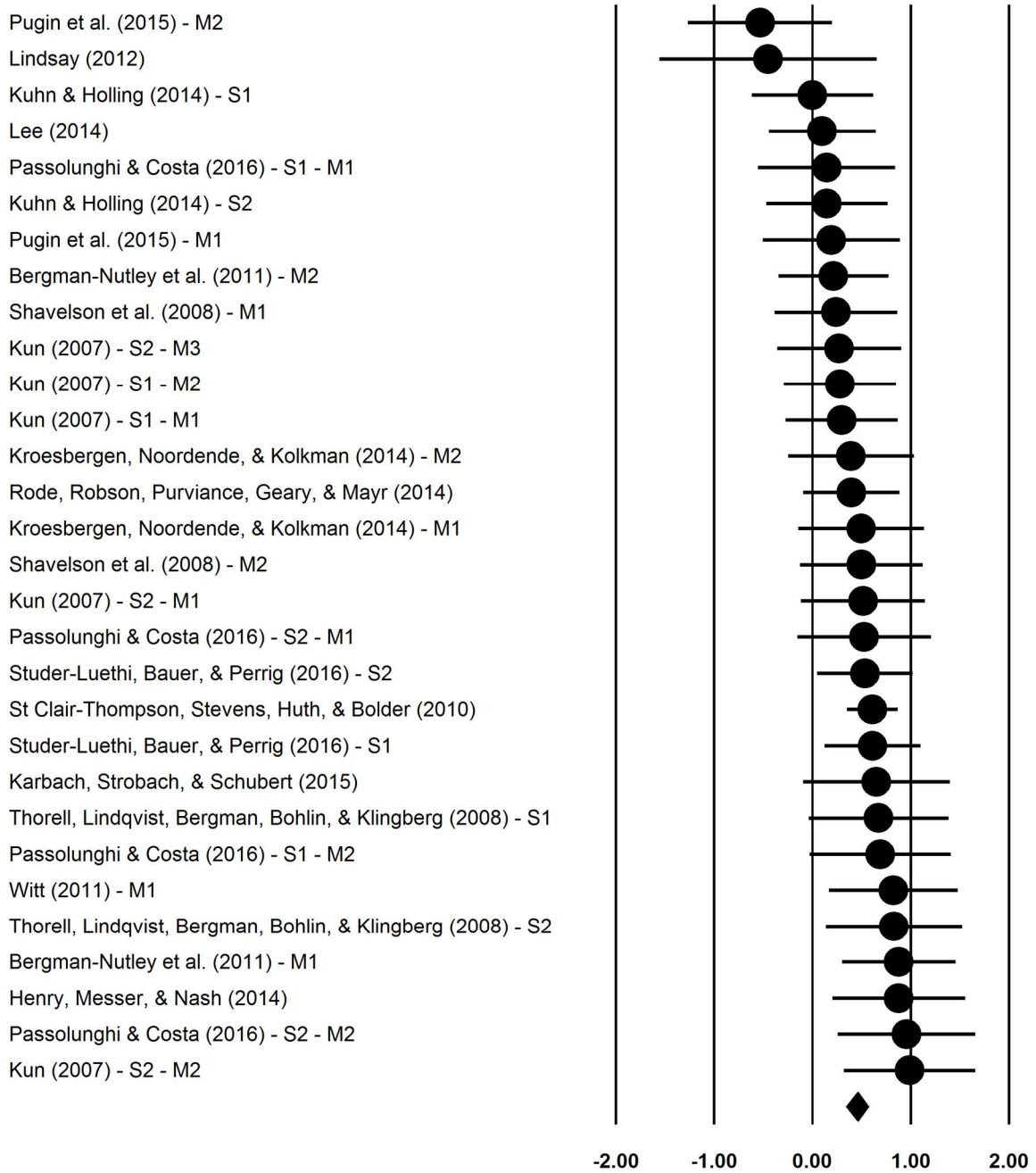


Figure 2. Forest plot of the near-transfer model. Hedges's g s (circles) and 95% CIs (lines) are shown for all the effects entered into the meta-analysis. The diamond at the bottom indicates the meta-analytically weighted mean \bar{g} . When studies had multiple samples, the table reports the result of each sample (S1, S2, etc.) separately. Similarly, when studies used multiple outcome measures, the table reports the result of each measure (M1, M2, etc.) separately.

3.1.1 Moderator analyses

Age was marginally significant, $Z(1) = -1.80$, $b = -0.03$, $p = .072$. None of the other three moderators were significant: Random allocation, $Z(1) = -0.58$, $b = -0.08$, $p = .562$; Type of control group, $Z(1) = -0.31$, $b = -0.04$, $p = .760$; and Duration of training, $Z(1) = 0.42$, $b = 0.01$, $p = .678$.

3.1.2 Publication bias analysis

To test whether our analysis was affected by publication bias, we examined a funnel plot representing the relation between effect sizes and standard errors. The contour-enhanced funnel plot (Peters, Sutton, Jones, Abrams, & Rushton, 2008) is shown in Figure 3.

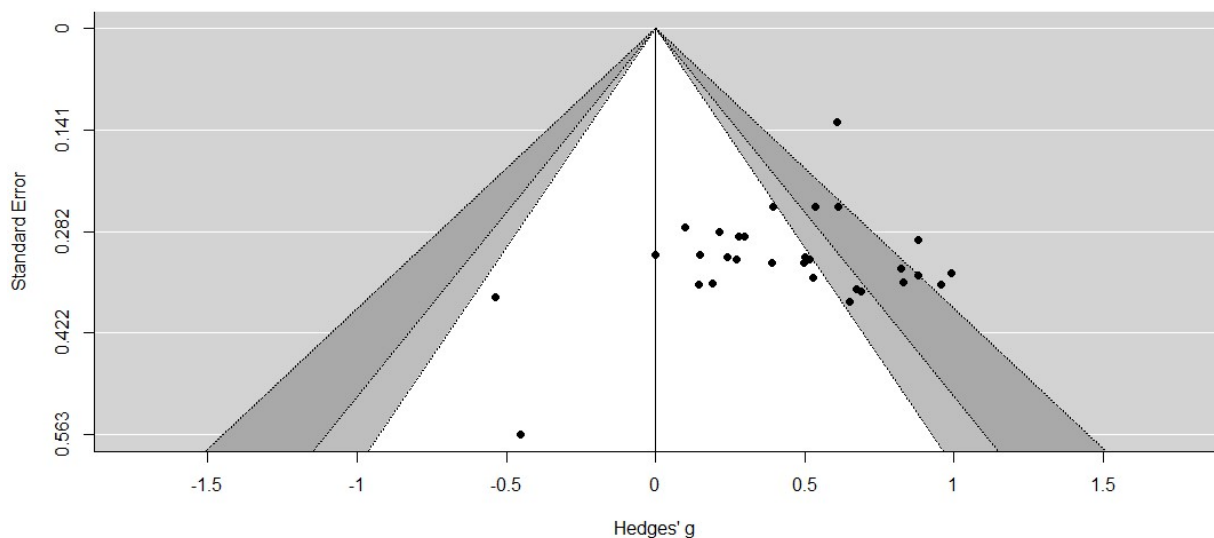
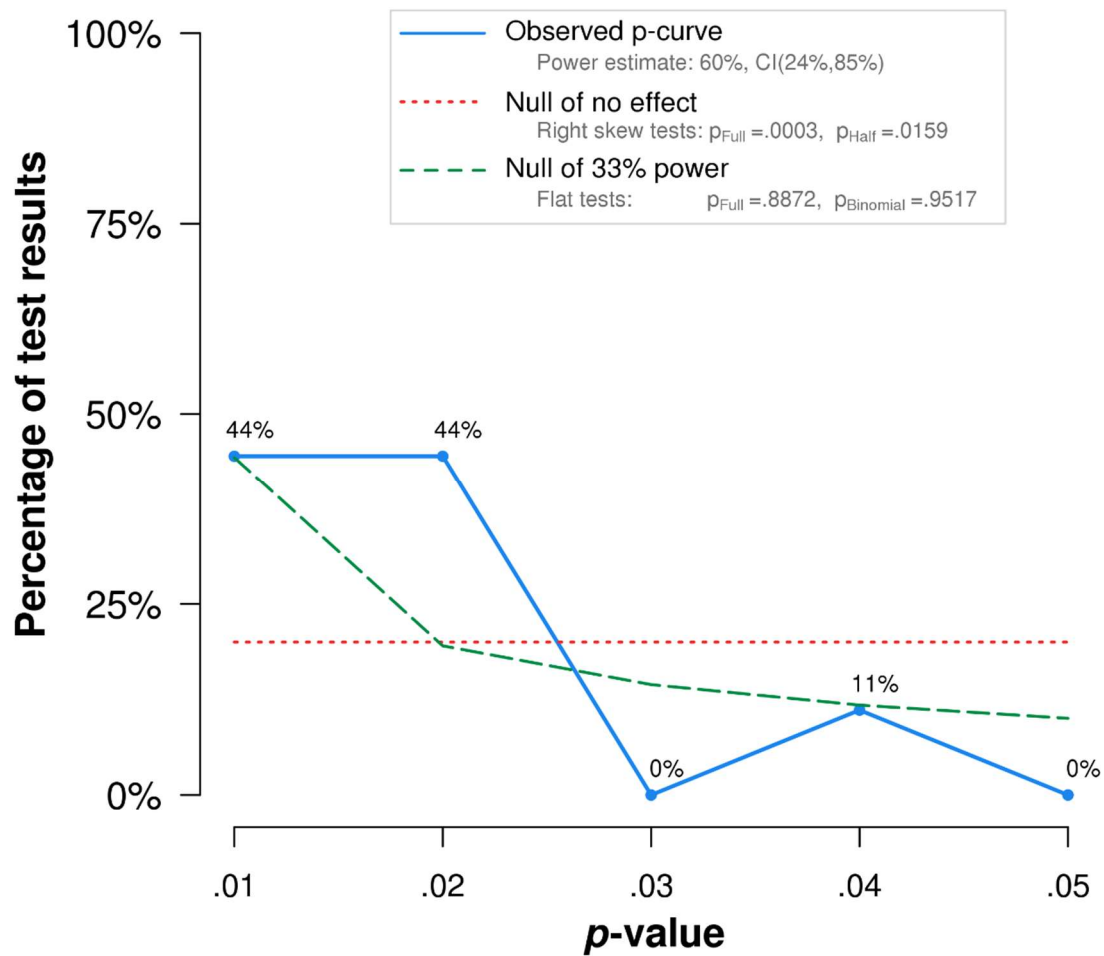


Figure 3. Contour-enhanced funnel plot of standard errors and effect sizes (Hedges's g s) in the near-transfer meta-analysis. The black circles represent the effect sizes included in the meta-analysis. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

The symmetry of the funnel plot around the meta-analytic mean was tested by Egger's regression test (Egger, Smith, Schneider, & Minder, 1997). The test showed no evidence of publication bias ($p = .217$). In addition, the trim-and-fill analysis (Duval & Tweedie, 2000) estimated no weaker-than-average missing study (left of the mean). Finally, a p -curve analysis was run with all the p -values $< .05$ related to positive effect sizes (Simonsohn, Nelson, & Simmons, 2014). The results showed evidential values (i.e., no evidence of

publication bias), $Z(9) = -3.39, p = .003$ (Figure 4).



Note: The observed p-curve includes 9 statistically significant ($p < .05$) results, of which 8 are $p < .025$. There were no non-significant results entered.

Figure 4. p-curve analysis. The blue (continuous) line shows that most of the significant p-values are smaller than .025, suggesting evidential value.

3.2 Far-Transfer Effects

The random-effects meta-analytic overall effect size was $\bar{g} = 0.12$, 95% CI [0.06; 0.18], $k = 74, p < .001$. The forest plot is shown in Figure 5. The degree of heterogeneity between effect sizes was $I^2 = 0.00$.

Study name

Hedges's g and 95% CI

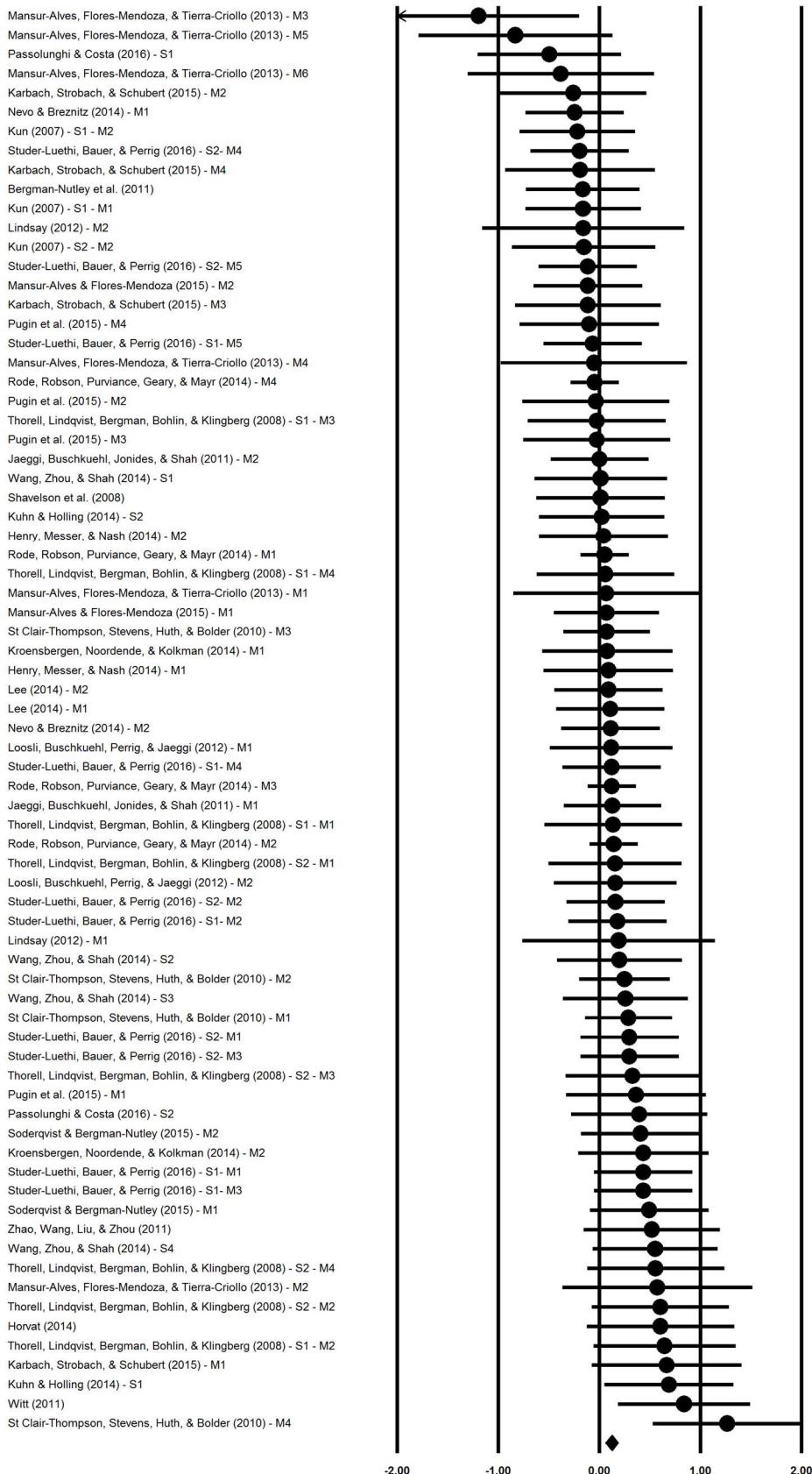


Figure 5. Forest plot of the far-transfer model. Hedges's g s (circles) and 95% CIs (lines) are shown for all the effects entered into the meta-analysis. The diamond at the bottom indicates the meta-analytically weighted mean \bar{g} . When studies had multiple samples, the table reports the result of each sample (S1, S2, etc.) separately. Similarly, when studies used multiple outcome measures, the table reports the result of each measure (M1, M2, etc.) separately.

3.2.1 Moderators analysis

Random Allocation was a significant moderator, $Z(1) = -2.76$, $b = -0.20$, $p = .006$.

The overall effect sizes in randomized and non-randomized samples were $\bar{g} = 0.07$, 95% CI [0.00; 0.14], $k = 50$, $p = .046$, and $\bar{g} = 0.27$, 95% CI [0.15; 0.39], $k = 24$, $p < .001$, respectively. Type of control group was marginally significant, $Z(1) = -1.83$, $b = -0.12$, $p = .067$. The overall effect sizes when WM training was compared to active and passive control groups were $\bar{g} = 0.05$, 95% CI [-0.05; 0.15], $k = 40$, $p = .311$, and $\bar{g} = 0.18$, 95% CI [0.09; 0.26], $k = 34$, $p < .001$, respectively. Also, the overall effect size in randomized samples with active control groups was $\bar{g} = 0.03$, CI [-0.07; 0.14], $k = 34$, $p = .521$. Finally, Duration of training was marginally significant, $Z(1) = -1.81$, $b = -0.02$, $p = .070$. No other moderator was significant: Age, $Z(1) = -1.60$, $b = -0.03$, $p = .110$; and Domain, $p = .703$.

3.2.2 Additional meta-analytic models

We calculated the random-effects meta-analytic overall effect sizes of each of the six domains. The only significant overall effect size was $\bar{g} = 0.20$, 95% CI [0.03; 0.36], $k = 17$, $p = .018$, for mathematics. To test the robustness of the result, we ran two moderator analyses for this domain. Random Allocation was a significant moderator, $Z(1) = -2.01$, $b = -0.35$, $p = .045$. The overall effect sizes in randomized and non-randomized samples were $\bar{g} = 0.10$, 95% CI [-0.05; 0.25], $k = 12$, $p = .193$, and $\bar{g} = 0.49$, 95% CI [0.11; 0.88], $k = 5$, $p = .012$, respectively. Type of control group was significant, $Z(1) = -2.41$, $b = -0.43$, $p = .016$. The overall effect sizes when WM training was compared to active and passive control groups

were $\bar{g} = -0.11$, 95% CI $[-0.38; 0.16]$, $k = 6$, $p = .426$, and $\bar{g} = 0.31$, 95% CI $[0.13; 0.49]$, $k = 11$, $p = .001$, respectively.

Literacy/WD overall effect size was marginally significant, $\bar{g} = 0.11$, 95% CI $[-0.00; 0.22]$, $k = 17$, $p = .055$. None of the other overall effect sizes was significant: $\bar{g} = 0.11$, 95% CI $[-0.02; 0.24]$, $k = 21$, $p = .101$ for fluid intelligence; $\bar{g} = 0.09$, 95% CI $[-0.08; 0.26]$, $k = 14$, $p = .302$ for cognitive control; $\bar{g} = -0.02$, 95% CI $[-0.75; 0.71]$, $k = 3$, $p = .956$ for crystallized intelligence; and $\bar{g} = -0.20$, 95% CI $[-0.65; 0.25]$, $k = 2$, $p = .386$ for science.

3.2.3 Publication bias analysis

The contour-enhanced funnel plot of the main model ($k = 74$) is shown in Figure 6.

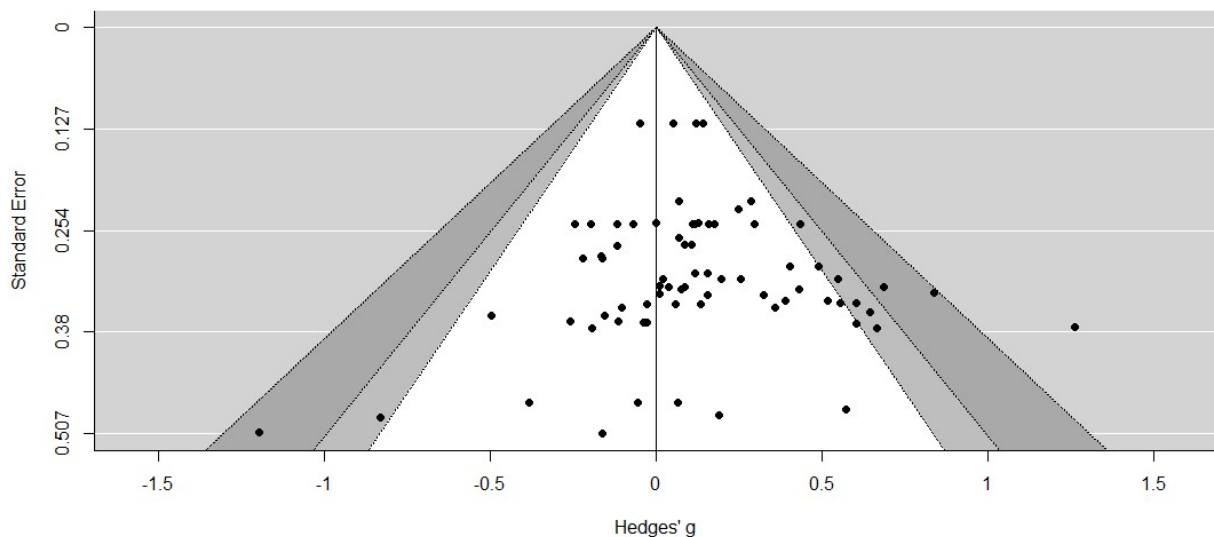
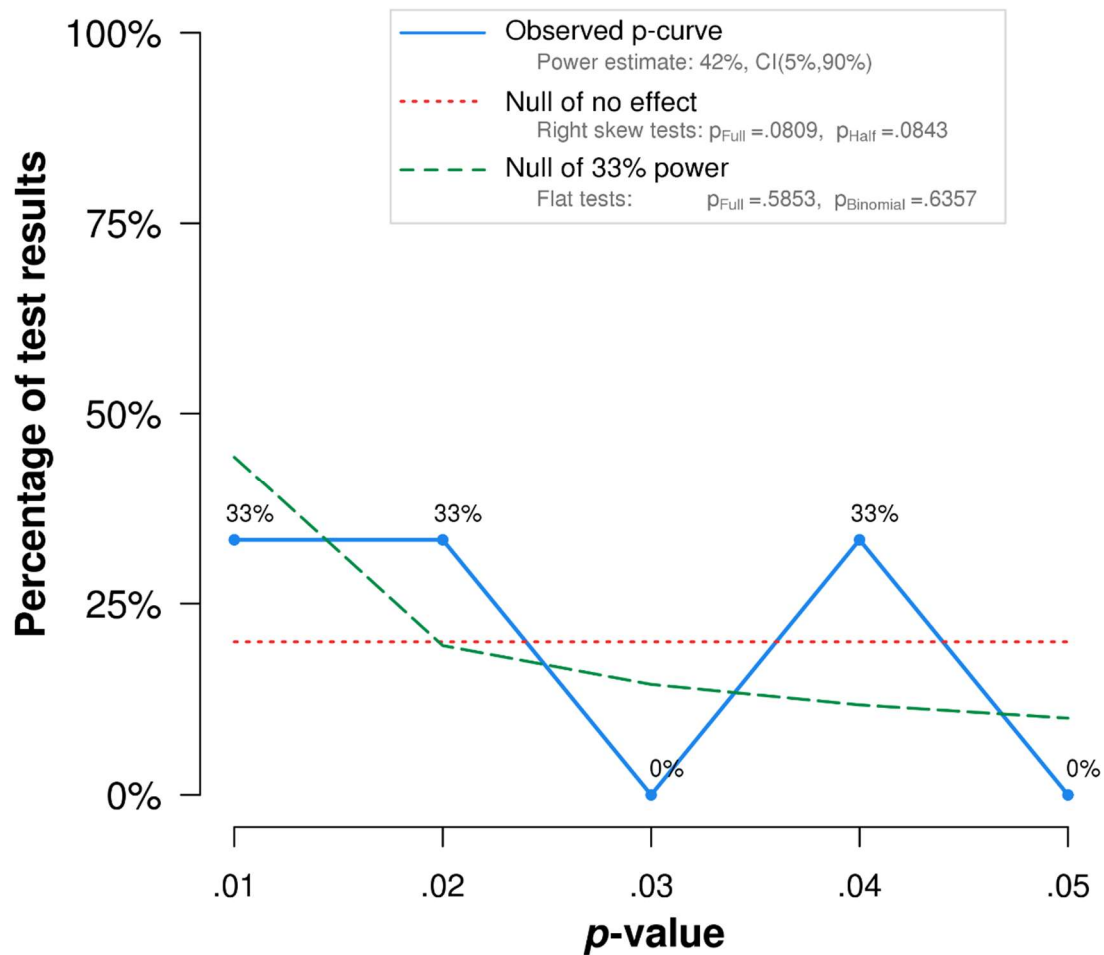


Figure 6. Contour-enhanced funnel plot of standard errors and effect sizes (g s) in the far-transfer meta-analysis. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

Egger's regression test showed no evidence of publication bias ($p = .511$). In addition, the trim-and-fill analysis estimated no weaker-than-average missing studies (left of the mean). Finally, we performed a p -curve analysis. Both the full and half p -curve tests were right

skewed with $p < .100$ ($Z(3) = -1.40, p = .081$ and $Z(3) = -1.38, p = .084$, respectively) suggesting evidential value (Simonsohn, Simmons, & Nelson, 2015; Figure 7).⁸



Note: The observed p-curve includes 3 statistically significant ($p < .05$) results, of which 2 are $p < .025$. There were no non-significant results entered.

Figure 7. *p*-curve analysis. The blue (continuous) line shows that most of the significant *p*-values are smaller than .025, suggesting evidential value.

⁸ Since only three values were inputted, the results of this *p*-curve analysis might be unreliable. However, it must be kept in mind that the occurrence of publication bias is quite unlikely when the overall effect size is close to zero.

A trim-and-fill analysis was performed for four additional meta-analytic models, (fluid intelligence, cognitive control, mathematics, and literacy/WD models). In the fluid intelligence model, five studies were filled in, and the point estimate was $\bar{g} = 0.03$, 95% CI [-0.09; 0.15]. In the literacy/word decoding model, two studies were filled in, and the point estimate was $\bar{g} = 0.08$, 95% CI [-0.03; 0.19]. No missing study was found in the other two models. Due to the scarcity of effect sizes, no publication bias analysis was run for the science and crystallized intelligence models.

3.3 Follow-Up Effects

For near-transfer follow-up effects, the random-effects meta-analytic overall effect size was $\bar{g} = 0.33$, 95% CI [0.00; 0.65], $k = 6$, $p = .049$. The degree of heterogeneity between effect sizes was $I^2 = 40.50$.

For far-transfer follow-up effects, the random-effects meta-analytic overall effect size was $\bar{g} = 0.09$, 95% CI [-0.02; 0.20], $k = 24$, $p = .122$. The degree of heterogeneity between effect sizes was $I^2 = 0.00$.

3.3.1 Moderator analyses

Due to the small number of effect sizes, no moderator analysis was run for the near-transfer effects model. (For the same reason, no publication bias analysis was carried out for this model.) Regarding the far-transfer effects model, no moderator was significant.

3.3.2 Publication bias analysis

In the far-transfer effect model, Egger's regression test showed no evidence of publication bias ($p = .345$). In addition, the trim-and-fill analysis estimated no weaker-than-average missing studies (left of the mean). No p -curve analysis was carried out because none of the effect sizes in the model reached statistical significance.

4. Discussion

The purpose of this meta-analysis was to evaluate the impact of WM training on TD children's cognitive and academic skills. The results showed a clear pattern. Similar to previous meta-analyses (e.g., Melby-Lervåg & Hulme, 2013; Schwaighofer et al., 2015), WM training significantly affected WM-related skills (post-test overall effect size, $\bar{g} = 0.46, p < .001$) and remained several months after the end of training (follow-up overall effect size, $\bar{g} = 0.33, p = .049$). However, we found little or no evidence that WM training enhances fluid intelligence or domain-general processes such as cognitive control. The same applied to academic abilities such as literacy or science. Only the mathematics-related overall effect size was significant, albeit quite modest ($\bar{g} = 0.20, p = .018$). However, methodological issues cast some doubts on the authenticity of the effect (we will take up this point below). Thus, the results of the meta-analysis do not support the hypothesis according to which WM training benefits cognitive or academic abilities in TD children.

Interestingly, WM training seems to produce approximately the same negligible effects on measures outside the domain of WM regardless of the age of participants and domain. The significant (or marginally significant) moderators in the far-transfer main model ($k = 74$) were the random allocation of the participants to the samples, the type of control group, and duration of training. The overall effect size was much smaller in randomized samples ($\bar{g} = 0.07, p = .046$) than in non-randomized samples ($\bar{g} = 0.27, p < .001$). This outcome suggests that episodes of self-selection in the experimental groups or differences at baseline level between experimental and control groups may have inflated the effect sizes in samples with no random allocation.⁹ Analogously, the overall effect size was smaller when

⁹ In the present case, the difference between groups at baseline level in some of the dependent variables seems to be the most likely explanation. In several studies (e.g., Thorell,

the experimental group was compared to an active control group ($\bar{g} = 0.05, p = .311$) than a passive control group ($\bar{g} = 0.18, p < .001$). This finding corroborates the idea that the positive effect sizes reported in some primary studies are due to placebos as well. Moreover, when only the effect sizes in randomized samples with active control groups were considered, the overall effect size was almost null ($\bar{g} = 0.03, p = .521$). Finally, the duration of training seems to be slightly inversely related to the size of the effects ($b = -0.02$). This result is difficult to interpret. However, the null degree of heterogeneity suggests caution in interpreting these outcomes. In fact, the moderator analyses may have detected effects due to random error rather than true heterogeneity between-effect sizes (see footnote 7). In any case, far transfer effects of WM training appear to be negligible or, at best, modest.

4.1 Theoretical and Practical Implications

The present meta-analysis reviewed the studies in which participants were TD children. For this reason, the results we reported do not apply to other populations – such as children with learning disabilities or adults. Nonetheless, the fact that, in the general population of children, WM training induces improvements in WM-related outcomes but not in other types of cognitive and academic measures suggests some theoretical and practical implications.

To begin with, if far-transfer is more likely to occur in children than adults when cognitive and academic skills are developing, then our findings cast serious doubts on the

Lindqvist, Bergman, Bohlin, & Klingberg, 2009), the control groups performed better than the experimental groups at the pre-test. The difference between the groups decreased at the post-test, suggesting that the positive effect size is probably due to some statistical artefact (e.g., regression to the mean, ceiling effect).

idea that training a domain-general mechanism such as WM improves fluid intelligence, cognitive control, or academic achievement.¹⁰ Second, and linked to the first point, the lack of an effect of WM training on fluid intelligence supports the idea that WM and fluid intelligence are two different constructs (Ackerman, Beier, & Boyle, 2005; Hornung, Brunner, Reuter, & Martin, 2011; Kane, Hambrick, & Conway, 2005).

However, it must be noted that the positive effects in near-transfer measures might reflect an improvement in WM tasks performance, rather than a genuine enhancement in WM capacity (Shipstead et al., 2012). In other words, participants learn how to do the task without improving their WM capacity. If this is the case, nothing can be inferred about the relationship between fluid intelligence (or any other far-transfer measure) and WM capacity. Moreover, following this line of reasoning, the absence of fluid intelligence enhancement could be interpreted as a failed improvement in WM capacity after the training (see also the discussion in Melby-Lervåg & Hulme, 2013). Regrettably, the information provided in the primary studies is not sufficient to solve the issue.

The fact that the participants showed improvements in a large variety of tasks different from the WM trained tasks (see Table S1.1 in the Supplemental Material available online; <http://psycnet.apa.org/record/2017-05288-001>) might suggest that WM capacity was actually boosted. However, pervasive improvement in WM-related measures may stem from amelioration in some general skill at performing WM tasks rather than an increased WM capacity. Thus, testing whether WM training enhances WM capacity requires not only a set

¹⁰ It must be noted that this argument does not apply to the population of older adults. In fact, the aim of WM training in the elderly is to slow down cognitive decline, not to extend developing cognitive abilities. For a review, see Karbach and Verhaeghen (2014).

of multivariate measures of WM capacity, but also that task-related improvements occur through a common factor that is measurement invariant across treatment and control groups (i.e., training effects that are proportional to the factor loadings in a structural equation model). If such conditions can be met in a well-powered single study, then it can be convincingly claimed that WM capacity has been enhanced.

Beyond these theoretical aspects, the most obvious practical implication of our results is that WM training, at the moment, cannot be recommended as an educational tool. WM training seems to have little or no effect on far-transfer measures of cognitive abilities and academic achievement. More generally, this meta-analysis provides further evidence that the occurrence of far-transfer is too infrequent to offer solid educational advantages. For this reason, cognitive and academic enhancement interventions should be as close as possible to the skills that are meant to be trained.

4.2 Limitations of the Present Meta-Analysis

Near-transfer effects seem to remain even a few months after the end of the training. However, the limited number of studies ($n = 4$) and effect sizes ($k = 6$) does not allow us to draw any reliable conclusion about this. The same limitation applies, to a lesser degree, to the far-transfer follow-up effects ($n = 6$, $k = 24$). In this case, however, the findings are consistent with the immediate post-test outcomes: modest or null effects in both the measures. In fact, it is hard to see why negligible effects immediately after training, such as those reported in this meta-analysis, should become significantly larger several months after the end of training.

Finally, other potential moderators – such as the type of training program – were not considered in the meta-analytic models because of the limited number of the effect sizes. However, the small degree of heterogeneity in both the near- and far-transfer models discourages us from thinking that other moderators could have affected the overall results.

4.3 Conclusions

The findings of the present meta-analysis do not invite optimism about the effectiveness of WM training at improving cognitive skills and academic achievement in TD children. WM training seems to enhance children's performance in WM- and STM-related measures. However, with regard to skills outside the domain of WM such as fluid intelligence, cognitive control, mathematics, and literacy, this training seems to have little or no effect. Consistent with Thorndike and Woodworth's (1901) common element theory, our findings show that the occurrence of far-transfer is, at best, sporadic.

Chapter 5: Meta-Analysis of Chess Training

Rationale for the Meta-Analysis in Chapter 5

Chapter 5 reports a meta-analysis on the effect of chess instruction on children's cognitive abilities and academic achievement. Chess is unanimously considered a "brain game," and chess skill is often associated with superior intellectual ability. Also, playing chess requires a considerable amount of cognitive effort. In fact, the game requires memorising thousands of positions, calculating hundreds of variants, and precise planning. Thus, chess is a perfect domain to test the hypothesis according to which engaging in intellectually demanding activities enhances domain-general cognitive function.

The studies included in this meta-analysis are listed in Appendix B.

1. Introduction

Recently, many concerns have been expressed about pupils' poor mathematics achievement both in the United States (Hanushek, Peterson, & Woessmann, 2012; Richland, Stigler, & Holyoak, 2012) and in Europe (Grek, 2009). Pupils' low mathematical skills have serious consequences well beyond the classroom, as the possibility of successfully majoring in Science, Technology, Engineering, and Mathematics (STEM) subjects, and hence obtaining STEM jobs, is limited by one's mathematical skills. The job market demands more graduates in STEM subjects than graduates in the humanities and has also become more competitive worldwide in recent years, with increasingly high mathematical competences being required (Halpern et al., 2007).

To address the issue of how to improve mathematics instruction, policy makers and researchers have explored a number of avenues. One such avenue is to teach chess in schools. Chess has recently started to become part of the school curriculum (as an optional subject) in several countries. Chess-related research and educational projects are currently ongoing in the United Kingdom, Spain, Turkey, Germany, and Italy, among other countries. Commenting on a large project having introduced chess in the curriculum of 175 schools in the UK, chess master Jerry Myers stated that chess "directly contributes to academic performance. Chess makes children smarter" (Garner, 2012). The European Parliament has expressed its favourable opinion on using chess courses in schools as educational tool (Binev, Attard-Montalto, Deva, Mauro, & Takkula, 2011) and, similarly, the Spanish Parliament has approved the implementation of chess courses during school hours. These initiatives have been conducted because chess is considered an effective educational tool able to improve not only mathematical skills, but also other academic skills such as reading and general cognitive abilities such as concentration and intelligence, and even children's heuristics and habits of

mind (Costa & Kallick, 2009). Critically, efforts to promote chess in schools take for granted that chess skill transfers to other domains.

1.1 Difficulty of Transfer

Transfer of learning occurs when a set of skills acquired in one domain generalizes to other domains or improves general cognitive abilities. Transfer is an important question both theoretically and practically. Mestre (2005) distinguishes between *near-transfer*, where transfer occurs between closely related domains (e.g., transfer from geometry to calculus) and *far-transfer*, where the source and target domains are only loosely related (e.g., transfer from Latin to geography). It has been proposed that transfer is a function of the extent to which two domains share common features (Thorndike & Woodworth, 1901) and cognitive elements (Anderson, 1990). In line with this hypothesis, near-transfer is often observed, although exceptions also exist. For example, research into expertise shows that transfer is only partial between subspecialties of expertise such as cardiology and neurology (Rikers, Schmidt, & Boshuizen, 2002). By contrast, substantial research in education and psychology suggests that far-transfer is difficult (Donovan, Bransford, & Pellegrino, 1999). This includes the research on teaching the computer language LOGO in order to improve children's thinking skills, which has obtained disappointing results (De Corte & Verschaffel, 1986; Gurtner, Gex, Gobet, Nunez, & Restchitzki, 1990). In addition, the higher the level of a skill, the more specific the features of a domain will be, and the lower the likelihood that there will be transfer (Ericsson & Charness, 1994), in particular because a large number of domain-specific perceptual chunks will be acquired (Gobet, 2016). Again, there are exceptions, and some individuals have excelled in several different domains (Gobet, 2011; 2016).

The difficulty of transferring knowledge and skills raises a number of significant practical issues, especially in education. Most educational interventions try to transmit knowledge which, to some extent, is meant to be transferable from one domain of learning to

another. In fact, transferability of skills is either a tacit assumption or a specific aim of nearly every educational program (Donovan et al., 1999; Perkins & Salomon, 1994). Therefore, educational institutions are interested in methodologies implementing school activities that teach and boost transferable skills. One approach is to teach general strategies, such as learning, problem-solving, and reasoning heuristics (Perkins & Grotzer, 2000), so that these skills can be easily transferred to other domains. Another approach is to teach a specific activity, with the hope that this activity will help individuals to develop skills that might be useable in other domains. The game of chess is one such activity that has been used in that way.

1.2 The issue of Transfer in Chess Research

A substantial amount of research has been devoted to understanding the cognitive processes underpinning chess skill, and much is known about chess players' perception, learning, memory, and problem solving (for reviews, see Gobet, 2016, and Gobet, de Voogt, & Retschitzki, 2004). Much less is known about the extent to which chess skill transfers to other domains of learning.

Several studies (Bilalić, McLeod, & Gobet, 2007; Doll & Mayr, 1987; Frydman & Lynn, 1992; Grabner, Stern, & Neubauer, 2007) have shown that chess players tend to be more intelligent than the general population. However, these studies were correlational in nature and cannot establish that chess skill is the actual cause of better intellectual abilities. In fact, the exact opposite causal explanation could be true: some individuals could excel at chess due to their superior intellectual abilities (Gobet & Campitelli, 2002).

Assuming that skills acquired in chess will lead to benefits in domains such as mathematics and reading clearly implies the presence of far transfer. In line with Thorndike and Woodworth's (1901) hypothesis, several studies have shown that chess players' skill

tends to be context-bound, suggesting that it is difficult to achieve far-transfer from chess to other domains. For example, memory for chess positions fails to transfer from chess to digits both in adults and children (Chi, 1978; Schneider, Gruber, Gold, & Opwis, 1993); chess players' perceptual skills do not transfer to visual memory of shapes (Waters, Gobet, & Leyden, 2002); chess skill does not predict performance in the economic game known as beauty contest (Bühren & Frank, 2010); and finally, chess planning skills do not help chess players to solve the Tower of London task (Unterrainer, Kaller, Leonhart, & Rahm, 2011).

1.3 Chess in School

In spite of these negative results, several researchers have pursued the hypothesis that skills acquired with chess can transfer to other domains. Two main explanations have been adduced to support this hypothesis. First, chess requires decision-making skills and high-level processes (such as acquiring and selecting relevant information from a problem) similar to those used in mathematics and reading (Margulies, 1992). Second, since chess is a cognitively demanding task involving focused attention and problem solving, playing chess should strengthen these cognitive abilities and thus be beneficial for children's school performance (Bart, 2014). However, convincing experimental evidence of the effectiveness of chess instruction is lacking. In a literature review, Gobet and Campitelli (2006) argued that there was no strong evidence for the cognitive and academic benefits of chess. They found only few studies, which included unpublished reports or master and doctoral theses. Most importantly, many of these studies had a quasi-experimental design (no random assignment to the experimental and control groups) and, in some cases, the experimental samples were self-selected.

The difficulty of transferring chess skill is consistent with the literature on the transfer of specific skills. At first blush, it is hard to see why knowing the strategic value of the bishop pair or the correct way to handle a minority attack should offer any advantage in

mathematics, understanding a text, or developing focused attention. Nevertheless, it is possible that chess practice enhances some abilities shared with other domains, such as those mentioned above, provided that chess is taught early on with children, when academic and cognitive abilities are at the beginning of their development. This is the reason why chess intervention studies have focused on the academic and cognitive skills of children rather than adults: Children's skills are less context-specific than adults', and thus transfer of learning is more likely in the former than in the latter.

Some recent studies (Sala, Gorini, & Pravettoni, 2015; Scholz et al., 2008; Trincherro, 2012; Trincherro & Sala, 2016) have provided more refined explanations as to why chess may effectively enhance cognitive and mathematical skills. According to these researchers, chess improves children's mathematical skills because the game has some elements in common with the mathematical domain and because it promotes suitable habits of mind (Costa & Kallick, 2009). Through chess, children train several context-independent skills (such as the ability to understand the existence of a problem or the need for correct reasoning), which may transfer to the mathematical domain. This is possible because (primary school) mathematics and chess share some common features (e.g., numerical and spatial relationships as well as quantity-based problems), strategies to solve problems (e.g., focusing and interpreting game/problem situations, selecting relevant information, or looking for correct arguments), cognitive skills (e.g., attention) and meta-cognitive skills (e.g., planning). The aim of our study is thus to test, comprehensively and quantitatively, these previous claims on the putative benefits of chess instruction in school.

2. Scope, Aims, and Hypotheses of the present Meta-Analysis

Given the considerable attention that research on chess in school is attracting and the potentially important implications for our understanding of transfer, it is important to provide a scientific evaluation of the effects of chess instruction on academic and cognitive skills. A

similar interest has been devoted to studies on the possibility that video-games improve cognitive skills and that the benefits transfer to other domains (Green, Li, & Bavelier, 2010; Green, Pouget, & Bavelier, 2010). Just like with the video-game literature, a possibility that will have to be kept in mind in our meta-analysis is that the observed transfer from the source domains to the target domains might be due to confounds such as the placebo effect (Boot, Blakely, & Simons, 2011; Gobet et al., 2014).

Our meta-analysis¹¹ is an investigation of studies regarding the potential benefits of chess for children with respect to (a) mathematics skills, (b) reading skills, and (c) several cognitive categories (general intelligence, meta-cognition, attention/concentration, and spatial abilities). We chose these three categories of skills because they were the three categories chess-related research has been focusing on.

Our study had two main aims. The first aim was to estimate the overall effect size of the benefits of chess instruction by comparing experimental groups to control groups. The second aim was to evaluate the potential role of several factors in moderating the effect of chess instruction in children. The first four moderators addressed substantive aspects of the studies, and the last two covered methodological aspects:

1. Outcome: Mathematics, reading, or cognitive skills;

¹¹ Two previous meta-analyses were carried out on the effect of chess instruction: Benson (2006) and Nicotera and Stuit (2014). Neither calculated an overall effect size nor ran a moderator analysis. Rather, they divided the meta-analytic means into sub-categories (such as mathematics with chess instruction). The results they obtained were optimistic compared to ours, as they included several studies that were not included in the present meta-analysis because they did not satisfy the selection criteria.

2. Duration of training (in hours);
3. Grade of the participants: Primary or secondary school;
4. Participants' category: Children with special educational needs or not;
5. Publication: Published or unpublished studies, where “published” is defined as having appeared in a peer-reviewed journal;
6. Design quality: Integer index (range 0 – 3, from poor to good) expressing the quality of the study design. The index measures three methodological characteristics: random allocation, administered pre-test, and avoidance of self-selection of the sample.

Along with the evaluation of the potential role of the above moderators, two specific sets of hypotheses were tested. The first pair of hypotheses dealt with the general question as to whether the skills acquired with chess instruction transfer to other domains. Two opposing hypotheses were tested. Hypothesis 1a predicted that, consistent with the literature on expertise and most of the literature on transfer, chess skill does not transfer to other domains, or at best the transfer is small and mostly due to unspecific factors (such as placebo effects). Hypothesis 1b, which reflects the view held by most researchers and practitioners in the field of chess instruction, predicted that there is substantial transfer. The second hypothesis dealt with the benefits of chess instruction on mathematics and reading. In line with Thorndike and Woodworth (1901), it was predicted that transfer is stronger with mathematics than with reading, as chess shares more elements with the former topic than with the latter.

3. Method

3.1 Literature Search

A systematic search strategy was used to find the relevant studies. The procedure is summarized in Figure 8. Google Scholar, ProQuest (Dissertations & Theses), ERIC and Psyc-Info databases were searched to identify all the potential relevant studies. In addition,

previous narrative reviews were examined, and we e-mailed researchers in the field asking for unpublished studies and data.

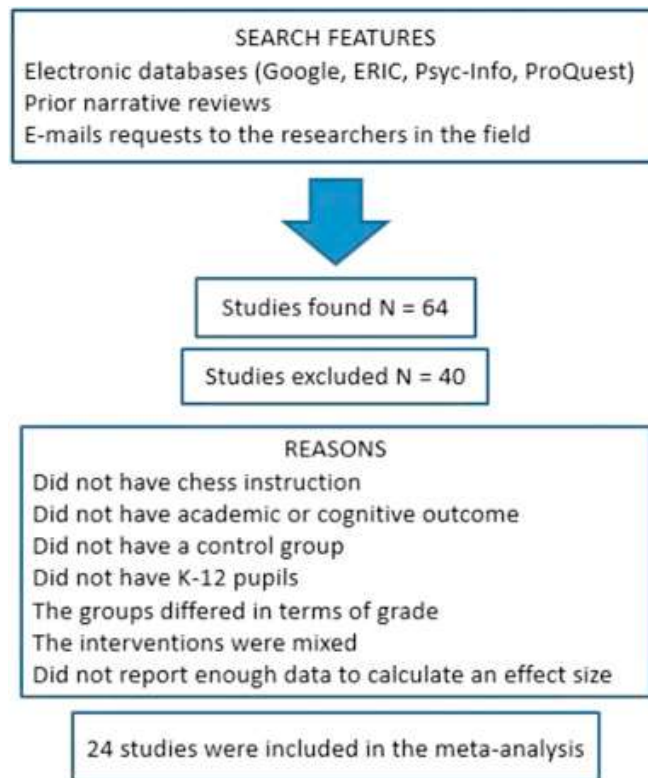


Figure 8. Flow diagram of the studies considered and ultimately included in the meta-analysis.

3.2 Inclusion/Exclusion Criteria

The studies were included according to the following seven criteria:

1. The design of the study was experimental or quasi-experimental; correlational and ex post facto studies were excluded.
2. The independent variable (chess instruction) was successfully isolated; the studies using chess instruction as one of several independent variables (such as other

activities) in the experimental group were excluded.

3. The study presented a comparison between a chess intervention group and at least one control group.
4. The treatment and the control groups did not differ in terms of grade (e.g. third graders compared to fourth or fifth graders).
5. During the study, a measure of mathematical, reading, or cognitive skill was collected.
6. The participants of the study were pupils from kindergarten to the 12th grade.
7. The data presented in the published study were sufficient to calculate an effect size or the author(s) of the study, after having been contacted, provided the necessary data.

We found 24 studies, conducted from July 1976 to July 2015, that met all the inclusion criteria (see Table 5). These studies included 25 independent samples and 40 effect sizes, and a total of 5,221 participants (2,788 in the experimental groups and 2,433 in the control groups).

Table 5

Summary of the 24 studies included in the meta-analysis

| STUDY | OUTCOME | PUBLISHED | HOURS | DESIGN QUALITY | SPECIAL EDUCATIONAL NEEDS | GRADE | OUTCOME MEASURE |
|-----------------------------------------------|-------------------|-----------|-------|-------------------|---------------------------------|-----------|-----------------|
| Aciego, Garcia, & Betancort (2012) | Cognitive | Yes | 96 | 1 | No | Both | WISC-R |
| Aydin (2015) | Maths & Cognitive | Yes | 48 | 1 | Yes | Secondary | Unknown |
| Barrett & Fish (2011) | Maths & Cognitive | Yes | 25 | 2 | Yes | Secondary | TAKS |
| Christiaen & Verhofstadt- Denève (1981) | Maths & Reading | Yes | 42 | 2 | No | Primary | DGB |
| DuCette (2009) | Maths & Reading | No | Not | 0 | No | Both | PSSA |

| Given | | | | | | | |
|-----------------------------------------------------|-------------------|-----|-----------|---|-----|-----------|--------------------------------------------------------------------|
| Eberhard (2003) | Cognitive | No | 60 | 1 | Yes | Secondary | CogAT; NNAT |
| Forrest, Davidson, Stucksmith, & Glendinning (2005) | Maths & Reading | No | 37 | 2 | No | Primary | WISC-R (arithmetic subtest); Neale test |
| Fried & Ginsburg (n.d.) | Cognitive | No | Not Given | 2 | Yes | Primary | WISC-R |
| Garcia (2008) | Maths & Reading | No | 90 | 1 | No | Primary | TAKS |
| Gliga & Flesner (2014) | Cognitive | Yes | 10 | 3 | No | Primary | Krapelin test; Rey test |
| Hong & Bart (2007) | Cognitive | Yes | 20 | 3 | Yes | Both | RPM |
| Kazemi, Yektayar, & Abad (2012) | Maths & Cognitive | Yes | 96 | 2 | No | Both | TIMSS (mathematical literacy); Panaoura, Philippou & Christou test |

| | | | | | | | |
|-------------------------------------------|-------------------|-----|-----------|---|-----|-----------|----------------------------------------------------------|
| Kramer & Filipp (n.d.) | Cognitive | No | 32 | 2 | No | Primary | Unknown |
| Margulies (1992) | Reading | No | Not Given | 1 | No | Primary | DRP |
| Rifner (1992) | Maths & Reading | No | 30 | 2 | No | Secondary | CTBS/4 |
| Romano (2011) | Maths | No | 25 | 3 | No | Primary | INVALSI |
| Sala & Trincherro (in preparation) | Maths & Cognitive | No | 10 | 3 | No | Primary | OCDE-Pisa (mathematical literacy) |
| Sala, Gorini, & Pravettoni (2015) | Maths | Yes | 18 | 3 | No | Primary | OCDE-Pisa (mathematical literacy) |
| Sala, Gobet, Trincherro, & Ventura (2016) | Maths & Cognitive | No | 15 | 3 | No | Primary | TIMSS (mathematical literacy); Panaoura & Philippou test |
| Scholz et al. (2008) | Maths & Cognitive | Yes | 24 | 3 | Yes | Primary | Arithmetic test designed by the authors; DL-KG |

| | | | | | | | |
|--------------------------------|-----------------|-----|----|---|----|---------|-----------------------------------|
| Sigirtmac (2012) | Cognitive | Yes | 50 | 0 | No | Primary | Unknown |
| Trincherro & Piscopo (2007) | Maths | No | 30 | 2 | No | Primary | Unknown |
| Trincherro & Sala (2016) | Maths | No | 19 | 3 | No | Primary | OCDE-Pisa (mathematical literacy) |
| Yap (2006) | Maths & Reading | No | 50 | 0 | No | Primary | Oregon State Assessment |

3.3 Effect Size¹²

For the studies with an *only-post-test design*, the standardized means difference (Cohen's d) was calculated with the following formula:

$$d = (M_e - M_c) / SD_{pooled} \quad (1)$$

where SD_{pooled} is the pooled standard deviation and M_e and M_c are the means of the experimental group and the control group, respectively.¹³ For the studies with a *repeated-measure design*, the standardized means difference was calculated with the following formula:

$$d = (M_{g-e} - M_{g-c}) / SD_{pooled-pre} \quad (2)$$

where $SD_{pooled-pre}$ is the pooled standard deviation of the two pre-test standard deviations, and M_{g-e} and M_{g-c} are the gain of the experimental group and of the control group, respectively. For the studies with an *ANCOVA design*, the standardized means difference was calculated with the following formula:

$$d = (M_{adj-e} - M_{adj-c}) / SD_{pooled} \quad (3)$$

where SD_{pooled} is the pooled standard deviation of the two standard deviations of the unadjusted means, and $M_{adj-e} - M_{adj-c}$ are the adjusted means of the experimental group and the control group, respectively. To correct for the upward bias, every Cohen's d was converted into Hedges's g by using the following formula:

¹² All the formulas we used were taken from Schmidt and Hunter (2015).

¹³ If the t or F statistics were provided, we used the regular formulas $d = t \times$

$\sqrt{(N_e + N_c) / (N_e \times N_c)}$ and $d = \sqrt{F \times (N_e + N_c) / (N_e \times N_c)}$.

$$g = d / (1 + 0.75 / (N - 3)) \quad (4)^{14}$$

where N is the sample size of the study.

Where reliability coefficients were available, the effect sizes were corrected for measurement error by using the following formula:

$$g' = g / a \quad (5)$$

where a is the square root of the reliability coefficient. It was possible to apply this correction to 31 effect sizes. Finally, there were three outliers whose residual errors had z scores greater than 4. These were Winsorized to z scores equal to 3.99.

Since we believed that the effect sizes had to reflect the actual improvement of the experimental groups and should not be the product of statistical artefacts, we adopted the following criterion: when the control group performance decreased in the post test, the effect size was calculated by considering M_{g-c} (control group gain) equal to 0. Finally, the Comprehensive Meta Analysis (Version 3.0; Biostat, Englewood, NJ) software package was used for computing the effect sizes and conducting the statistical analyses.

4. Results

A random model ($k = 40$) was built to calculate an overall effect size¹⁵. The overall effect size was $\bar{g} = 0.338$, 95% CI [0.242; 0.435], $p < .001$. The degree of heterogeneity

¹⁴ This formula is an acceptable approximation of the one converting Cohen's d s into Hedges's g s presented in Chapter 3.

¹⁵ Twelve studies had more than one effect size. However, according to Tracz, Elmore, and Pohlmann (1992), violations of statistical independence have little or no effect on means, standard deviations, and confidence intervals.

between effect sizes was between moderate and high ($I^2 = 57.227$), suggesting the potential effect of some moderators. A trim-and-fill analysis showed that there was no publication bias. Consistent with this, a funnel plot analysis, depicting the relationship between standard error and effect size, was approximately symmetrical (see Figure 9).

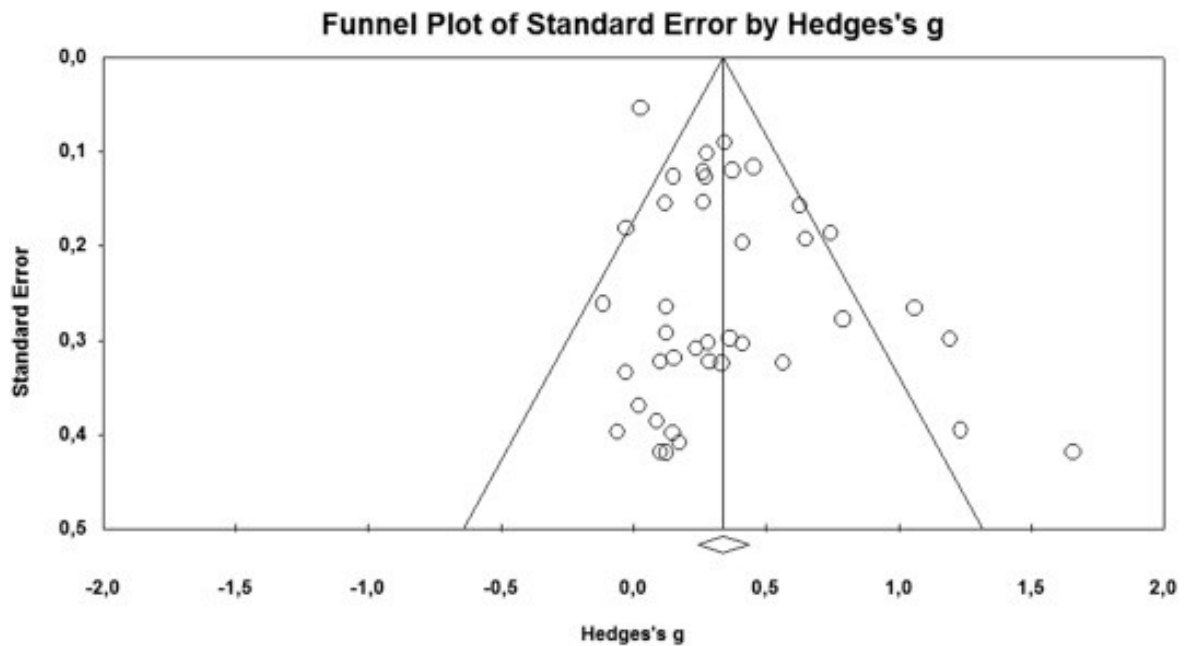


Figure 9. Funnel plot of standard errors and effect sizes (g). The diamond at bottom represents the meta-analytically weighted mean Hedges's \bar{g} .

4.1 Moderator Analyses

The only two statistically significant moderators were Duration of Training, which positively affected the effect sizes ($Q(1) = 3.89$, $b = 0.0038$, $p < .05$, two tailed, $k = 35$), and Publication, which also positively affected the effect size ($Q(1) = 10.17$, $b = 0.2941$, $p < .01$, two tailed, $k = 40$).

Following Trinchero's (2012) suggestion (see Discussion), we considered 25 hours as a threshold for the moderator Duration of Training. The overall effect size in studies with 25 or more hours of treatment was $\bar{g} = 0.427$, 95% CI [0.271; 0.583], $p < .001$, $k = 23$, while the overall effect size in studies with less than 25 hours of training was $\bar{g} = 0.303$, 95% CI

[0.189; 0.417], $p < .001$, $k = 12$. Regarding the moderator Publication, the overall effect size of the published studies was $\bar{g} = 0.540$, 95% CI [0.346; 0.735], $p < .001$, $k = 17$, while the overall effect size of the unpublished studies was $\bar{g} = 0.230$, 95% CI [0.149; 0.311], $p < .001$, $k = 23$.

4.2 Additional Meta-Analytic Models

Although outcome was not a significant moderator, we ran three additional random models – one for each outcome – in order to investigate whether any outcome shows an overall effect size appreciably superior (or inferior, see discussion) to the others, as stated in Hypothesis 2.

The first model included the 17 mathematics-related effect sizes. The overall effect size was $\bar{g} = 0.382$, 95% CI [0.229; 0.535], $p < .001$. A trim-and-fill analysis showed that there was no publication bias. The second model included the 16 cognitive-related effect sizes. The overall effect size was $\bar{g} = 0.330$, 95% CI [0.130; 0.529], $p = .001$. A trim-and-fill analysis indicated that there was no publication bias. Finally, the third model included the seven reading-related effect sizes. The overall effect size was $\bar{g} = 0.248$, 95% CI [0.128; 0.368], $p < .001$. A trim-and-fill analysis showed a possible publication bias (one study trimmed, left to the mean). The analysis showed that the point estimate was $\bar{g} = 0.241$, 95% CI [0.122; 0.359].

5. Discussion

There is currently much research and excitement about the benefits of teaching chess in schools. The issue is theoretically important, since chess researchers' and practitioners' claims about the presence of far transfer are at variance with main theories of learning and expertise, which consider far transfer as difficult. In order to evaluate these diverging

predictions, the current meta-analysis examined the effect exerted by chess instruction on academic (mathematics and reading) and cognitive abilities in children.

5.1 Substantive Results

The first hypothesis predicted overall transfer beyond placebo effects. The results of the current meta-analysis suggest that chess instruction improves children's mathematical, reading, and cognitive skills moderately. Although this outcome seems promising, two considerations should be borne in mind. First, the overall effect size is not large enough to convincingly establish the effectiveness of chess instruction in enhancing the skills in consideration. By using Hattie's (2009) categorization, an overall effect size of $\bar{g} = 0.338$ is not in the so-called "zone of desired effects," that is $d \geq 0.4$, which is the median value of the effectiveness of educational interventions estimated by Hattie's second-order meta-analysis. This suggests that chess instruction is no more effective in enhancing children's cognitive and academic skills than many (at least more than 50%) other possible educational interventions. Moreover, the observed difference between treatment and control groups might be due to chess instructors' passion rather than chess itself, because the potential role of placebo effects was rarely, if ever, controlled for in the studies under consideration (we will take up this methodological point below). Thus, consistent with Thorndike and Woodworth's (1901) common-element theory, the results tend to lend more support to Hypothesis 1a (chess skill does not transfer to other domains) than Hypothesis 1b (transfer will be substantial), which is largely held by the field of chess-in-school research. These considerations – along with the overall results of the meta-analysis – lead us to think that learning activities should be as close as possible to the skills to train; for example, mathematics instruction should be used to teach mathematical skills.

However, the positive influence of the hours of treatment on the results seems to support the idea that chess skill does transfer to other domains. Trinchero (2012) has

suggested that appreciable positive effects occur only after 25 – 30 hours of chess instruction. For studies with a minimum of 25 hours of instruction, the overall \bar{g} effect size was 0.427, which is a value in the “zone of desired effects” (see above). It is thus unlikely that this positive outcome is only the consequence of placebo effects, although this possibility cannot be ruled out completely. This suggests that 25 – 30 hours of chess instruction is the minimum amount of instruction in order to obtain a significant transfer of learning from chess to other domains.

The second hypothesis, which was a more direct test of Thorndike and Woodworth’s (1901) theory, predicted that transfer from chess should be stronger to mathematics than to reading, as chess shares more common elements with the former than the latter. Consistent with the hypothesis, the overall effect size was larger with mathematics than with reading ($\bar{g} = 0.382$ vs. $\bar{g} = 0.248$). Although outcome was not a significant moderator, reading seemed to benefit less from chess instruction than mathematics, as the effect size was substantially lower; this was despite the fact that five of the seven studies on reading used a long duration (30 hours or more; no information about duration was available in the other two studies).

In the introduction, we presented Thorndike and Woodworth’s (1901) view that transfer of skills occurs only between two domains that share components. It is plausible to argue that chess and mathematics have some components in common, such as their problem-solving nature and the importance of quantitative relationships. Therefore, the hypothesis that chess is a medium (in the sense of Feuerstein, Feuerstein, Falik, & Rand, 2006) through which cognitive skills are trained with some benefit for mathematics is plausible, even though it has not yet been convincingly supported by empirical research. However, with respect to reading, it is difficult to identify what components are shared with chess, unless we focus on very general commonalities (e.g., chess playing and reading are both decision-making activities). In their study of the effects of chess instruction on reading, Forrest, Davidson,

Stucksmith, and Glendinning (2005) suggested that chess interventions enabled participants with low self-esteem to gain more confidence, which improved their literacy skills. If true, this suggestion – along with the small effect size ($\bar{g} = 0.248$) – upholds the idea that the effects of chess interventions on reading are non-specific.

5.2 Methodological Moderators

The index of design quality was not a significant moderator. This fact suggests that the results have not been significantly biased by the design used in the studies included in the meta-analysis. Nevertheless, as previously mentioned, the absence of an active control group in almost all the studies was a potential design-related confound we could not control for. The moderator Publication indicated that studies published in peer-reviewed journals have greater effect sizes. That studies with good results are more likely to be published is a common pattern in the literature (Schmidt & Hunter, 2015).

5.3 Limitations of this Study

Regrettably, like the vast majority of studies carried out to assess the effect of educational methods, none of the studies considered in this review employed what Gobet and Campitelli (2006) called the “ideal design.” This design includes the following requirements in addition to a treatment group: pre-test and post-test; two control groups (a do-nothing group and an active control group, necessary for removing the possibility of a placebo effect); random allocation to group; different personnel for conducting the pre-test, the treatment, and the post-test; and ideally – but nearly impossible to do in practice – experimenters’ and testers’ unawareness of the nature of the assignment into groups, and participants’ unawareness of the goal of the experiment and the fact that they take part in an experiment. The presence of an active control group is crucial for controlling the possibility of placebo effects, and thus establishing the causal role of chess instruction in far transfer. Mechanisms that could produce “placebo effects” include instructors’ motivation, the state of motivation

induced by a novel activity, and educators' expectations (e.g., Boot, Simons, Stothart, & Stutts, 2013; Gobet & Campitelli, 2006). Without any active control group, it is not possible to exclude the possibility that positive results are due to such confounds, rather than to chess itself. It remains unknown whether a study with a more rigorous design would yield the same results as the studies previously conducted. Since nearly no study in the current meta-analysis had an active control group, which is necessary for ruling out possible placebo effects, the effects of chess instruction could have been systematically overestimated.

Another limitation of this field of research is that too few studies reliably controlled for moderator effects. In addition, the dependent variables were often very different between the studies: for example, basic arithmetic skills and mathematical problem-solving skills are not the same thing, and the same applies to meta-cognition, general intelligence, attention, and spatial abilities. We classified the studies using three broad kinds of outcomes (mathematical, reading, and cognitive skills) because, unfortunately, the small number of studies did not allow us to reliably evaluate the specific skills assessed as potential moderators.

5.4 Conclusions and Recommendations for Future Research

Even if chess, under specific circumstances, seems to positively affect children's skills, there still are serious doubts about the real effectiveness of its practice. There is a need to clarify whether this positive influence is due to placebo effects or to chess instruction itself. In the latter case, research should identify the mechanisms underpinning the link between chess, the specific cognitive abilities involved and enhanced by the practice of the game, and their potential influence on mathematics and reading skills. In addition, the field should develop a detailed causal model explaining the cognitive processes that mediate learning and transfer. Finally, the data suggest that chess enhances children's mathematical skills and cognitive abilities more than reading skills, although the moderator analysis was

not statistically significant. With reading skills, both the data and the explanations provided by researchers suggest that the positive effects of chess on children's reading skills are due to placebo effects. Further research should establish the reliability of these results.

Regarding future studies, we recommend to use an experimental design (random allocation, pre-tests and post-tests) with two control groups (a do-nothing group and an active control group). While logistically more complex, such a design is necessary in order to establish whether the benefits putatively provided by chess instruction are genuine and not caused by non-specific factors (e.g., placebo effect). Another important goal is to identify the specific characteristics of chess that might improve children's abilities, and which abilities they foster (e.g., attention, spatial abilities, quantitative reasoning, or meta-cognition). For example, is it the diversity of pieces on the board that help maintain attention? Does the movement of the pieces help to boost visuo-spatial abilities? Does chess ideally combine numerical, spatial, temporal, and combinatorial aspects? Does chess promote a better and more conscious way of thinking? In particular, it is important to demonstrate whether these features are common or not to other activities and games. Specifically, one should understand whether some features (e.g., quantitative relationships between pieces and problem-solving situations) are shared by other board games.

Thus, researchers should include (at least) two dependent variables – one academic and one cognitive – in their experimental designs, in order to shed some light on the causal relationships between chess instruction, and cognitive and academic skills. Many researchers, for instance, have claimed that chess enhances mathematical skills because chess practice relies on cognitive skills and mechanisms that, in turn, underlie mathematical skills. While this hypothesis is plausible, too few studies have directly addressed the question by assessing both a cognitive and an academic outcome, and the results have been contradictory. For example, Scholz et al. (2008) and Sala and Trincherò (in preparation) found no effect of

chess on focused attention and meta-cognition respectively, whereas Kazemi, Yektayar, and Abad (2012) found a positive effect of chess practice on meta-cognitive abilities both in primary and in secondary school participants.

Finally, since the effectiveness of chess in enhancing children's intellectual skills seems to be dependent on the duration of the training, it would be useful to directly manipulate this variable in future studies, by systematically varying the duration of treatments between groups. This would ascertain the minimal and optimal amounts of chess instruction for far transfer: too short a duration might not provide enough time for progress, while too long a duration might lead to diminishing returns. Other worthwhile topics of investigation include a comparative study of different teaching methods with respect to their efficiency (e.g., is instruction better with computers or without computers? Are group activities preferable to individual activities, or is it the opposite? Are there more efficient orders of covering the material?). Finally, there has been little research that has explicitly mapped between chess and aspects of mathematics. Possible examples include bridging the chess board with the Cartesian graph and bridging the way the king moves in chess with block distance (as opposed to Euclidean distance). As it is known that awareness makes transfer more likely (Gick & Holyoak, 1980; Salomon & Perkins, 1989), it is plausible that making explicit the links between chess and mathematics could facilitate transfer.

In conclusion, the game of chess seems to exert a slight positive influence on both academic and cognitive abilities. Further research is needed to shed light on the relationship between cognitive and academic improvements, to evaluate the role of potential moderators and confounds, and to understand the role, if any, of placebo effects and game elements non-specific to chess.

Chapter 6: Meta-Analysis of Music Training

Rationale for the Meta-Analysis in Chapter 6

Chapter 6 reports a meta-analysis on the effect of music training on children and young adolescents' cognitive abilities and academic achievement. Like chess, music is considered a cognitively demanding activity, and correlational evidence links music skill with superior cognitive ability. Music training is thus another relevant domain in which to test the cognitive-training theory.

The studies included in this meta-analysis are listed in Appendix C.

1. Introduction

Recently, the question of whether music-related activities in school improve young people's cognitive and academic skills has raised much interest among researchers, educators, and policy makers. Several studies have tried to establish the effectiveness of music training in enhancing children's and young adolescents' general intelligence (Rickard, Bambrick, & Gill, 2012), memory (Roden, Kreutz, & Bongard, 2012), spatial ability and mathematics (Mehr, Schachner, Katz, & Spelke, 2013), and literacy skills (Slater et al., 2014), among others (for a review, see Miendlarzewska & Trost, 2013). Music training comprises activities such as singing songs, playing instruments, clapping, and rhythm games beyond many others. Notably, several specific curricula have been designed to develop those cognitive skills involved in playing music (e.g., Kodály method; Houlahan & Tacka, 2015). The educational implications of this research are evident. If music training enhances children's and young adolescents' cognitive skills and school grades, then schools might consider implementing additional musical activities.

1.1 The Question of Transfer of Skills

Crucially, the importance of establishing whether music training provides any educational advantage is not limited to the field of education. In fact, this topic addresses the broader psychological question of *transfer* of skills. Transfer of learning takes place when skills learned in one particular area either generalize to new areas or increase general cognitive abilities. It is customary to distinguish between *near-* and *far-transfer* (Barnett & Ceci, 2002; Mestre, 2005). Whilst near-transfer takes place between areas that are tightly related (e.g., driving two different car models), far-transfer occurs where the relationship between source and target areas is weak (e.g., transfer from music to mathematics). Thus, postulating that music skill generalizes to other non-music-related cognitive and academic abilities means assuming the occurrence of a far-transfer.

According to Thorndike and Woodworth's (1901) common-element theory, transfer depends on the number of features that are shared between two areas; these features are hypothesized to engage common cognitive elements (Anderson, 1990). A direct consequence of this theory, well supported by empirical data in psychology and education, is that, while near-transfer should be frequent, far-transfer should be rare (Donovan, Bransford, & Pellegrino, 1999; Sala & Gobet, 2016).

1.2 Why Should Music Skill Transfer to non-Music Skills?

Music training has been claimed to enhance various cognitive and academic skills. Given the well-known difficulty of far-transfer to occur, it is possible that music training boosts context-independent cognitive mechanisms, which may, in turn, improve other non-music cognitive and academic skills. According to Schellenberg (2004, 2006), the most likely explanation for the alleged diverse benefits of music interventions is that such training enhances individuals' general intelligence, which correlates with many cognitive and academic skills (Deary, Strand, Smith, & Fernandes, 2007; Rohde & Thompson, 2007). Music training requires focused attention, learning complex visual patterns, memory, and fine motor skills. Thus, such a demanding activity may enhance children's and young adolescents' overall cognitive skill, which, in turn, would increase their academic performance. This hypothesis is corroborated by the fact that formal exposure to music in childhood appears to correlate with IQ scores and academic attainment (Schellenberg, 2006).

Another possible explanation relies on executive functions. Cognitive skills such as working memory, cognitive control, and cognitive flexibility are important predictors of academic achievement (e.g., Conway & Engle, 1996; Peng, Namkung, Barnes, & Sun, 2016). Learning to play an instrument engages executive functions (Bialystok & Depape, 2009; George & Coch, 2011) and it is not impossible that such improvements generalize to non-music skills.

1.3 Does Music Skill Transfer to non-Music Skills? A Look at the Empirical Evidence

Several correlational studies have shown that music skill is associated with non-music-specific skills such as literacy (Anvari, Trainor, Woodside, & Levy, 2002; Forgeard et al., 2008), mathematics (Cheek & Smith, 1999), short-term and working memory (Lee, Lu, & Ko, 2007), and general intelligence (Lynn, Wilson, & Gault, 1989; Schellenberg, 2006; Schellenberg & Mankarious, 2012). Anvari et al. (2002) found that music perception skills predicted reading abilities in preschool children. Similarly, Forgeard et al. (2008) reported that music discrimination ability correlated with phonological processing skill in a sample of typically developing and dyslexic children. Concerning mathematical ability, Cheek and Smith (1999) showed that students who had received private lessons of music performed better in the mathematics portion of the Iowa Test of Basic Skills. Consistent with the latter two studies, Wetter, Koerner, and Schwaninger (2009) found a positive relationship between being engaged in music activities and overall academic achievement.

Music skill seems to be positively associated to cognitive ability too. In Lee et al.'s (2007) study, music-trained children and adults were compared to age-matched control groups in a series of digit span and spatial span tasks. The music-trained groups outperformed the controls in all the measures. Regarding overall cognitive ability, a convincing amount of evidence suggests that music skill and general intelligence are significantly related. Lynn et al. (1989) found a correlation between the scores on Raven's Standard Progressive Matrices (Raven, 1960) and a series of music tests in a group of 9-11-year-old children. Moreover, Schellenberg (2006) reported a positive correlation between duration of the music training and IQ in children and undergraduate students. Crucially, this result remained even after controlling for potentially confounding variables, such as parental income and education. Finally, this finding was confirmed in a more recent study involving 7- and 8-year-old children (Schellenberg & Mankarious, 2012).

However, such correlational studies cannot ascertain any far-transfer of skill from music training to other areas, because no direction of causality can be inferred. For example, both music and non-music skills could stem from innate intellectual abilities. Stronger conclusions can be drawn from studies using an experimental design, where an experimental group without previous formal musical instruction receives musical training. However, the experimental studies on the benefits of music training have provided mixed results. For example, while some studies have reported positive results (Kaviani, Mirbaha, Pournaseh, & Sagan, 2014; Portowitz, Lichtenstein, Egorova, & Brand, 2009), others have showed modest evidence of music training on children's performance on intelligence tests (Rickard, et al., 2012; Schellenberg, 2004). Analogously, studies investigating the effect of music training on cognitive ability such as spatial- and memory-related skills have provided no clear pattern of results. For example, in Bowels (2003), music training exerted a strong effect on children's visuospatial ability. Analogously, in Degé, Wehrum, Stark, and Schwarzer (2011) music training significantly enhanced the participants' visual and auditory memory. However, Rickard et al. (2012) failed to find any effect in either of the above measure. With regard to academic achievement, previous meta-analyses suggest that music training slightly enhances students' mathematical (Hetland & Winner, 2001; Vaughn, 2000) and literacy skills (Gordon, Fehd, & McCandliss, 2015). However, the overall effect sizes reported in these meta-analyses are modest, and the variability between studies is quite pronounced. Put simply, the effects of music training on skills such as spatial ability, memory, academic performance, and general intelligence are still controversial, and positive results have not always been replicated (Miendlarzewska & Trost, 2013).

Such variability in the results may be due to the design features of the studies, including (a) the age of the participants, (b) the random (or non-random) assignment to the treatment and control groups, and (c) the presence (or absence) of a group engaged in an

alternative activity to control for non-music-specific effects, such as placebos. Age may affect the occurrence of transfer of skills in two ways. First, transfer effects may be a function of neural plasticity (Buschkuhl, Jaeggi, & Jonides, 2012), which, in turn, is a function of age. Second, as students grow up, the level of specificity of the activities they are engaged in increases (e.g., mathematics, literacy, etc.). Crucially, research on expertise has shown that the higher the level of a particular ability, the more specific the features of that ability will be, and consequently, the lower the likelihood that transfer will occur (Ericsson & Charness, 1994; Gobet, 2016).

Quality design-related features may be important moderators too. Without random allocation of the participants, it is not always possible to ensure the baseline equivalence between experimental and control groups, especially if the experimental group is self-selected. Controlling for placebo effects could be even more important. In fact, the experience of a new activity such as music training may cause, *ipso facto*, an enhancement in children's and young adolescents' cognitive and academic skills. Music-related activities are usually a novelty for young students and may induce a state of motivation and excitement, which, in turn, may be the real cause of the observed (and temporary) improvements. Comparing music training with other enrichment activities is thus essential to understand whether the observed benefits are specifically due to music, or just the consequence of non-specific placebo effects.

1.4 Aims of the Present Meta-Analysis

Because of the theoretical implications for theories of transfer, the possible educational applications, and the current general interest in this topic, it is imperative to rigorously evaluate the putative benefits of music training for academic and cognitive skills. Similar claims have been made about the possibility of obtaining transferable benefits, both cognitive and academic, from playing video-games (Green, Li, & Bavelier, 2010; Green,

Pouget, & Bavelier, 2010), working memory training (Melby-Lervåg & Hulme, 2013, 2016), and playing chess in schools (Gobet & Campitelli, 2006; Sala & Gobet, 2016). However, research in these fields suggests that optimism about the positive effects of music training must be tempered by the possibility that the observed effects are due to confounding factors such as placebo effects (Boot, Blakely, & Simons, 2011; Gobet et al., 2014; Sala & Gobet, 2016) and lack of random assignment of the participants to the groups.

Our meta-analysis, then, examines the potential cognitive and academic benefits of music training for the general population of children and young adolescents (see 2.2. Inclusion/Exclusion Criteria). In a first stage, we estimated the overall size of the effects of music training on non-music cognitive and academic skills by comparing experimental groups to control groups. In a second phase, we assessed the potential role of several possible moderators on the effectiveness of music training. The analysis of these factors – along with the estimation of an overall effect size – aimed to test: (a) whether music training enhances students' cognitive and academic skills, or whether far-transfer from music to other areas is null or negligible; (b) whether music training improves some specific skills more than others; (c) whether students' age affects the benefits of music training; and (d) whether the methodological quality of the studies reviewed – i.e., random allocation of participants and comparisons with active (i.e., do-other) control groups to rule out placebo effects – influence the results.

Points a) and b) were tested by calculating a general overall effect size (see Section 3. Results) and the measure-specific overall effect sizes (see Sections 3.1 and 3.2), respectively. Points c) and d) were addressed by performing a meta-regression analyses (see Section 3.1).

2. Method

2.1 Literature Search

In line with the PRISMA statement (Moher, Liberati, Tetzlaff, & Altman, 2009), a systematic search strategy was used to find the relevant studies. Using the following combination of the keywords “music” AND (“training” OR “instruction” OR “education” OR “intervention”), Google Scholar, ERIC, Psyc-Info, ProQuest Dissertation & Theses, and Scopus databases were searched to identify all the potentially relevant studies. Also, previous narrative reviews were examined, and we e-mailed researchers in the field ($n = 11$) asking for unpublished studies and inaccessible data.¹⁶

2.2 Inclusion/Exclusion Criteria

The studies were included according to the following nine criteria:

1. The design of the study included music training; correlational and ex-post facto studies were excluded;
2. The independent variable (music training) was successfully isolated; the studies using integrated curricula (e.g., lessons of music and reading in the same intervention) were excluded;
3. The study presented a comparison between a music-treated group and, at least, one control group;
4. Music training was not merely environmental (e.g., background music, music videos);
5. During the study, a measure of academic and/or cognitive skill non-related to

¹⁶ Unfortunately, no author replied to our e-mails.

music was collected;

6. The participants of the study were pupils aged three to 16;
7. The participants of the study were pupils without any previous formal musical training (as stated by the authors of the included studies);
8. The participants of the study were pupils without any specific learning disability (e.g., developmental dyslexia) or clinical condition (e.g., autism);
9. The data presented in the study were sufficient to calculate an effect size.

To identify studies meeting these criteria, we searched for relevant published and unpublished articles in the last 30 years (from January 1, 1986, through March 1, 2016), and scanned reference lists.

Among the studies screened ($n = 166$), we found 38 studies, conducted from 1986 to 2016, that met all the inclusion criteria. These studies included 40 independent samples and 118 effect sizes, and a total of 3,085 participants.

2.3 Moderators

We selected four potential moderators. The first two, which we termed *theoretical moderators*, referred to features of the dependent variables and the participants of the studies, while the last two, which we termed *methodological moderators*, addressed more general methodological aspects:

1. Outcome measure (categorical variable): This variable includes literacy, mathematics, memory, intelligence, phonological processing, and spatial skills.¹⁷ Effect sizes that

¹⁷ These broad categories were built by aggregating different outcomes related to a particular cognitive or academic ability (e.g., reading and writing both under the category of literacy).

were not related to these categories (e.g., visual-auditory learning and visual attention) were labelled as *others*;

2. Age: The age of the participants in years (continuous variable);
3. Random allocation (dichotomous variable): Whether participants were fully randomly allocated to the groups;¹⁸
4. Presence of active control group (dichotomous variable): Whether the music training group was compared to another activity.

Table 6

Studies, dependent variables, and moderators of the 118 effect sizes included in the meta-analysis

For all the details about the dependent variables of the reviewed studies, see Table 1. See the Table S1 in the Supplemental material for more details about the descriptive statistics of the studies.

¹⁸ The category of “non-random” encompasses both pre-post-test studies and only-post-test studies. Two studies reported only post-test results: Cardarelli (2003) and Geoghegan and Mitchelmore (1996).

| Study | Dependent variable | Outcome measure | Age (years)^a | Randomization | Active control group |
|---------------------------------------|--------------------------------|----------------------------|--------------------------------|----------------------|-----------------------------|
| Bhide, Power, and Goswami (2013) - M1 | Working Memory (digit span) | Memory | 6.8 | No | Yes |
| Bhide, Power, and Goswami (2013) - M2 | Phonological awareness | Phonological processing | 6.8 | No | Yes |
| Bhide, Power, and Goswami (2013) - M3 | Spelling | Literacy | 6.8 | No | Yes |
| Bhide, Power, and Goswami (2013) - M4 | Reading | Literacy | 6.8 | No | Yes |
| Bilhartz, Bruhn, and Olson (2000) | Intelligence (vocabulary) | Intelligence | 4.5 | No | No |
| Bowels (2003) - M1 | Spatial temporal ability | Spatial | 6.5 | Yes | No |
| Bowels (2003) - M2 | Reading | Reading | 6.5 | Yes | No |
| Bowels (2003) - M3 | Phonological awareness | Phonological processing | 6.5 | Yes | No |

| | | | | | |
|---------------------------------------------------------|------------------------|-------------------------|------|-----|-----|
| Cardarelli (2003) - M1 | Reading | Reading | 9.0 | No | No |
| Cardarelli (2003) - M2 | Mathematics | Mathematics | 9.0 | No | No |
| Cogo-Moreira, de Avila, Ploubidis, and Mari (2013) - M1 | Phonological awareness | Phonological processing | 9.2 | Yes | No |
| Cogo-Moreira, de Avila, Ploubidis, and Mari (2013) - M2 | Reading | Literacy | 9.2 | Yes | No |
| Costa-Giomi (2004) - M1 | Mathematics | Mathematics | 9.0 | Yes | No |
| Costa-Giomi (2004) - M2 | Language | Literacy | 9.0 | Yes | No |
| Costa-Giomi (2004) - M3 | Mathematics | Mathematics | 9.0 | Yes | No |
| Degé and Schwarzer (2011) - S1 | Phonological awareness | Phonological processing | 5.8 | Yes | Yes |
| Degé and Schwarzer (2011) - S2 | Phonological awareness | Phonological processing | 5.8 | Yes | No |
| Degé, Wehrum, Stark, and Schwarzer (2011) - M1 | Visual memory | Memory | 10.8 | No | No |
| Degé, Wehrum, Stark, and Schwarzer (2011) - M2 | Memory (auditory) | Memory | 10.8 | No | No |
| Geoghegan and Mitchelmore (1996) | Mathematics | Mathematics | 4.5 | No | No |

| | | | | | |
|-----------------------------------------------------------------------------|-----------------------------|----------------------------|-----|-----|-----|
| | | Phonological | | | |
| Gromko (2005) | Phonemic awareness | processing | 5.5 | No | No |
| Gromko and Poorman (1998) | Intelligence | Intelligence | 3.5 | No | No |
| Hanson (2001) - M1 - S1 | Intelligence | Intelligence | 5.5 | Yes | No |
| Hanson (2001) - M1 - S2 | Intelligence | Intelligence | 5.5 | No | Yes |
| Hanson (2001) - M2 - S1 | Spatial-temporal ability | Spatial | 5.5 | Yes | No |
| Hanson (2001) - M2 - S2 | Spatial-temporal ability | Spatial | 5.5 | No | Yes |
| Hanson (2001) - M3 - S1 | Spatial recognition | Spatial | 5.5 | Yes | No |
| Hanson (2001) - M3 - S2 | Spatial recognition | Spatial | 5.5 | No | Yes |
| Herrera, Lorenzo, Defior, Fernandez-Smith, and Costa-Giomi (2011) - M1 - S1 | Phonological awareness | Phonological processing | 4.6 | Yes | Yes |
| Herrera, Lorenzo, Defior, Fernandez-Smith, and Costa-Giomi (2011) - M2 - S1 | Naming speed | Phonological processing | 4.6 | Yes | Yes |
| Herrera, Lorenzo, Defior, Fernandez-Smith, and Costa-Giomi (2011) - M1 - S2 | Phonological awareness | Phonological processing | 4.6 | Yes | Yes |

| | | | | | |
|-----------------------------------------------------------------------------|------------------------------------|-------------------------|-----|-----|-----|
| Herrera, Lorenzo, Defior, Fernandez-Smith, and Costa-Giomi (2011) - M2 - S2 | Naming speed | Phonological processing | 4.6 | Yes | Yes |
| Hole (2013) | Reading | Reading | 8.9 | No | Yes |
| Hunt (2012) | Phonological discrimination | Phonological processing | 3.5 | Yes | No |
| Janus, Lee, Moreno, and Bialystok (2016) - M1 | Working Memory (verbal) | Memory | 5.0 | No | Yes |
| Janus, Lee, Moreno, and Bialystok (2016) - M2 | Working Memory (spatial) | Memory | 5.0 | No | Yes |
| Janus, Lee, Moreno, and Bialystok (2016) - M3 | Executive Control (verbal fluency) | Other | 5.0 | No | Yes |
| Janus, Lee, Moreno, and Bialystok (2016) - M4 | Attention (sentence judgement) | Other | 5.0 | No | Yes |
| Janus, Lee, Moreno, and Bialystok (2016) - M5 | Attention (visual search) | Other | 5.0 | No | Yes |
| Kaviani, Mirbaha, Pournaseh, and Sagan (2014) - M1 | Intelligence (IQ) | Intelligence | 5.5 | Yes | No |
| Kaviani, Mirbaha, Pournaseh, and Sagan (2014) - M2 | Abstract reasoning | Intelligence | 5.5 | Yes | No |

| | | | | | |
|----------------------------------------------------|----------------------|--------------|-----|-----|-----|
| Kaviani, Mirbaha, Pournaseh, and Sagan (2014) - M3 | Verbal reasoning | Intelligence | 5.5 | Yes | No |
| | Quantitative | | | | |
| Kaviani, Mirbaha, Pournaseh, and Sagan (2014) - M4 | reasoning | Intelligence | 5.5 | Yes | No |
| Kaviani, Mirbaha, Pournaseh, and Sagan (2014) - M5 | Short-term memory | Memory | 5.5 | Yes | No |
| Legette (1993) - M1 | Mathematics | Mathematics | 6.0 | No | No |
| Legette (1993) - M2 | Reading | Reading | 6.0 | No | No |
| Lu (1986) | Reading | Reading | 6.0 | No | Yes |
| | Spatial navigation | | | | |
| Mehr, Schachner, Katz, and Spelke (2013) - M1 - S1 | reasoning | Spatial | 4.0 | Yes | Yes |
| | Spatial navigation | | | | |
| Mehr, Schachner, Katz, and Spelke (2013) - M1 - S2 | reasoning | Spatial | 4.0 | Yes | No |
| Mehr, Schachner, Katz, and Spelke (2013) - M2 - S1 | Visual form analysis | Spatial | 4.0 | Yes | Yes |
| Mehr, Schachner, Katz, and Spelke (2013) - M2 - S2 | Visual form analysis | Spatial | 4.0 | Yes | No |
| | Numerical | | | | |
| Mehr, Schachner, Katz, and Spelke (2013) - M3 - S1 | discrimination | Mathematics | 4.0 | Yes | Yes |
| Mehr, Schachner, Katz, and Spelke (2013) - M3 - S2 | Numerical | Mathematics | 4.0 | Yes | No |

| | | | | | |
|-------------------------------------------------------------|--------------------------|-------------------------|-----|-----|-----|
| | discrimination | | | | |
| Mehr, Schachner, Katz, and Spelke (2013) - M4 - S1 | Receptive vocabulary | Literacy | 4.0 | Yes | Yes |
| Mehr, Schachner, Katz, and Spelke (2013) - M4 - S2 | Receptive vocabulary | Literacy | 4.0 | Yes | No |
| Moreno et al. (2009) | Reading | Literacy | 8.3 | No | Yes |
| Moreno, Friesen, and Bialystok (2011) - M1 | Rhyming | Phonological processing | 5.3 | Yes | Yes |
| Moreno, Friesen, and Bialystok (2011) - M2 | Visual-auditory learning | Other | 5.3 | Yes | Yes |
| Moritz, Yampolsky, Papadelis, Thomson, and Wolf (2013) - M1 | Rhyming | Phonological processing | 5.6 | No | No |
| Moritz, Yampolsky, Papadelis, Thomson, and Wolf (2013) - M2 | Isolation of phonemes | Phonological processing | 5.6 | No | No |
| Myant, Armstrong, and Healy (2008) - M1 | Alliteration | Phonological processing | 4.3 | No | No |
| Myant, Armstrong, and Healy (2008) - M2 | Rhyming | Phonological processing | 4.3 | No | No |
| Portowitz, Lichtenstein, Egorova, and Brand (2009) - | Intelligence | Intelligence | 8.0 | No | No |

| | | | | | |
|------------------------------------------------------|------------------|--------------|-----|----|-----|
| M1 | | | | | |
| Portowitz, Lichtenstein, Egorova, and Brand (2009) - | | | | | |
| M2 | Memory | Memory | 8.0 | No | No |
| | Working Memory | | | | |
| Portowitz, Pepler, and Downton (2014) | (spatial) | Memory | 9.5 | No | No |
| Rauscher and Zupan (2000) - M1 | Working memory | Memory | 5.5 | No | No |
| | Spatial-temporal | | | | |
| Rauscher and Zupan (2000) - M2 | ability | Spatial | 5.5 | No | No |
| | | Phonological | | | |
| Register (2004) - M1 - S1 | Letter naming | processing | 6.0 | No | Yes |
| | | Phonological | | | |
| Register (2004) - M1 - S2 | Letter naming | processing | 6.0 | No | No |
| | | Phonological | | | |
| Register (2004) - M2 - S1 | Sounds fluency | processing | 6.0 | No | Yes |
| | | Phonological | | | |
| Register (2004) - M2 - S2 | Sounds fluency | processing | 6.0 | No | No |
| Register (2004) - M3 - S1 | Reading | Literacy | 6.0 | No | Yes |

| | | | | | |
|----------------------------------------------|-------------------|--------------|------|-----|-----|
| Register (2004) - M3 - S2 | Reading | Literacy | 6.0 | No | No |
| Rickard, Bambrick, and Gill (2012) - M1 - S1 | Memory | Memory | 12.7 | No | Yes |
| Rickard, Bambrick, and Gill (2012) - M1 - S2 | Memory | Memory | 12.7 | No | Yes |
| Rickard, Bambrick, and Gill (2012) - M2 - S1 | Intelligence (IQ) | Intelligence | 12.7 | No | Yes |
| Rickard, Bambrick, and Gill (2012) - M2 - S2 | Intelligence (IQ) | Intelligence | 12.7 | No | Yes |
| Rickard, Bambrick, and Gill (2012) - M3 - S3 | Reading | Literacy | 10.9 | Yes | Yes |
| Rickard, Bambrick, and Gill (2012) - M3 - S4 | Reading | Literacy | 10.9 | Yes | No |
| Rickard, Bambrick, and Gill (2012) - M4 - S3 | Writing | Literacy | 10.9 | Yes | Yes |
| Rickard, Bambrick, and Gill (2012) - M4 - S4 | Writing | Literacy | 10.9 | Yes | No |
| Rickard, Bambrick, and Gill (2012) - M5 - S3 | Speaking | Literacy | 10.9 | Yes | Yes |
| Rickard, Bambrick, and Gill (2012) - M5 - S4 | Speaking | Literacy | 10.9 | Yes | No |
| Rickard, Bambrick, and Gill (2012) - M6 - S3 | Space | Spatial | 10.9 | Yes | Yes |
| Rickard, Bambrick, and Gill (2012) - M6 - S4 | Space | Spatial | 10.9 | Yes | No |
| Rickard, Bambrick, and Gill (2012) - M7 - S3 | Number | Mathematics | 10.9 | Yes | Yes |
| Rickard, Bambrick, and Gill (2012) - M7 - S4 | Number | Mathematics | 10.9 | Yes | No |

| | | | | | |
|-----------------------------------------------|----------------------------------|-------------|------|-----|-----|
| Rickard, Bambrick, and Gill (2012) - M8 - S3 | Structure | Mathematics | 10.9 | Yes | Yes |
| Rickard, Bambrick, and Gill (2012) - M8 - S4 | Structure | Mathematics | 10.9 | Yes | No |
| Rickard, Bambrick, and Gill (2012) - M9 - S3 | Measurement | Mathematics | 10.9 | Yes | Yes |
| Rickard, Bambrick, and Gill (2012) - M9 - S4 | Measurement | Mathematics | 10.9 | Yes | No |
| Rickard, Bambrick, and Gill (2012) - M10 - S3 | Mathematics | Mathematics | 10.9 | Yes | Yes |
| Rickard, Bambrick, and Gill (2012) - M10 - S4 | Mathematics | Mathematics | 10.9 | Yes | No |
| Roden et al. (2014) - M1 | Visual attention | Other | 7.9 | No | Yes |
| Roden et al. (2014) - M2 | Speed processing | Other | 7.9 | No | Yes |
| Roden, Grube, Bongard, and Kreutz (2014) - M1 | Working memory (visuospatial) | Memory | 7.5 | No | Yes |
| Roden, Grube, Bongard, and Kreutz (2014) - M2 | Working memory (phonological) | Memory | 7.5 | No | Yes |
| Roden, Grube, Bongard, and Kreutz (2014) - M3 | Working memory (CE) | Memory | 7.5 | No | Yes |
| Roden, Kreutz, and Bongard (2012) - M1 - S1 | Verbal memory | Memory | 7.7 | No | Yes |
| Roden, Kreutz, and Bongard (2012) - M1 - S2 | Verbal memory | Memory | 7.7 | No | No |

| | | | | | |
|-----------------------------------------------------|-----------------------------------|-------------------------|-----|-----|-----|
| Roden, Kreutz, and Bongard (2012) - M2 - S1 | Visual memory | Memory | 7.7 | No | Yes |
| Roden, Kreutz, and Bongard (2012) - M2 - S2 | Visual memory | Memory | 7.7 | No | No |
| Schellenberg (2004) - S1 | Intelligence (IQ) | Intelligence | 6.0 | Yes | Yes |
| Schellenberg (2004) - S2 | Intelligence (IQ) | Intelligence | 6.0 | Yes | No |
| Schellenberg, Corrigan, Dys, and Malti (2015) | Vocabulary | Literacy | 8.7 | No | No |
| Slater et al. (2014) - M1 | Reading | Reading | 8.3 | No | No |
| Slater et al. (2014) - M2 | Phonological awareness | Phonological processing | 8.3 | No | No |
| Slater et al. (2014) - M3 | Phonological memory | Phonological processing | 8.3 | No | No |
| Slater et al. (2014) - M4 | Rapid naming | Phonological processing | 8.3 | No | No |
| Thompson, Schellenberg, and Husain (2004) - M1 - S1 | Speech Prosody (spoken utterance) | Phonological processing | 7.0 | Yes | Yes |
| Thompson, Schellenberg, and Husain (2004) - M1 - S2 | Speech Prosody (spoken utterance) | Phonological processing | 7.0 | Yes | No |

| | | | | | |
|-----------------------------------------------------|--------------------------------|-------------------------|------|-----|-----|
| Thompson, Schellenberg, and Husain (2004) - M2 - S1 | Speech Prosody (tone sequence) | Phonological processing | 7.0 | Yes | Yes |
| Thompson, Schellenberg, and Husain (2004) - M2 - S2 | Speech Prosody (tone sequence) | Phonological processing | 7.0 | Yes | No |
| Tierney, Krizman, and Kraus (2015) - M1 | Phonological awareness | Phonological processing | 14.7 | No | Yes |
| Tierney, Krizman, and Kraus (2015) - M2 | Phonological memory | Phonological processing | 14.7 | No | Yes |
| Tierney, Krizman, and Kraus (2015) - M3 | Phonological awareness | Phonological processing | 14.7 | No | Yes |
| Yazejian and Peisner-Feinberg (2009) - M1 | Phoneme deletion | Phonological processing | 4.4 | Yes | No |
| Yazejian and Peisner-Feinberg (2009) - M2 | Rhyming | Phonological processing | 4.4 | Yes | No |

Note. For studies with multiple samples, the result of each sample (S1, S2, etc.) is reported separately, and for studies with multiple outcome measures, the result of each measure (M1, M2, etc.) is reported separately.

^a When the mean age was not provided, the mid-point of the range was inserted in the model. Similarly, when the grade of the students was provided the mid-point of the range was considered (e.g., first graders, six-year-olds).

2.4 Effect Size¹⁹

For the studies with an *only-post-test design*, the standardized means difference (Cohen's d) was calculated with the following formula:

$$d = (M_e - M_c) / SD_{pooled} \quad (1)$$

where SD_{pooled} is the pooled standard deviation, and M_e and M_c are the means of the experimental group and the control group, respectively.²⁰ For the studies with a *repeated-measure design*, the standardized means difference was calculated with the following formula:

$$d = (M_{g-e} - M_{g-c}) / SD_{pooled-pre} \quad (2)$$

where $SD_{pooled-pre}$ is the pooled standard deviation of the two pre-test standard deviations, and M_{g-e} and M_{g-c} are the gain of the experimental group and the control group, respectively (Schmidt & Hunter, 2015, p. 353).

Analogously to other recent meta-analyses (e.g., Macnamara, Hambrick, & Oswald, 2014), the effect sizes with z -scores greater than 3 ($n = 9$) or smaller than -3 ($n = 2$) were Winsorized to z -scores equal to 2.99 and -2.99 , respectively.²¹ This procedure was adopted to reduce the weight of potential outliers in the analysis (Lipsey & Wilson, 2001; Schmidt & Hunter, 2015, pp. 235-236; Tukey, 1962). Finally, the Comprehensive Meta-Analysis (Version 3.0; Biostat, Englewood, NJ) software package was used for computing the effect sizes and conducting statistical analyses.

¹⁹ All the formulas we used were taken from Schmidt and Hunter (2015).

²⁰ If the t statistic was provided, we used the regular formula $d = t \times \sqrt{(n_1 + n_2) / (n_1 \times n_2)}$.

²¹ We also performed additional analyses without Winsorizing the 11 effect sizes. No significant difference was found in the overall results (for details, see Section S1 and Table S3 in the Supplemental material available online).

2.5 Statistical Dependence of the Samples

The effect sizes were calculated for each dependent variable reported in the studies (Schmidt & Hunter, 2015). Moreover, when the study presented a comparison between one experimental group and two control groups (do-nothing and active), two effect sizes were calculated (one for each comparison with experimental and control groups; see Table 6). As this procedure violates the principle of statistical independence, the method designed by Cheung and Chan (2004) was applied to both the main and the additional models (see Sections 3.1 and 3.2). This method reduces the weight in the analysis of dependent samples by calculating an adjusted (i.e., smaller) N . Since Cheung and Chan's (2004) method cannot be used for partially dependent samples, we ran our analyses as if the comparisons between experimental samples and two different control groups were statistically independent. However, it must be noted that the violation of statistical independence has little or no effect on means, standard deviations, and confidence intervals (Bijmolt & Pieters, 2001; Tracz, Elmore, & Pohlmann, 1992). Thus, the entire procedure is a reliable way to deal with the statistical dependence of part of the samples. For the list of the studies and the adjusted N s, see Table S2 in the Supplemental material available online (<http://www.sciencedirect.com/science/article/pii/S1747938X16300641>).

3. Results

The random-effects meta-analytic overall effect size was $\bar{d} = 0.16$, CI [0.09; 0.22], $k = 118$, $p < .001$. The degree of heterogeneity (Borenstein, Hedges, Higgins, & Rothstein, 2009) between effect sizes was $I^2 = 46.94$, suggesting that some moderators had a potential effect.²²

²² A degree of heterogeneity (I^2) around 50.00 is considered moderate, around 25.00 low, and around 75.00 high.

3.1 Meta-Regression Analysis

A meta-regression model including all the four moderators was run. The model fitted the data significantly, $Q(9) = 49.06$, $R^2 = .65$, $p < .001$. Age was not a significant moderator, $p = .944$. The statistically significant moderators were Outcome measure, $Q(6) = 21.78$, $p = .001$, Random allocation, $b = -0.16$, $p = .010$, and Presence of active control group, $b = -0.25$, $p < .001$. The last two moderators show that studies with random allocation of participants and studies comparing music treatment to another activity (active control group) tended to have weaker effect sizes. The overall effect sizes in randomized and non-randomized samples were $\bar{d} = 0.09$, CI $[-0.01; 0.18]$, $k = 57$, $p = .084$, and $\bar{d} = 0.23$, CI $[0.14; 0.31]$, $k = 61$, $p < .001$, respectively. The overall effect sizes when music training was compared to active control and do-nothing control groups were $\bar{d} = 0.03$, CI $[-0.07; 0.12]$, $k = 54$, $p = .562$, and $\bar{d} = 0.25$, CI $[0.17; 0.34]$, $k = 64$, $p < .001$, respectively. Finally, the overall effect size in randomized samples with active control groups was $\bar{d} = -0.12$, CI $[-0.27; 0.03]$, $k = 22$, $p = .113$, while the overall effect size in the non-randomized samples without active control group was $\bar{d} = 0.33$, CI $[0.23; 0.44]$, $k = 29$, $p < .001$.

3.2 Additional Meta-Analytic Models

Since Outcome measure was a significant moderator, we calculated the random-effects meta-analytic overall effect size of each of the seven measures, in order to investigate whether any measure showed an overall effect size appreciably larger (or smaller) than the others. The overall effect sizes are summarized in Table 7.

The meta-regression analysis showed that only memory- and intelligence-related overall effect sizes were significantly different compared to the other measures ($b = 0.26$, $p = .041$, and $b = 0.30$, $p = .029$, respectively).

Table 7

Overall effect sizes, confidence intervals, ks, and p-values in each outcome measure

| Outcome measure | Effect size (\bar{d}) | 95% CI | <i>k</i> | <i>p</i>-value |
|-------------------------|-------------------------------------------|---------------|-----------------|-----------------------|
| Literacy | -0.07 | [-0.23; 0.09] | 22 | .386 |
| Mathematics | 0.17 | [-0.02; 0.36] | 15 | .085 |
| Memory | 0.34 | [0.20; 0.48] | 18 | <.001 |
| Intelligence | 0.35 | [0.21; 0.49] | 13 | <.001 |
| Phonological Processing | 0.17 | [0.04; 0.29] | 32 | .008 |
| Spatial | 0.14 | [-0.06; 0.34] | 12 | .168 |
| Others | -0.01 | [-0.25; 0.23] | 6 | .919 |

3.3 Publication Bias Analysis

Begg and Mazumdar's (1994) rank correlation test showed no evidence of publication bias ($p = .433$, one-tailed). In addition, to test the robustness of results (Kepes & McDaniel, 2015), we ran a p -curve analysis for the detection of publication bias (Simonsohn, Nelson, & Simmons, 2014). We selected the p s according the following two rules: (a) only positive results (i.e., $z > 0$) were considered; and (b) to avoid redundancy, only one $p < .01$ per study was inserted into the analysis. The results had evidential value (i.e., no evidence of publication bias) because we found more low p -values ($p < .01$) than high p -values ($.01 < p < .05$), $z(14) = -4.24$, $p < .001$ (Figure 10).

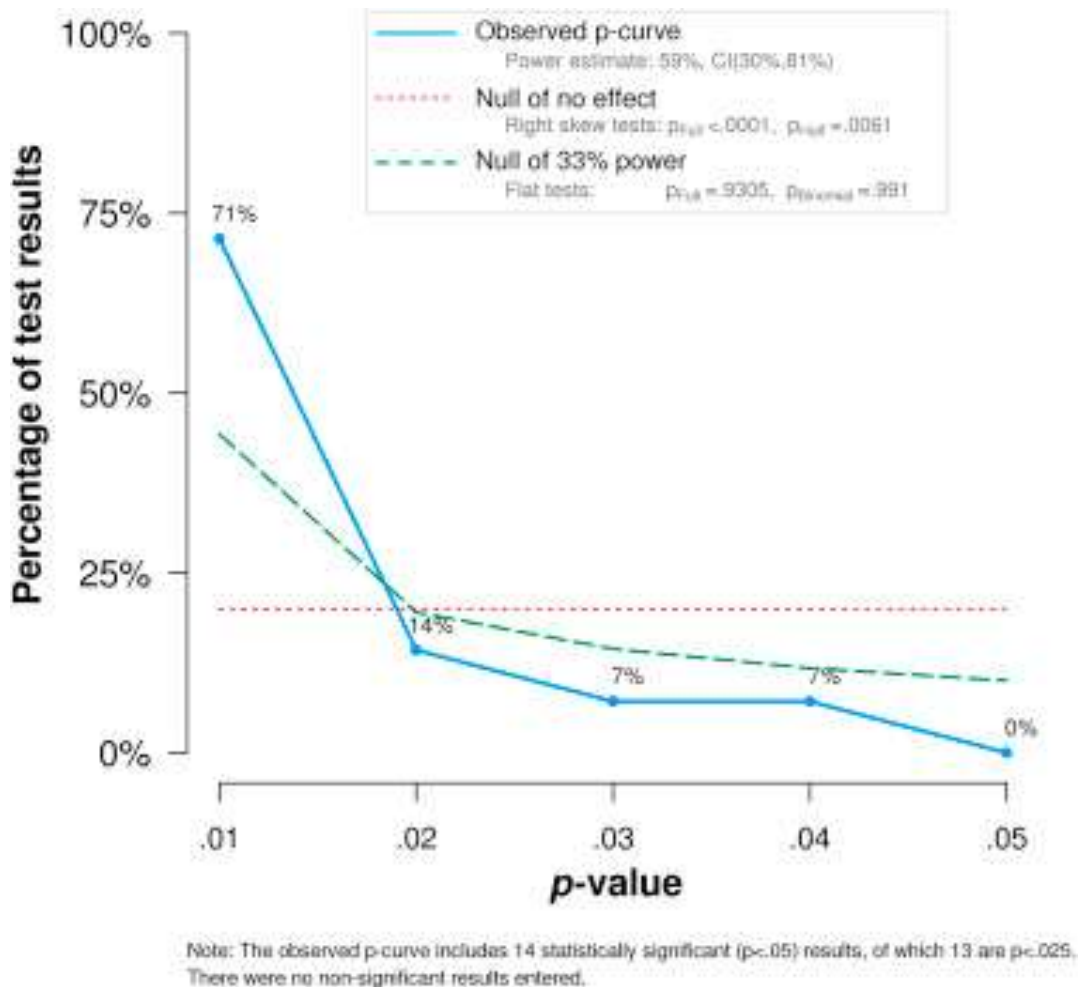


Figure 10. *p*-curve analysis of the studies reporting significantly positive results. The blue (continuous) line shows that most of the significant *p*-values are smaller than .01.

Finally, Duval and Tweedie’s (2000) method found no publication bias in any of the seven models (i.e., no studies trimmed left of the mean).

3.4 Sensitivity Analysis

Since Rickard et al.’s (2012) study reported a large number of effect sizes ($k = 20$), we conducted a sensitivity analysis by excluding those effect sizes from all the models. The random-effects meta-analytic overall effect size was still modest, $\bar{d} = 0.20$, CI [0.14; 0.27], $k = 98$, $p < .001$. The degree of heterogeneity between effect sizes was $I^2 = 39.31$, suggesting that some moderators had a potential effect. For the list of the studies and the adjusted *N*s, see Table S4 in the Supplemental material available online.

A meta-regression model including all the four moderators was run. The model fitted the data significantly, $Q(9) = 36.94$, $R^2 = .74$, $p < .001$. The only two statistically significant moderators were Outcome measure, $Q(6) = 20.16$, $p = .003$ and Presence of active control group, $b = -0.17$, $p = .0014$. The overall effect sizes when music training was compared to do-nothing control and active control groups were $\bar{d} = 0.28$, CI [0.19; 0.36], $k = 56$, $p < .001$, and $\bar{d} = 0.08$, CI [-0.03; 0.19], $k = 42$, $p = .139$, respectively. Compared to Table 7, no significant difference was found in six of the seven Outcome measure-related overall effect sizes. The only exception was Mathematics ($\bar{d} = 0.35$ vs $\bar{d} = 0.17$; Table 8).

Table 8

Overall effect sizes, confidence intervals, ks, and p-values in each outcome measure

| Outcome measure | Effect size (\bar{d}) | 95% CI | <i>k</i> | <i>p</i>-value |
|-------------------------|-------------------------------------------|---------------|-----------------|-----------------------|
| Literacy | 0.07 | [-0.07; 0.21] | 16 | .307 |
| Mathematics | 0.35 | [0.16; 0.54] | 7 | <.001 |
| Memory | 0.39 | [0.25; 0.54] | 16 | <.001 |
| Intelligence | 0.37 | [0.21; 0.53] | 11 | <.001 |
| Phonological Processing | 0.17 | [0.04; 0.29] | 32 | .008 |
| Spatial | 0.15 | [-0.10; 0.40] | 10 | .248 |
| Others | -0.01 | [-0.25; 0.23] | 6 | .919 |

Note. The 20 effect sizes calculated from Rickard et al. (2012) were excluded.

4. Discussion

The present meta-analysis aimed to test the hypothesis that music training improves children's and young adolescents' cognitive and academic skills, and to evaluate the potential role of moderating variables. Along with a small overall effect size ($\bar{d} = 0.16$, CI [0.09; 0.22]), which

indicates that far-transfer from music to non-music skills was limited, the results showed a slightly greater positive effect of music training on some of the cognitive skills (i.e., intelligence and memory) and a non-significant effect on all the academic skills. Moreover, the design quality of the studies significantly affected the magnitude of the effects. A similar pattern of results was obtained in the sensitivity analysis model.

We did not correct for attenuation due to measurement error because only about half of the studies provided reliability coefficients. However, correcting for measurement error would not significantly affect the effect sizes. For example, if we assume that the reliability coefficients are between .80 and .90, then the corrected estimate of the overall effect size of the main model (i.e., $\bar{d} = 0.16$) would be between 0.17 and 0.18, a difference of only 0.01 or 0.02 standard deviations.

4.1 Substantive Results

The outcomes of the present meta-analysis allow us to draw some important conclusions. First, the small overall effect size upholds Thorndike and Woodworth's (1901) common-element theory. In line with previous research (Donovan et al., 1999; Sala & Gobet, 2016), far-transfer from music to other cognitive or academic abilities seems to be small or null. Second, music training appears to moderately foster intelligence- and memory-related outcomes. However, no significant effect on academic skills was found (literacy, $\bar{d} = -0.07$, CI [-0.23; 0.09], $p = .386$; mathematics, $\bar{d} = 0.17$, CI [-0.02; 0.36], $p = .085$). This outcome suggests that improvements in memory and intelligence do not generalize to academic skills. Alternatively, and more likely, the observed positive effects of music training in intelligence- and memory-related outcomes are due to confounding variables (we will take up this point below). Either way, the hypothesis according to which the multiple benefits of music training, including academic benefits, stem from an improvement in general intelligence (or overall cognitive skill) is not corroborated. Third, the age of the participants is not a statistically significant moderator. Fourth, the meta-regression model accounts for a large proportion of the variance ($R^2 = .65$) between the effect sizes. The latter result

implies that the statistically significant moderators explain, to a large extent, why the research on the effects of music training on children's and young adolescents' skills has produced mixed results up to now.

4.2 Methodological Results

The meta-regression analysis shows that both methodological moderators (i.e., random allocation of participants to the treatment groups and comparison to an active control group) affected the effect sizes. In other words, the better the design quality, the smaller the effect sizes. This outcome lends further support to the idea that the observed positive effects, when any, of music training on non-music-related outcomes, are probably due to confounding variables, such as placebo effects and lack of random allocation of participants.

Unfortunately, this conclusion seems to apply to memory- and intelligence-related effect sizes too. In fact, despite the greater overall effect sizes in these two outcome measures ($\bar{d} = 0.34$, CI [0.20; 0.48] and $\bar{d} = 0.35$, CI [0.21; 0.49], respectively), the reliability of these positive results seems questionable. Only one study (Schellenberg, 2004) tested the effect of music training on children's intelligence using a rigorous experimental design (i.e., random allocation of participants and active control group), and the effect was found to be modest ($d = 0.16$). Concerning the memory-related outcomes, none of the reviewed studies adopted such a design. Furthermore, as pointed out above, a genuine – i.e., not due to confounding variables – improvement in such critical cognitive skills should leave a trace in students' academic skills, at least to some degree.

The sensitivity analysis (Section 3.4) showed that when Rickard et al.'s (2012) study and all its effect sizes were excluded, the overall effect size in mathematics became significantly positive. However, the only study comparing a music training group to an active control group and with random allocation of the participants to the groups – i.e., Mehr et al. (2013) – found a negative effect size ($d = -0.25$). These considerations uphold the conclusion that music training does not substantially enhance any non-music-related cognitive skill.

4.3 Conclusions and Recommendations for Future Research

The results of this meta-analysis fail to support the hypothesis that music skill transfers to cognitive or academic skills in the general population of children and young adolescents. Together with previous findings in psychology and education, these results suggest a sobering conclusion: when the potential occurrence of far-transfer is tested rigorously, the results are often, if not always, disappointing. Thus, this study lends further support to the hypothesis according to which far-transfer rarely occurs. Even when music training appears to foster some of the participants' cognitive skills (intelligence and memory), the reliability of the results is doubtful. In fact, only one study investigated, with a proper design, the effects exerted by music training on the participants' intelligence- and memory-related skills.

Due to the lack of well-designed studies, the question of whether music training enhances children's and young adolescents' intelligence- and memory-related skills is still unanswered. For this reason, future studies should strive for proper designs that include both random allocation of the participants and an active control group. Furthermore, future investigations should evaluate the effects of music training on both cognitive (especially intelligence and memory) and academic skills. Such a design makes it possible to empirically assess whether the potential benefits of music training on youngsters' cognitive skills generalize to academic performance. Nonetheless, considering the previous unsatisfactory outcomes and the scarcity of far-transfer in the literature, it is our opinion that future experiments will show results in line with those presented in this meta-analysis.

Chapter 7: Meta-Analysis of Video-Game Training

Rationale for the Meta-Analysis in Chapter 7

Chapter 7 explores the relationship between the practice of video-games and cognitive function. Meta-analysis 1 assesses the correlation between video-game skill and several cognitive abilities. Meta-analysis 2 compares the performance of video-game players and non-players in a variety of cognitive tasks. Meta-analysis 3 examines the effects of video-game training experimentally (i.e., studies with an intervention design).

Considering the relevance of the field of video-game playing, the large number of studies carried out, and the substantial disagreement among researchers, such broader meta-analytic investigation is necessary. My aim is to investigate the correlational and cross-sectional evidence – along with the experimental studies – regarding video-game practice and cognitive abilities to test the claims about the benefits of video-game training.

Please note that the original paper included a broad introduction to the general questions of transfer and cognitive training. The introduction has been moved to Chapter 2 to avoid redundancy.

The studies included in Meta-analysis 1, Meta-analysis 2, and Meta-analysis 3 are listed in Appendix D, Appendix E, and Appendix F, respectively.

1. Introduction

As just seen, recent experimental evidence and meta-analytic reviews have highlighted the limitations, rather than the benefits, of many different types of cognitive training. Cognitive-training regimens seem to affect only the trained skills, while no effect is exerted on non-trained tasks. This applies to both those activities that specifically train cognitive abilities (e.g., *n*-back tasks in WM training, spatial training, and brain-training programs) and cognitively demanding activities such as chess and music. The converging evidence provided by the research into expertise acquisition and cognitive training strongly suggests that the occurrence of far transfer is rare at best.

Video game training offers another potential avenue for cognitive enhancement. Unlike chess and music training, where the number of studies is limited, video game training has been extensively studied for the last 20 years. The deep interest of scientists and policy makers for this activity has made the research on video games one of the most important domains in which to test the occurrence of far transfer. Action video-game players have been found to outperform non-players in a variety of attentional and perceptual tasks (Green, Li, & Bavelier, 2010). Crucially, several experimental studies (e.g., Bejjanki et al., 2014; Green & Bavelier 2003) have provided some evidence of a causal relationship between action video game training and improvement in cognitive ability. Notably, even the US Navy has been attracted by these promising results (Hsu, 2010).

The most influential explanation proposed to account for those positive results is the “learning to learn” theory (Bavelier, Green, Pouget, & Schrater, 2012). According to this theory, experience with action video games leads to an improvement in probabilistic inference. It is argued that the tasks that are used to compare the performance of video game players and non-video game players all require computing the probability of a choice being true given the available information. In other words, video game players are better at using such information, and this improved computational ability leads to better performance across tasks.

Finally, non-action video game training seems to offer some benefits as well. For example, Okagaki and Frensch (1994) reported that a 6-hour training of the game Tetris improved the spatial abilities in a group of older adolescents.

Playing video games also seems associated with neural changes (functional and anatomical). For example, enhanced attentional control due to video gaming is consistent with several fMRI studies revealing that video game players have superior functional integration between working memory and attention networks involving frontoparietal areas (Gong et al., 2016), as well as enhanced white matter connectivity from the visual area to frontal cortex (Kim et al., 2015). Wu et al. (2012) trained non-video game players with an action video game (Medal of Honor) for 10 hours and measured event-related potentials. After video game training, high-performing players showed larger amplitudes of P3 waves, which have been implicated in top-down control of attention.

The research into video game training has, however, failed to consistently replicate the abovementioned positive results. Terlecki, Newcombe, and Little (2008) found no difference in mental rotation ability between the training group (playing Tetris) and the control group. Similarly, Minear et al.'s (2016) study of real-time strategy video game provided no evidence of training effects on several measures of WM, short-term memory, spatial ability, and fluid intelligence. Boot, Kramer, Simons, Fabiani, and Gratton (2008) questioned the effectiveness of action video game training at enhancing a broad set of cognitive abilities (e.g., enumeration, span, and *n*-back tasks). Finally, Oei and Patterson (2013, 2014, 2015) have challenged the "learning to learn" hypothesis and claimed that action video game training fosters, at best, those cognitive abilities necessary to play a particular video game. To test this hypothesis, Oei and Patterson (2015) used, as training tasks, four different action video games, differing from each other with regard to their cognitive demands (e.g., different speed and levels of selective attention). In line with Thorndike and Woodworth's (1901) common elements theory, participants' improvements were restricted to the cognitive abilities targeted by the video game they played.

Other researchers have also raised doubts about the alleged superior cognitive ability of video game players over non-players. For example, in Gobet et al. (2014), the group of action video game players failed to outperform the non-video game players in a flanker task and a change detection task. Similarly, Murphy and Spencer (2009) found no difference between a group of action video game players and a group of non-players in a set of visual-attention tasks. Comparable outcomes were obtained by Castel, Pratt, and Drummond (2005) and Irons, Remington, and McLean (2011).

A further source of scepticism about the relationship between video game playing and superior cognitive ability comes from several correlational studies. If video game training is effective, more skilled and experienced video game players should show superior cognitive ability compared to novice video game players. However, Hambrick, Oswald, Darowski, Rench, and Brou (2010) reported near-zero correlations between the participants' video game experience and several measures of processing speed, WM capacity, and fluid reasoning. Hambrick et al.'s (2010) results were replicated by Unsworth et al. (2015).

2. The Meta-Analytic Evidence

The research about video game and cognitive ability has provided mixed results in both experimental, quasi-experimental (i.e., comparison between players and non-players), and correlational studies. To disentangle these discrepancies, Powers, Brooks, Aldrich, Palladino, and Alfieri (2013) ran two meta-analyses collecting the available evidence about the effects of playing video games on cognitive ability.²³ The first meta-analysis, comparing players to non-players,

²³ More recently, two other smaller meta-analyses were carried out. Toril, Reales, and Ballesteros (2014) reviewed 20 studies regarding the effects of video game training on older adults' cognitive ability, while Wang et al. (2016) meta-analysed 19 studies regarding action video game training in healthy adults. Both meta-analyses reported moderate effects of training on participants' cognitive

reported medium to large effect sizes showing that video-game players were superior to non-players in measures of visual processing, executive functioning, and spatial imagery, among others. The second meta-analysis, focusing on true experiments, found positive, yet slightly smaller, effects of video game training on the same measures. Overall, the results suggested optimism about the ability of video game training to enhance a broad range of cognitive abilities.

However, several serious methodological flaws make Powers et al.'s (2013) findings unreliable. The inclusion criteria appear too loose, especially because of the inclusion of training studies without a control group controlling for testing effects, studies mixing video game experience with general computer use (e.g., Li & Atkins, 2004), and studies dealing with the effects of exergaming (i.e., games for physical training; e.g., Staiano, Abraham, & Calvert, 2012). Another problem was that Powers et al.'s (2013) meta-analysis included, when assessing the differences between video game players and non-video game players, studies reporting correlations between video game players' experience/skill and cognitive ability (e.g., Hambrick et al., 2010).

Most importantly, in the two meta-analytic models (and hence in the sub-models), too many of the effect sizes (up to 28) were extracted from the same samples and were often referring to the same cognitive construct, without any correction for statistical dependence. Even if the violation of the assumption of statistical independence does not necessarily cause a systematic bias in the estimation of overall meta-analytic means (i.e., \bar{r} and \bar{d} ; Schmidt & Hunter, 2015), the features of a particular meta-analytic model may lead to an accidental inflation (or reduction) of the overall means. Moreover, the violation of the assumption of statistical independence is associated with an underestimation of sampling error inflating the variability between studies (Schmidt & Hunter, 2015), with possible consequent biases in moderator analysis.

3. The Present Meta-Analytic Investigation

The field of video game training might be a significant exception to Thorndike and Woodworth's (1901) common elements hypothesis. The potential theoretical and practical implications of such an anomaly would be huge. It is thus imperative to test – comprehensively and with rigorous statistical methods – the claim that video game training produces far-transfer effects.

We thus ran three meta-analytic models. The first meta-analysis assessed the correlation between video-game skill and cognitive ability. To the best of our knowledge, no such meta-analysis has ever been carried out before. The second meta-analysis tested whether the population of video game players significantly differed from the population of non-video game players in terms of cognitive ability. The third meta-analysis dealt with the effects of video game training on cognitive ability. The first two meta-analyses represent an important (if not necessary) test for the hypothesis according to which video game training exerts a positive influence on cognitive ability. If video game experience/skill is not correlated with cognitive ability, or video game players are not better than non-video game players, then it is difficult to claim that video game training enhances cognitive ability.

4. General Method

4.1 Literature Search

A systematic search strategy was used to find the relevant studies (PRISMA statement; Moher, Liberati, Tetzlaff, & Altman, 2009). ERIC, PsycINFO, MEDLINE, JSTOR, Science Direct, and ProQuest Dissertation&Theses databases were searched to identify all the potentially relevant studies, using the following combination of keywords: ("video gam*" OR videogame) AND (intelligen* OR IQ OR "executive function*" OR percept* OR cognit* OR attention* OR visual* OR vision OR inhibition OR memory OR motor OR "dual task" OR "switching task" OR flanker OR "object tracking" OR "spatial"). Also, previous reviews were examined, and we e-mailed researchers in the field ($n = 135$) asking for inaccessible data.

4.2 Inclusion Criteria

The studies were included in accordance with the following four general criteria:

1. The variable of interest (e.g., video game experience, skill, and training) was successfully isolated. For example, studies reporting correlations and comparisons between treated and non-treated groups regarding the general use of digital media were excluded. Similarly, studies regarding the effects of video games involving physical training (i.e., exergames) on the participants' cognitive abilities were excluded;
2. During the study, at least one measure of domain-general cognitive ability non-related to video gaming was collected;
3. The participants of the study suffered from no specific learning disability (e.g., developmental dyslexia), behavioural disorder (e.g., aggressive behaviour), or clinical condition (e.g., video game addiction, amblyopia);
4. The data presented in the study, or provided by the authors, were sufficient to calculate an effect size.

The additional criteria for each of the three meta-analyses are reported in the three relevant Method sections.

To identify studies meeting these criteria, we searched for relevant published and unpublished articles until December 31st, 2016, and scanned reference lists. Forty-two authors replied to our e-mails. Twenty-five provided unpublished data.

We found 66 studies (Appendix D) reporting correlations between cognitive ability and video game skill, including 8,141 participants and 310 effect sizes. We found 98 studies (Appendix E) reporting comparisons (i.e., quasi-experimental design) between players and non-players, including 6,166 participants and 315 effect sizes. Finally, we found 63 studies (Appendix F)

regarding the effects of video game training on cognitive ability, including 3,286 participants and 359 effect sizes. The procedure is summarized in Figure 11.

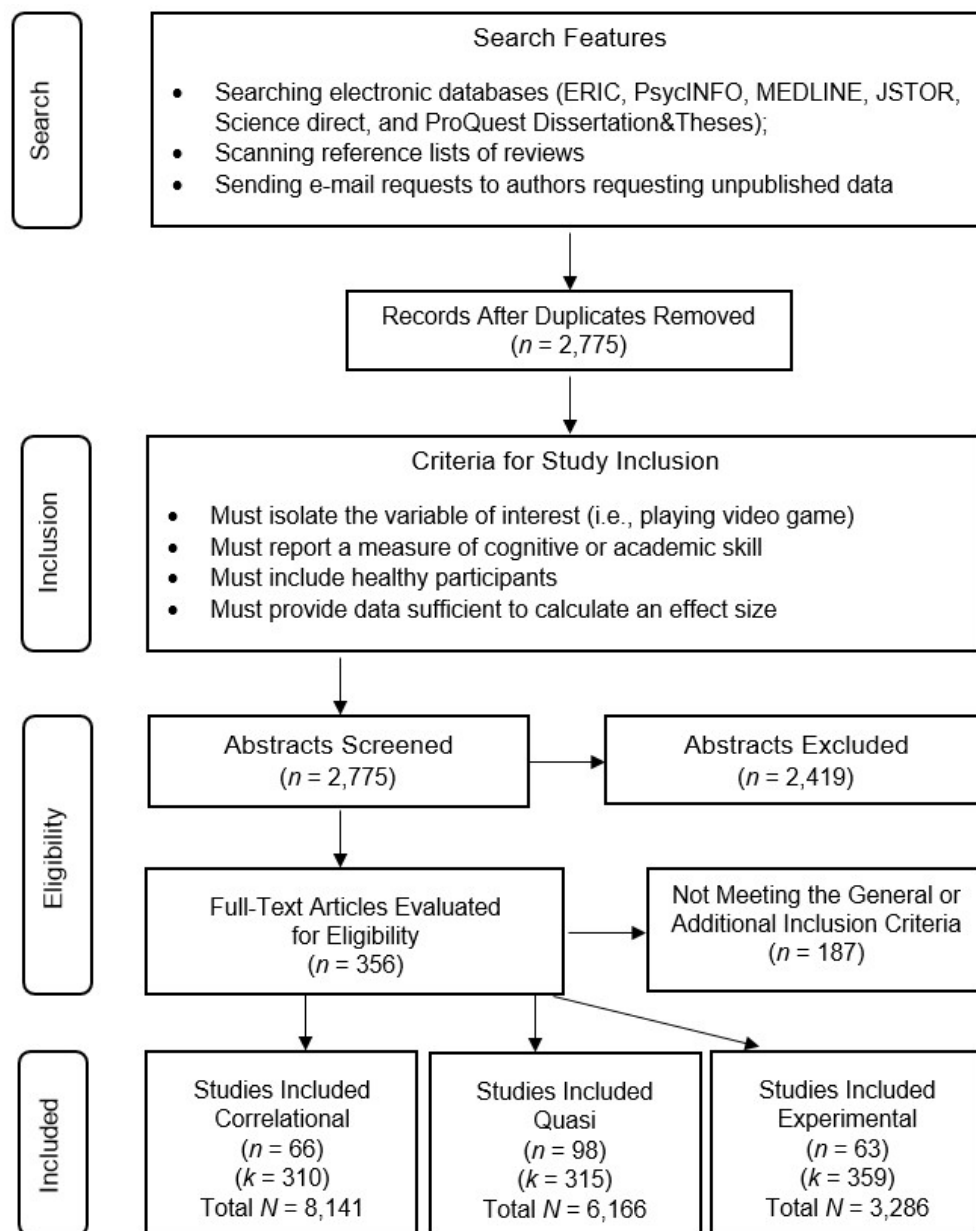


Figure 11. Flow diagram of the studies included in the meta-analyses.

4.3 Outcome Measures

The effect sizes were categorized into five broad measures: (a) *Visual attention/processing*, including all those tests measuring visual-perception skills (e.g., visual search tasks, flanker task, useful field of view [UFOV] tasks, and change detection tasks); (b) *Spatial ability*, including tests

such as mental rotation and folding tasks; (c) *Cognitive control*, including tests such as task switching, go/no-go, Simon, and Stroop tasks; (d) *Memory*, including tests such as span, *n*-back, and recall tasks; and (e) *Intelligence/reasoning*, including tests of fluid intelligence/reasoning (e.g., Raven's matrices) and comprehension knowledge (e.g., verbal fluency).

This categorization was used as the main moderator and named Outcome measure in all the three meta-analyses. When analysing the other categorical moderators, the effect sizes were sorted by Outcome measure, and the relative overall meta-analytic means were calculated.

The first author and the second author coded each effect size independently. The Cohen's kappa was $\kappa = .85$, 95% CI [.82; .88]. The authors resolved every discrepancy.

4.4 Statistical Dependence of the Samples

The effect sizes were calculated for each dependent variable reported in the studies. For each independent sample, those effect sizes referring to the same type of measure (e.g., reaction times) and extracted from the same test (e.g., different stimulus onset asynchronies in the UFOV task) were merged into one effect size. This procedure was used to calculate more reliable estimates and reduce the number of statistically dependent effect sizes in the model (Schmidt & Hunter, 2015). For those effect sizes that were statistically dependent and referred to different constructs or were extracted from different tests, Cheung and Chan's (2004) correction for statistically dependent samples was applied. This method decreased the weight of dependent samples in the analysis by calculating an adjusted (i.e., smaller) *N* in each meta-analytic model.

When the study presented multiple-group comparisons – for example, between one group (e.g., action video game players) and several comparison groups (e.g., non-video game players, non-action video game players) – we calculated as many effect sizes as the number of comparisons. Since Cheung and Chan's (2004) method cannot be used for partially dependent samples, we ran our analyses as if these effect sizes were statistically independent. This relatively minor limitation

was nearly absent when the effect sizes were sorted by type of video game. Thus, the entire procedure was a reliable way to deal with statistical dependent effect sizes and the related potential biases (e.g., overestimation of between-study variability).

4.5 Calculations of the Overall Meta-Analytic Means

Random-effect models were used to estimate the overall meta-analytic means. First, a model including all the effect sizes (main model) was run for each of the three meta-analyses. The overall meta-analytic means of the three main models represented a measure of the relationship between video game playing and overall cognitive ability. A series of meta-analytical sub-models were built to assess the effects of categorical moderators in all the three meta-analyses. To run the models, we used the Comprehensive Meta-Analysis (CMA; Version 3.3; Biostat, Englewood, NJ) software package.

4.6 Publication Bias Analysis

Publication bias occurs when non-significant results are systematically suppressed from the literature. This problem has been documented in the field of video gaming (e.g., Boot, Blakely, & Simons, 2011). Moreover, since the response's rate to our e-mails requesting for unpublished data was modest (25 positive responses out of 135 requests), a rigorous analysis of the effects of publication bias was imperative.

To investigate whether the results were affected by publication bias, we used Duval and Tweedie's (2000) trim-and-fill analysis and Vevea and Woods's (2005) selection model analysis. Trim-and-fill analysis estimates the symmetry of a funnel plot representing the relation between effect size and standard error. In the presence of publication bias, the trim-and-fill analysis estimates the number of missing studies from the funnel plot – either left or right of the meta-analytic mean – and imputes missing effect sizes based on the data's asymmetry to generate a more symmetrical funnel plot. CMA was used to perform trim-and-fill analyses.

Vevea and Woods's (2005) selection model analysis estimates four adjusted values by pre-weighted functions of p -values' distributions. If all (or most of) the four adjusted values are shown not to differ significantly from the meta-analytic mean, then it can be reliably concluded that the results are not affected by publication bias (Schmidt & Hunter, 2015). Also, this analysis stays reliable even when the number of effect sizes is modest. For this reason, only this publication bias analysis was run in those models that had fewer than 30 effect sizes. Finally, the trim-and-fill and selection model analyses can estimate adjusted values both smaller and greater than the meta-analytic mean. The Metafor software package (Viechtbauer, 2010) was used for conducting selection model analyses.

4.7 Influential Cases Analysis

Finally, to evaluate whether some effect sizes had an unusually large influence on the meta-analytic means, Viechtbauer and Cheung's (2010) influential cases analysis was performed in every meta-analytic model. Together with publication bias analysis, influential cases analysis was adopted to test the robustness of the overall results. The Metafor software package was used for conducting these analyses.

5. Meta-Analysis 1: Meta-Analysis of Correlational Data Among Video Game Players

Here, we report the first ever meta-analysis examining the relationship between video game skill and cognitive ability in video game players. As stated in the introduction, a positive correlation between video game skill and cognitive ability is a necessary, yet not sufficient, condition for the hypothesis according to which video game training exerts positive effects on cognitive ability. Also, the results of the present meta-analysis are a significant contribution to the study of the cognitive correlates of video game expertise.

5.1 Method

5.1.1 Additional inclusion criteria

The studies were included in the present meta-analysis when meeting the following two additional criteria:

1. The study provided information about how video game skill was assessed;
2. The participants had some experience of video games. For example, participants reporting zero hours of video game play per week were excluded.

5.1.2 Additional moderators

Along with Outcome measure, the effects of two additional moderators were analysed:

1. Skill measure (categorical moderator). This variable has two levels: (a) video game skill measured by the frequency of video game play (hours per week), and (b) video game skill measured by video game score obtained;
2. Type of video game (categorical moderator). This variable has three levels: (a) Action video games, (b) Non-action video games, and (c) Mixed video games. The category of Action video games refers to those video games classified as shooter (e.g., Unreal Tournament) and racing (e.g., Mario Kart) video games. The category of Non-action video games includes those video games that are not classifiable as action video games. Finally, the category of Mixed video games refers to general video game experience rather than the practice of a specific genre of video game. The first and the second authors coded each effect size independently. The Cohen's kappa was $\kappa = .96$, 95% CI [.94; .99]. The authors resolved every discrepancy.

5.1.3 Effect sizes

The correlations between video game skill and cognitive outcomes were taken from the data reported in the primary studies or calculated with the data provided by the authors. When group-

level comparisons (e.g., intermediates vs. experts) were reported ($k = 16$), we calculated point-biserial correlations. When the data were extrapolated from experimental studies with both pre- and post-test assessments, we used the correlations between performance on the cognitive test at the post-test assessment and either difference between post-test and pre-test video-game performance or, when provided, video-game post-test scores.²⁴ Finally, when possible, the samples were sorted by type of video game and gender.

5.2 Results

As described in the General Method section, we adopted a systematic approach to examine the correlation between video game skill and cognitive ability. First, we calculated the overall correlation with all the effect sizes. Then, we investigated the potential effects of the moderators and ran the relative sub-models. The robustness of the results of each model was tested with the abovementioned publication bias and influential case analyses.

5.2.1 Main model

A model comprising all the correlations was run. The random-effects meta-analytic overall correlation was $\bar{r} = .07$, 95% CI [.05; .09], $k = 310$, $p < .001$. The degree of heterogeneity between effect sizes was $I^2 = 52.19$,²⁵ suggesting that some moderators had a potential effect. The contour-

²⁴ The experimental studies included participants with no previous experience of the training video games. Thus, post-test scores and score gains between post- and pre-test scores were expected to be highly correlated and, therefore, equally valid measures of video game skills.

²⁵ The I^2 statistic refers to the percentage of between-study variance due to true heterogeneity and not to random error (Higgins, Thompson, Deeks, & Altman, 2003). The higher the value of the I^2 statistic, the higher the percentage of between-study variance due to true heterogeneity. A degree of heterogeneity (I^2) around 25.00 is considered low, around 50.00 moderate, and around 75.00 high (Higgins et al., 2003).

enhanced funnel plot (Peters, Sutton, Jones, Abrams, & Rushton, 2008) depicting the relation between effect size and standard error is shown in Figure 12.

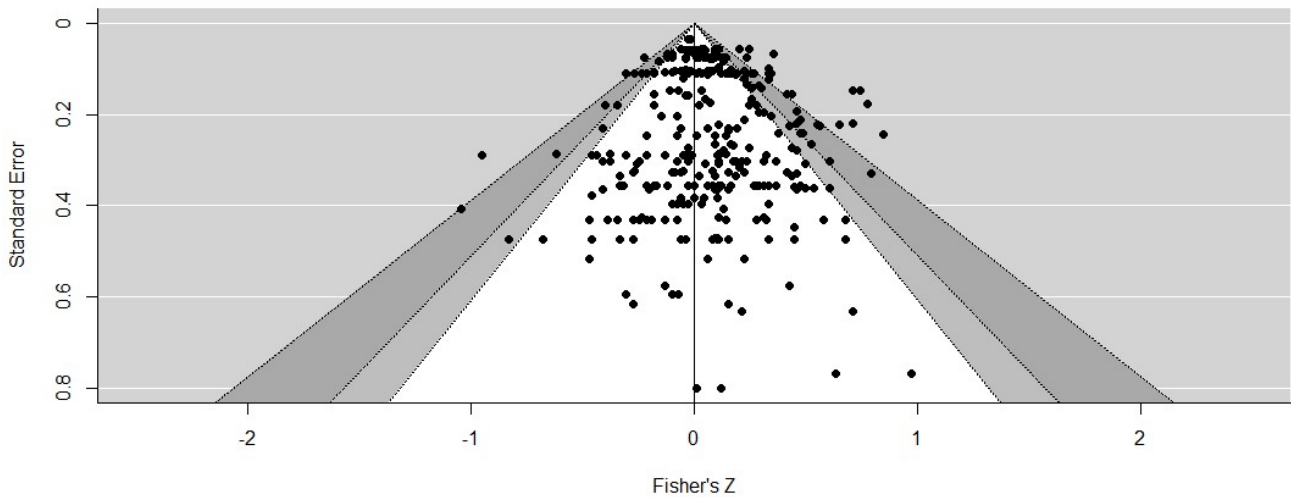


Figure 12. Contour-enhanced funnel plot of standard errors and effect sizes (Fisher's Zs) in the meta-analysis of the correlational data. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

The trim-and-fill analysis filled 16 studies left of the mean. The estimated correlation was $\bar{r} = .05$, 95% CI [.03; .08]. The estimates of the selection model analysis were $\bar{r} = .03$, $\bar{r} = .01$, $\bar{r} = .04$, and $\bar{r} = .04$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. The two publication bias analyses thus suggested that the overall correlation ($\bar{r} = .07$) was a slight overestimation.

Finally, Viechtbauer and Cheung's (2010) analysis detected four influential effect sizes. The overall correlation without these effect sizes was $\bar{r} = .06$, 95% CI [.04; .08], $k = 306$, $p < .001$, $I^2 = 41.08$. Therefore, the exclusion of the influential cases did not substantially alter the results.

5.2.2 Moderator analysis

Given the presence of some true heterogeneity in the main model, a meta-regression model including all the three moderators was run, $Q(7) = 58.68$, $k = 310$, $p < .001$.²⁶ Outcome measure and Skill measure were significant moderators ($p = .005$ and $p < .001$, respectively). Type of video game was marginally significant ($p = .059$).

We calculated the overall correlations of the five outcome measures. The results provided near-zero correlations in four measures and a small correlation ($\bar{r} = .18$) in spatial ability. The publication bias and influential case analyses did not evidence any substantial difference with the unadjusted correlations. The results are summarized in Table 9.

²⁶ Running separate analyses for each moderator does not control for potential interactions between moderators. Thus, when the power of the model is sufficient, including all the moderators in a single analysis should be preferred.

Table 9

Meta-Analytic and Publication Bias Results of the Main Model Sorted by Outcome Measure (Meta-Analysis 1)

| Outcome Measure | <i>Model</i> | | | | | | <i>Model without influential cases^a</i> | | | |
|-----------------------------|--------------|---------------------|-----------------|-----------------------|----------------------|----------------------------|----------------------------------------------------|---------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 122 | .06 [.03; .10] | .001 | 24.23 | .07 [.03; .11] | .04; .00; .06; .05; | 106 | .07 [.02; .11] | .003 | 11.52 |
| Spatial ability | 50 | .18 [.13; .23] | < .001 | 60.03 | .14 [.09; .19] | .14; .12; .14; .13; | 50 | .18 [.13; .23] | < .001 | 60.03 |
| Cognitive control | 38 | -.02 [-.09; .06] | .693 | 32.33 | -.11 [-.20; -.02] | -.06; -.10; -.03; -.02; | 33 | -.04 [-.14; .05] | .335 | .00 |
| Memory | 43 | .01 [-.03; .04] | .623 | .00 | .01 [-.03; .05] | -.01; -.04; .01; .01; | 39 | .02 [-.02; .07] | .317 | .00 |
| Intelligence/reasoning | 57 | .05 [.00; .10] | .033 | 75.43 | .00 [-.05; .06] | .00; -.02; .01; .01; | 54 | .01 [-.02; .04] | .536 | 29.12 |

Note. *k* = number of the effect sizes; \bar{r} = random effects meta-analytic overall correlation with 95% confidence intervals (in brackets); *p*-value of the meta-analytic overall correlation; *I*² = ratio of true heterogeneity; T&F = trim-and-fill estimates with 95% confidence intervals (in brackets); SM = moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection models estimates.

^a When no influential cases are found, the statistics are the same as in the uncorrected model.

5.2.2.1 Skill measure

To examine the effect of Skill measure further, two sub-models were run. The first sub-model comprised all the correlations between the outcome measures and video game skill measured by frequency of video game playing. The random-effects meta-analytic overall effect size was $\bar{r} = .03$, 95% CI [.00; .05], $k = 156$, $p = .024$. The degree of heterogeneity between effect sizes was $I^2 = 40.05$.

Trim-and-fill analysis filled four studies left of the mean. The estimated correlation was $\bar{r} = .02$, 95% CI [.00; .04]. The estimates of the selection model analysis were $\bar{r} = .01$, $\bar{r} = -.02$, $\bar{r} = .02$, and $\bar{r} = .01$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. The estimates of the publication bias analyses thus did not significantly differ from the overall correlation in this model.

Three influential effect sizes were detected. The overall correlation without these effect sizes was $\bar{r} = .02$, 95% CI [.00; .04], $k = 153$, $p = .048$, $I^2 = 21.56$. Therefore, the exclusion of the influential cases did not substantially alter the results.

We finally calculated the overall correlations sorted by Outcome measure. All the five overall correlations were close to zero. The publication bias and influential case analyses did not evidence any substantial difference with the unadjusted correlations. The results are summarized in Table 10.

Table 10

Meta-Analytic and Publication Bias Results Sorted by Outcome Measure, with Video Game Skill measured by Frequency of Video Game Playing (Meta-Analysis 1)

| Outcome Measure | <i>Model</i> | | | | | | <i>Model without influential cases</i> | | | |
|-----------------------------|--------------|---------------------|-----------------|-----------------------|--------------------|----------------------------|----------------------------------------|----------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 65 | .05 [.01; .09] | .006 | .00 | .05 [.01; .09] | .03; -.01; .04; .03; | 60 | .07 [.03; .11] | .001 | .00 |
| Spatial ability | 15 | .09 [.01; .17] | .028 | 72.55 | - | .10; .09; .11; .10; | 15 | .09 [.01; .17] | .028 | 72.55 |
| Cognitive control | 19 | -.06 [-.15; .03] | .212 | 38.94 | - | -.08; -.12; -.05; -.04; | 18 | -.09 [-.17; -.02] | .016 | 10.79 |
| Memory | 23 | .00 [-.04; .04] | .936 | .00 | - | -.02; -.04; .00; .00; | 23 | .00 [-.04; .04] | .936 | .00 |
| Intelligence/reasoning | 34 | .01 [-.02; .05] | .498 | 36.25 | .01 [-.03; .04] | -.01; -.03; .00; .00; | 32 | -.01 [-.04; .02] | .417 | .00 |

Note. See Note to Table 9 for abbreviations.

The same analysis was carried out for the correlation between cognitive ability and video game scores as measure of skill. A model comprising all the correlations between the outcome measures and video game skill measured with video game scores was run. The random-effects meta-analytic overall effect size was $\bar{r} = .16$, 95% CI [.12; .20], $k = 154$, $p < .001$. The degree of heterogeneity between effect sizes was $I^2 = 48.06$.

Trim-and-fill analysis filled 19 studies right of the mean. The estimated correlation was $\bar{r} = .21$, 95% CI [.16; .26]. The estimates of the selection model analysis were $\bar{r} = .12$, $\bar{r} = .09$, $\bar{r} = .13$, and $\bar{r} = .12$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. The two publication bias analyses thus provided a different pattern of results. All the estimated overall correlations were small but greater than zero.

Three influential effect sizes were detected. The overall correlation without these effect sizes was $\bar{r} = .17$, 95% CI [.13; .21], $k = 151$, $p < .001$, $I^2 = 34.88$. Therefore, the exclusion of the influential cases did not substantially alter the results.

We finally calculated the overall correlations sorted by Outcome measure. Three overall correlations (i.e., spatial ability, cognitive control, and intelligence/reasoning) appeared to be greater than the others. The influential case analysis showed that removing the influential case detected in the cognitive control model significantly lowered the estimated overall correlation (from $\bar{r} = .16$ to $\bar{r} = .07$; $p = .044$ and $p = .445$, respectively). Regarding the spatial ability overall correlation ($\bar{r} = .24$), the publication bias analyses calculated slightly smaller estimates (ranging between .15 and .18). Finally, the overall correlation between intelligence/reasoning ($\bar{r} = .14$) was found to be moderately underestimated (between .17 and .22, according to publication bias analysis). The results are summarized in Table 11.

Table 11

Meta-Analytic and Publication Bias Results Sorted by Outcome Measure, with Video Game Skill Measured by Video Game Scores (Meta-Analysis I)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|--------------------|-----------------|-----------------------|-------------------|-------------------------|---------------------------------|--------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 57 | .07 [-.01; .16] | .098 | 33.33 | .15 [.05; .24] | .08; .04; .10; .09; | 53 | .06 [-.03; .16] | .204 | 12.39 |
| Spatial ability | 35 | .24 [.18; .30] | < .001 | 43.40 | .18 [.12; .25] | .17; .15; .18; .16; | 35 | .24 [.18; .30] | < .001 | 43.40 |
| Cognitive control | 19 | .16 [.00; .30] | .044 | 4.51 | - | .10; .01; .14; .11; | 18 | .07 [-.11; .24] | .445 | .00 |
| Memory | 20 | .05 [-.04; .14] | .302 | .00 | - | .02; -.02; .04; .03; | 18 | .09 [-.10; .27] | .358 | .00 |
| Intelligence/reasoning | 23 | .14 [-.06; .33] | .162 | 63.75 | - | .21; .17; .22; .19; | 23 | .14 [-.06; .33] | .162 | 63.75 |

Note. See Note to Table 9 for abbreviations.

5.2.2.2 Type of video game

We carried out a set of analyses to examine the potential moderating role of type of video game. First, a model comprising all the correlations referring to action video games was run. The random-effects meta-analytic overall effect size was $\bar{r} = .11$, 95% CI [.06; .16], $k = 69$, $p < .001$. The degree of heterogeneity between effect sizes was $I^2 = 38.32$.

Trim-and-fill analysis filled seven studies right of the mean. The estimated correlation was $\bar{r} = .13$, 95% CI [.08; .18]. The estimates of the selection model analysis were $\bar{r} = .06$, $\bar{r} = .03$, $\bar{r} = .07$, and $\bar{r} = .06$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. All the estimates of the publication bias analyses were small (between $\bar{r} = .03$ and $\bar{r} = .13$) and thus did not substantially differ from the overall correlation in this model ($\bar{r} = .11$).

Four influential effect sizes were detected. The overall correlation without these effect sizes was $\bar{r} = .15$, 95% CI [.11; .19], $k = 65$, $p < .001$, $I^2 = 31.99$. Therefore, the exclusion of the influential cases showed that the overall correlation calculated for this model ($\bar{r} = .11$) might have been a moderate underestimation.

We finally calculated the overall correlations sorted by Outcome measure. The only overall correlation significantly different from zero was the one concerned with spatial ability ($\bar{r} = .30$). According to the publication bias analyses, this value was probably an overestimation (between $\bar{r} = .18$ and $\bar{r} = .26$). The results are summarized in Table 12.

Table 12

Meta-Analytic and Publication Bias Results Sorted by Outcome Measure, for the Correlations Referring to Action Video Games (Meta-Analysis 1)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|----------------------|-----------------|-----------------------|--------------------|-------------------------|---------------------------------|---------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 39 | .05 [-.01; .11] | .087 | .00 | .04 [-.02; .10] | .02; -.02; .04; .04; | 36 | .09 [.00; .18] | .058 | .00 |
| Spatial ability | 11 | .30 [.20; .39] | < .001 | 19.87 | - | .26; .24; .26; .24; | 10 | .32 [.23; .41] | < .001 | .00 |
| Cognitive control | 6 | -.17 [-.28; -.05] | .005 | .00 | - | - | 5 | -.07 [-.26; .13] | .505 | .00 |
| Memory | 7 | -.01 [-.08; .06] | .810 | .00 | - | - | 6 | .05 [-.08; .17] | .444 | .00 |
| Intelligence/reasoning | 6 | .12 [-.04; .26] | .142 | 50.85 | - | - | 5 | .21 [.08; .33] | .002 | .00 |

Note. See Note to Table 9 for abbreviations.

The same set of analyses was carried out for the non-action video games. We first ran model comprising all the correlations referring to non-action video games. The random-effects meta-analytic overall effect size was $\bar{r} = .07$, 95% CI [.04; .10], $k = 144$, $p < .001$. The degree of heterogeneity between effect sizes was $I^2 = 30.69$.

Trim-and-fill analysis filled nine studies left of the mean. The estimated correlation was $\bar{r} = .06$, 95% CI [.03; .10]. The estimates of the selection model analysis were $\bar{r} = .04$, $\bar{r} = .01$, $\bar{r} = .05$, and $\bar{r} = .04$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. All the estimates of the publication bias analyses were close to zero and did not substantially differ from the unadjusted overall correlation ($\bar{r} = .07$).

Twenty-five influential effect sizes were detected. The overall correlation without these effect sizes was $\bar{r} = .11$, 95% CI [.05; .16], $k = 119$, $p < .001$, $I^2 = 25.48$. Therefore, the exclusion of the influential cases moderately increased the overall correlation (from $\bar{r} = .07$ to $\bar{r} = .11$). In summary, all the estimated overall correlations ranged from $\bar{r} = .01$ to $\bar{r} = .11$.

We finally calculated the overall correlations sorted by Outcome measure. The overall correlation referring to spatial ability ($\bar{r} = .19$) was greater than the other ones (all smaller than $\bar{r} = .10$). The publication bias analyses provided significantly smaller estimates (between $\bar{r} = .06$ and $\bar{r} = .10$). The results are summarized in Table 13.

Table 13

Meta-Analytic and Publication Bias Results Sorted by Outcome Measure, for the Correlations Referring to Non-Action Video Games (Meta-Analysis 1)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|--------------------|-----------------|-----------------------|-------------------|-------------------------|---------------------------------|--------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 53 | .08 [.01; .14] | .019 | 33.14 | .09 [.02; .16] | .06; .03; .07; .06; | 43 | .09 [-.02; .20] | .096 | 11.70 |
| Spatial ability | 25 | .19 [.10; .27] | < .001 | 55.30 | - | .09; .06; .10; .08; | 25 | .19 [.10; .27] | < .001 | 55.30 |
| Cognitive control | 17 | .09 [-.09; .27] | .310 | .00 | - | .02; -.08; .08; .06; | 17 | .09 [-.09; .27] | .310 | .00 |
| Memory | 24 | .02 [-.03; .07] | .427 | .00 | - | .00; -.03; .02; .01; | 22 | .03 [-.03; .09] | .342 | .00 |
| Intelligence/reasoning | 25 | .02 [-.04; .08] | .500 | 17.30 | - | .00; -.03; .02; .01; | 22 | .07 [.00; .14] | .064 | .00 |

Note. See Note to Table 9 for abbreviations.

Finally, a model comprising all the correlations referring to mixed video games was run. The random-effects meta-analytic overall effect size was $\bar{r} = .04$, 95% CI [.01;0.08], $k = 97$, $p = .024$. The degree of heterogeneity between effect sizes was $I^2 = 69.22$.

Trim-and-fill analysis filled four studies right of the mean. The estimated correlation was $\bar{r} = .05$, 95% CI [.01; .09]. The estimates of the selection model analysis were $\bar{r} = .01$, $\bar{r} = -.01$, $\bar{r} = .02$, and $\bar{r} = .02$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. The estimates of the publication bias analyses thus did not significantly differ from the overall correlation in this model.

Four influential effect sizes were detected. The overall correlation without these effect sizes was $\bar{r} = .01$, 95% CI [-.01; .04], $k = 93$, $p = .332$, $I^2 = 37.38$. Therefore, the exclusion of the influential cases did not substantially alter the results.

We finally calculated the overall correlations sorted by Outcome measure. Four overall correlations were not significantly different from zero. The only exception was the small overall correlation referring to spatial ability ($\bar{r} = .11$). The publication bias and influential case analyses did not evidence any substantial difference with the unadjusted correlations. The results are summarized in Table 14.

Table 14

Meta-Analytic and Publication Bias Results Sorted by Outcome Measure, for the Correlations Referring to Mixed Video Games (Meta-Analysis 1)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|---------------------|-----------------|-----------------------|--------------------|----------------------------|---------------------------------|---------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{r} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 30 | .04 [-.02; .11] | .208 | 40.83 | .05 [-.02; .12] | .02; -.01; .04; .03; | 29 | .06 [.00; .13] | .067 | .00 |
| Spatial ability | 14 | .11 [.04; .18] | .004 | 66.21 | - | .12; .11; .13; .12; | 14 | .11 [.04; .18] | .004 | 66.21 |
| Cognitive control | 15 | .00 [-.12; .12] | .987 | 56.16 | - | -.02; -.06; .00; .00; | 15 | .00 [-.12; .12] | .987 | 56.16 |
| Memory | 12 | -.02 [-.10; .06] | .610 | .00 | - | -.10; -.21; -.03; -.02; | 11 | -.02 [-.13; .09] | .742 | .00 |
| Intelligence/reasoning | 26 | .06 [-.01; .13] | .082 | 90.01 | - | -.01; -.03; .00; .00; | 23 | -.02 [-.05; .01] | .147 | .00 |

Note. See Note to Table 9 for abbreviations.

5.3 Discussion

The main model and most of the sub-models showed weak correlations between video game skill and cognitive ability. For example, the overall correlations between video game skill and visual attention/processing measures are all smaller than .10. Similarly, none of the correlations regarding the measures cognitive control, memory, or intelligence/reasoning were greater than .16.

The only exception to this pattern of results was spatial ability. The overall correlations between video game skill and spatial ability were all significant with a range of values between .09 and .30. Given that $\bar{r} = .30$ is probably an overestimation (see publication bias estimates, Table 12), video game skill explains approximately between 1% and 6% of the variance in the participants' spatial ability. The correlation between video game skill and spatial ability, although limited in size, may represent a characteristic trait of video game expertise. In support of this hypothesis, the correlation between spatial ability and video game skill was stronger when one specific genre of video games was considered (i.e., action video games; Table 12).

As expected, the overall correlation was higher when video game skill was measured with scores rather than hours ($\bar{r} = .16$ and $\bar{r} = .03$, respectively). Score tends to be a more reliable measure of video game skill than the weekly frequency of play reported in a questionnaire. Thus, it is possible that the correlation between cognitive ability and hours of play per week was more affected by measurement error than the correlation between cognitive ability and video game scores.

Importantly, the influential cases analysis showed no substantial differences in the overall correlations between the models with and without influential effect sizes. Regarding the publication bias analysis, most of the corrected estimates were only slightly smaller (or, in

a few cases, greater) than the random-effects overall correlations. Thus, the results were robust. Overall, the results suggest that video game skill is not related or only weakly related to cognitive ability in general.

6. Meta-Analysis 2: Meta-Analysis of Quasi-Experimental Data

The meta-analysis of the correlational data showed little evidence of the cognitive benefits of playing video games. However, it is possible that such benefits occur regardless of skill, as long as individuals engage in video game playing. It is thus necessary to examine whether video game players outperform non-video game players in the cognitive measures examined above. Like in the previous meta-analysis, this condition is necessary, yet not sufficient, for the hypothesis that video game training positively impacts cognitive ability. Finally, like in the meta-analysis of the correlational data, the results of this meta-analysis will contribute to the research into the cognitive correlates of video game expertise.

6.1 Method

6.1.1 Additional inclusion criteria

The studies were included in the present meta-analysis when meeting the following two additional criteria:

1. The study provided clear information about how video game status was assessed (e.g., hours of play per week);²⁷

²⁷ Some studies did not explicitly report the precise cut-off point that was used (e.g., minimum 5 hours of video game playing per week in order to be included as a video game player) but rather referred to the criterion used in a previous study such as Green and Bavelier (2003).

2. The study compared participants with experience of video game playing (in general or in a specific genre of video game) with participants with negligible or null experience in video game playing (in general, or in that specific genre of video game).

6.1.2 Additional moderators

Along with the Outcome measure, the effects of one additional moderator were analysed:

1. Type of video game (categorical moderator). This variable has three levels: (a) Action video game, (b) Non-action video game, and (c) Mixed video game. Action video game refers to the comparisons between action video game (shooter and racing) players vs. non-action video game players and non-video game players.²⁸ The category of Non-action video game includes the comparisons between non-action video game players and non-players. Finally, the category of Mixed video game refers to the comparisons between video game players (with no specific genre specialization) and non-video game players. The first and the second authors coded each effect size independently. The Cohen's kappa was $\kappa = .95$, 95% CI [.92; .99]. The authors resolved every discrepancy.

6.1.3 Effect sizes

We calculated the standardized mean difference (i.e., Cohen's d) between the two groups with the following formula:

$$d = (M_e - M_c) / SD_{pooled} \quad (1)$$

²⁸ Most of the studies involving action video game players adopted Green and Bavelier's (2003) criterion (see Footnote 27). Thus, action video game players were compared to non-action video game players without distinguishing between non-players and players of non-action video games.

where SD_{pooled} is the pooled standard deviation, and M_e and M_c are the means of the experimental group (i.e., video game players) and the control group (i.e., non-video game players), respectively. When t - or F -values were provided, we used CMA to convert them into Cohen's ds . Finally, to correct the effect sizes for upward bias, CMA was used to convert Cohen's ds into Hedges's gs (Hedges & Olkin, 1985).

6.2 Results

The systematic approach described in the General Method section was adopted to examine the difference between video game players and non-video game players in terms of cognitive ability. First, we calculated the overall effect sizes including all the effects. Then, we investigated the potential effects of the moderators and ran the relative sub-models. The robustness of the results of each model was tested with the two publication bias analyses and Viechtbauer and Cheung's (2010) influential case analysis.

6.2.1 Main model

In the model comprising all the effect sizes, the random-effects meta-analytic overall effect size was $\bar{g} = 0.33$, 95% CI [0.28; 0.39], $k = 315$, $p < .001$. The degree of heterogeneity between effect sizes was low, $I^2 = 33.79$. The contour-enhanced funnel plot is shown in Figure 13.

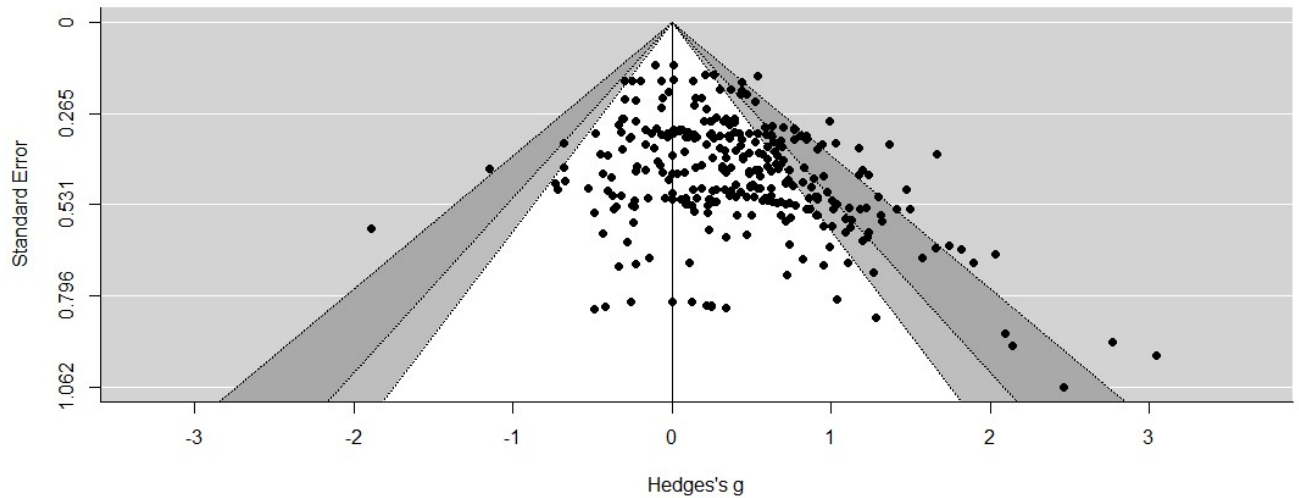


Figure 13. Contour-enhanced funnel plot of standard errors and effect sizes (g s) in the meta-analysis of the quasi-experimental data. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

The trim-and-fill analysis filled 73 studies left of the mean. The estimated effect size was $\bar{g} = 0.18$, 95% CI [0.12; 0.24]. The estimates of the selection model analysis were $\bar{g} = 0.24$, $\bar{g} = 0.17$, $\bar{g} = 0.27$, and $\bar{g} = 0.23$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. The two publication bias analyses thus suggested that the unadjusted overall effect size ($\bar{g} = 0.33$) was significantly inflated by the suppression from the literature of several smaller-than-average effect sizes.

Finally, Viechtbauer and Cheung's (2010) analysis detected two influential effect sizes. The overall effect size without these effect sizes was $\bar{g} = 0.32$, 95% CI [0.27; 0.38], $k = 313$, $p < .001$, $I^2 = 30.59$. Therefore, the exclusion of the influential cases did not substantially alter the results.

6.2.2 Moderator analysis

A meta-regression model including the two moderators was run, $Q(6) = 13.72$, $k = 315$, $p = .033$. Neither Outcome measure, nor Type of game was significant ($p = .174$ and $p = .101$, respectively).

The overall effect sizes were calculated for the five outcome measures. The results showed small to medium²⁹ effect sizes in all the measures. The influential case analyses did not highlight any substantial difference with the unadjusted overall effect sizes. By contrast, the estimates provided by the publication bias analyses were systematically smaller than the unadjusted values. This pattern of results was particularly evident in the visual attention/processing- and memory-related measures. The results are summarized in Table 15.

²⁹ According to Cohen's (1988) categorization, effect sizes of 0.20, 0.50, and 0.80 are considered small, medium, and large, respectively.

Table 15

Meta-Analytic and Publication Bias Results of the Main Model Sorted by Outcome Measure (Meta-Analysis 2)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|----------------------|-----------------|-----------------------|------------------------|----------------------------|---------------------------------|----------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 186 | 0.41 [0.33; 0.49] | < .001 | 32.89 | 0.27 [0.19; 0.35] | 0.34; 0.26; 0.36; 0.31; | 184 | 0.39 [0.32; 0.47] | < .001 | 28.15 |
| Spatial ability | 28 | 0.24 [0.13; 0.34] | < .001 | 25.33 | - | 0.18; 0.13; 0.20; 0.18; | 28 | 0.24 [0.13; 0.34] | < .001 | 25.33 |
| Cognitive control | 53 | 0.24 [0.12; 0.36] | < .001 | 25.68 | 0.17 [0.04; 0.30] | 0.18; 0.12; 0.21; 0.18; | 53 | 0.24 [0.12; 0.36] | < .001 | 25.68 |
| Memory | 32 | 0.20 [0.03; 0.37] | .019 | 45.14 | -0.04 [-0.23; 0.15] | 0.09; 0.01; 0.13; 0.10; | 32 | .20 [0.03; 0.37] | .019 | 45.14 |
| Intelligence/reasoning | 16 | 0.19 [0.00; 0.38] | .055 | 7.37 | - | 0.13; 0.05; 0.16; 0.13; | 16 | 0.19 [0.00; 0.38] | .055 | 7.37 |

Note. k = number of the effect sizes; \bar{g} = random effects meta-analytic mean with 95% confidence intervals (in brackets); p -value of the meta-analytic overall effect size; I^2 = ratio of true heterogeneity; T&F = trim-and-fill estimates with 95% confidence intervals (in brackets); SM = moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection models estimates.

6.2.2.1 Type of video game

Like in the meta-analysis regarding the correlational data, we ran a series of analyses to examine the potential moderating role of type of video game. First, a sub-model comprising all the effect sizes referring to action video games was run. The random-effects meta-analytic overall effect size was $\bar{g} = 0.40$, 95% CI [0.33; 0.47], $k = 199$, $p < .001$. The degree of heterogeneity between effect sizes was $I^2 = 33.10$.

Trim-and-fill analysis filled 38 studies left of the mean. The estimated overall effect size was $\bar{g} = 0.26$, 95% CI [0.18; 0.34]. The estimates of the selection model analysis were $\bar{g} = 0.34$, $\bar{g} = 0.26$, $\bar{g} = 0.37$, and $\bar{g} = 0.31$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. Thus, the publication bias analyses suggested that the unadjusted overall effect size ($\bar{g} = 0.40$) was an overestimation.

Two influential effect sizes were detected. The overall effect size without these effect sizes was $\bar{g} = 0.38$, 95% CI [0.31; 0.45], $k = 197$, $p < .001$, $I^2 = 28.37$. Therefore, the exclusion of the influential cases did not substantially alter the results.

We finally calculated the overall effect sizes sorted by Outcome measure. Four measures provided statistically significant and small-to-medium overall effect sizes (the only exception was intelligence/reasoning). Viechtbauer and Cheung's (2010) influential case analysis evidenced no significant differences between adjusted and unadjusted values. The publication bias analyses estimated smaller overall effect sizes in all the measures. The results are summarized in Table 16.

Table 16

Meta-Analytic and Publication Bias Results Sorted by Outcome Measure, for the Effect Sizes Referring to Action Video Games (Meta-Analysis 2)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|-----------------------|-----------------|-----------------------|-----------------------|-----------------------------|---------------------------------|-----------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 132 | 0.45 [0.36; 0.54] | < .001 | 33.10 | 0.27 [0.17; 0.38] | 0.38; 0.30; 0.41; 0.34; | 131 | 0.43 [0.34; 0.52] | < .001 | 29.52 |
| Spatial ability | 8 | 0.47 [0.21; 0.74] | .001 | .00 | - | 0.43; 0.36; 0.44; 0.38; | 8 | 0.47 [0.21; 0.74] | .001 | .00 |
| Cognitive control | 33 | 0.27 [0.09; 0.46] | .004 | 28.07 | 0.18 [-0.01; 0.38] | 0.19; 0.10; 0.23; 0.19; | 33 | 0.27 [0.09; 0.46] | .004 | 28.07 |
| Memory | 17 | 0.31 [0.06; 0.57] | .017 | 50.93 | - | 0.22; 0.14; 0.25; 0.21; | 17 | 0.31 [0.06; 0.57] | .017 | 50.93 |
| Intelligence/reasoning | 9 | 0.17 [-0.21; 0.54] | .377 | 51.92 | - | 0.09; -0.01; 0.14; 0.11; | 9 | 0.17 [-0.21; 0.54] | .377 | 51.92 |

Note. See Note to Table 15 for abbreviations.

Second, a model comprising all the effect sizes concerned with non-action video games was run. The random-effects meta-analytic overall effect size was $\bar{g} = 0.33$, 95% CI [0.11; 0.55], $k = 14$, $p = .003$. The degree of heterogeneity between effect sizes was $I^2 = .00$.

The estimates of the selection model analysis were $\bar{g} = 0.27$, $\bar{g} = 0.19$, $\bar{g} = 0.30$, and $\bar{g} = 0.25$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. No outlier was detected. Due to the scarcity of the effect sizes, no sub-models of outcome measures were run.

Finally, a systematic set of analyses was carried out for mixed video games. For the model comprising all the effect sizes referring to mixed video games, the random-effects meta-analytic overall effect size was $\bar{g} = 0.23$, 95% CI [0.15; 0.31], $k = 102$, $p < .001$. The degree of heterogeneity between effect sizes was $I^2 = 33.15$.

Trim-and-fill analysis filled 25 studies left of the mean. The estimated overall effect size was $\bar{g} = 0.12$, 95% CI [0.04; 0.20]. The estimates of the selection model analysis were $\bar{g} = 0.15$, $\bar{g} = 0.07$, $\bar{g} = 0.17$, and $\bar{g} = 0.14$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. Thus, the publication bias analyses once again suggested that the unadjusted overall effect size ($\bar{g} = 0.23$) was an overestimation. Only one influential effect size was detected. The overall effect size without this effect size was $\bar{g} = 0.22$, 95% CI [0.14; 0.30], $k = 101$, $p < .001$, $I^2 = 29.87$.

We finally calculated the overall effect sizes sorted by Outcome measure. All the overall effect sizes were small (see Table 17). The influential case analysis evidenced no significant differences between adjusted and unadjusted values. The publication bias analyses estimated smaller overall effect sizes in all the measures.

Table 17

Meta-Analytic and Publication Bias Results Sorted by Outcome Measure, for the Effect Sizes Referring to Mixed Video Games (Meta-Analysis 2)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|--------------------|-----------------|-----------------------|-------------------|--------------------------|---------------------------------|---------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 44 | .33 [.18; .47] | < .001 | 35.35 | .21 [.05; .36] | .25; .18; .28; .23; | 44 | .33 [.18; .47] | < .001 | 35.35 |
| Spatial ability | 16 | .21 [.06; .36] | .005 | 37.63 | - | .13; .08; .15; .12; | 16 | .21 [.06; .36] | .005 | 37.63 |
| Cognitive control | 20 | .21 [.06; .36] | .006 | 30.14 | - | .18; .13; .20; .17; | 18 | .20 [.07; .34] | .002 | .00 |
| Memory | 15 | .06 [-.13; .26] | .523 | 36.25 | - | -.02; -.10; .03; .02; | 14 | -.05 [-.21; .12] | .578 | 1.50 |
| Intelligence/reasoning | 7 | .20 [.00; .40] | .052 | .00 | - | .15; .09; .18; .15; | 7 | .20 [.00; .40] | .052 | .00 |

Note. See Note to Table 15 for abbreviations.

6.3 Discussion

The overall effect sizes of the main model and sub-models showed small to medium effect sizes, indicating that video game players outperformed non-players in all the five broad measures of cognitive ability. This superiority occurred regardless of the type of game considered. However, the publication bias analysis calculated a reduced estimate for many of the greatest overall effect sizes. Most of these corrected overall effect sizes remained significant or marginally significant. The influential cases analysis did not meaningfully modify the overall effect sizes.

Overall, the results suggest that video game players do differ from non-video game players in terms of cognitive ability. Nonetheless, the size of the effects is substantially smaller than the ones reported in Powers et al. (2013). Although quasi-experiments do not allow any strong inference with respect to causality, the outcomes of this meta-analysis suggest that engagement in video games exerts some modest effects on overall cognitive ability. However, the results do not exclude the possibility that individuals with higher cognitive abilities are more likely to play videogames. If so, no causal relationship between playing video games and superior cognitive abilities needs to be postulated to account for these results.

7. Meta-Analysis 3: Meta-Analysis of Experimental Training Data

Overall, the two previous meta-analyses provided weak evidence in favour of the hypothesis according to which playing video games enhances cognitive ability. This hypothesis, however, cannot be properly evaluated without testing it directly. This meta-analysis thus examines the effects of video game interventions on participants' cognitive ability.

7.1 Method

7.1.1 Additional inclusion criteria

Studies were included in the present meta-analysis when meeting the following additional criteria:

1. The study included at least one control group;
2. The study included participants with no (or negligible) experience, at the beginning of the experiment, in the video game(s) used during training;
3. The training video game was not purposely designed to improve cognitive ability (e.g., Lumosity® brain-training video games).³⁰

7.1.2 Additional moderators

Along with Outcome measure, the effects of four additional moderators were analysed:

1. Random allocation (dichotomous moderator): whether participants were allocated or not to the groups by randomization. This moderator was included to control for potential confounding effects due to differences at baseline level.
2. Hours of training (continuous moderator): the duration of training in hours;
3. Type of video game (categorical moderator). This variable has three levels: (a) Action vs. non-action video game players, where the action video game training (e.g., Unreal Tournament) was compared with an active control group training in a non-action video game (e.g., The Sims); (b) Action video game training, where the action video game training was compared with a control group not engaged in video game playing; and (c) Non-action video game training, where the non-action video game training (e.g., Tetris) was compared with a control group not engaged in video game playing. A small group of effect sizes ($k = 16$) from four studies did not fit any of the above categories and were excluded from the analyses regarding Type of video game. The first and the second authors coded each effect size independently. No discrepancies were found;³¹

³⁰ For a systematic review of the effects of brain-training programs, see Simons et al. (2016).

³¹ No moderator distinguishing between active and passive control groups was included. In most of the cases, active control groups consisted of people playing another type of video game. Thus,

4. Age (categorical moderator). This variable has three levels: (a) Adult, where the participants were aged 18 to 55; (b) Old, where the participants were older than 55; and (c) Young, where the participants were younger than 18.

7.1.3 Effect sizes

We calculated the standardized mean difference (i.e., Cohen's d) between the two groups with the following formula:

$$d = (M_{g-e} - M_{g-c}) / SD_{pooled-pre} \quad (2)$$

where $SD_{pooled-pre}$ is the pooled standard deviation of the two pre-test standard deviations, and M_{g-e} and M_{g-c} are the gain of the experimental group and the control group, respectively (Schmidt & Hunter, 2015, p. 353). When means and standard deviations were not available, t - or F -values³² were converted into Cohen's d s with CMA. Finally, to correct the effect sizes for upward bias, CMA was used to convert Cohen's d s into Hedges's g s (Hedges & Olkin, 1985).

7.2 Results

A set of analyses was run to investigate whether video game training provided any benefit for the participants' cognitive ability. Like in the two previous meta-analyses, we first calculated the overall effect sizes including all the effects. Then, we examined the potential effects of the

running models sorted by the type of control group (i.e., active or passive) would substantially duplicate the results of the moderator Type of video game.

³² The t - and F -statistics referring to pre-post improvements within groups were converted to d s and then subtracted to calculate the standardized mean difference between the experimental and control groups. Alternatively, the statistics referring to between-group differences at pre- and post-tests were converted to d s and then subtracted. The statistics referring to interactions between group and others conditions were excluded.

moderators and ran the relative sub-models. The robustness of the results of each model was tested with the two publication bias analyses and Viechtbauer and Cheung's (2010) influential case analysis.

7.2.1 Main model

The random-effects meta-analytic overall effect size was $\bar{g} = 0.07$, 95% CI [0.02; 0.12], $k = 359$, $p = .004$. The degree of heterogeneity between effect sizes was $I^2 = 17.90$. The contour-enhanced funnel plot is shown in Figure 14.

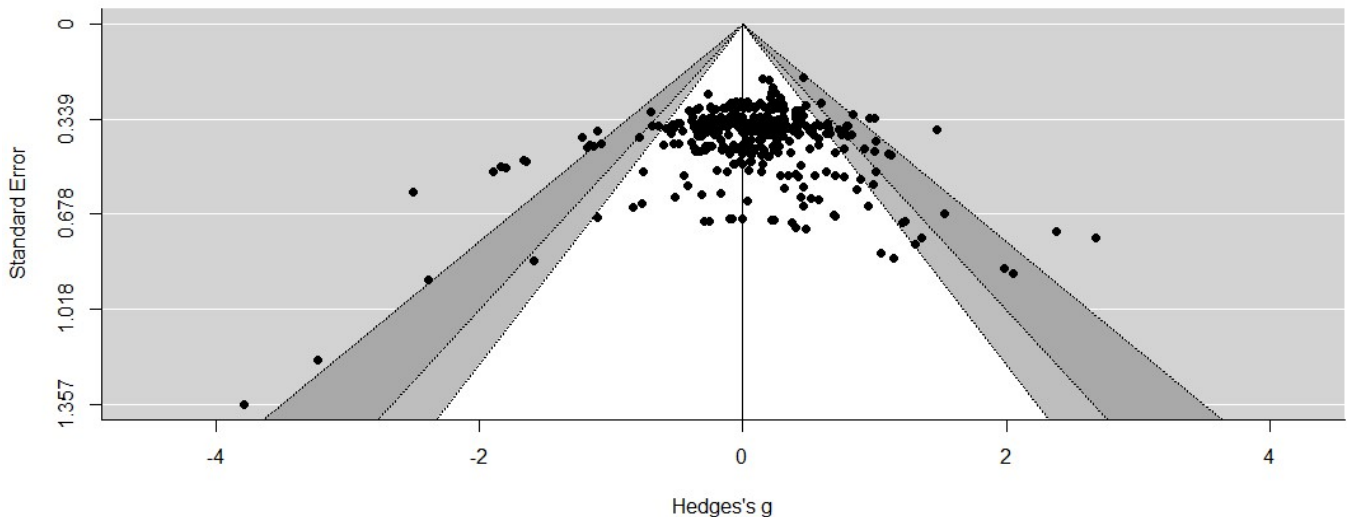


Figure 14. Contour-enhanced funnel plot of standard errors and effect sizes (g s) in the meta-analysis of the experimental data. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

The two publication bias analyses lowered the already small effect size further. The trim-and-fill analysis filled 35 studies left of the mean. The estimated effect size was $\bar{g} = 0.00$, 95% CI [-0.05; 0.05]. The estimates of the selection model analysis were $\bar{g} = 0.00$, $\bar{g} = -0.11$, $\bar{g} = 0.06$, and $\bar{g} = 0.05$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively.

Viechtbauer and Cheung's (2010) analysis detected three influential effect sizes. The overall effect size without these effect sizes was unaltered, $\bar{g} = 0.07$, 95% CI [0.03; 0.12], $k = 356$, $p = .002$, $I^2 = 10.64$.

7.2.2 Moderator analysis

A meta-regression model including all the five moderators was run, $Q(10) = 34.13$, $k = 341$,³³ $p < .001$. In line with the small degree of heterogeneity, the effect of the moderators was modest. Random allocation and hours of training were not significant moderators, $p = .273$ and $p = .927$, respectively. Outcome measure and Age were marginally significant, $p = .058$ and $p = .068$, respectively. Type of video game was the only significant moderator, $p = .006$.

Similar to the other two meta-analyses, we calculated overall effect sizes for the five outcome measures. The results showed null or small effect sizes in all the measures. No substantial difference emerged from the influential case and publication bias analyses. The results are summarized in Table 18.

³³ Most of the missing values ($k = 16$) in the model were due to the moderator Type of video game. The remaining three missing values came from other moderators. Given the small percentage of missing values (about 5%), the results of this moderator analysis can be considered highly reliable.

Table 18

Meta-Analytic and Publication Bias Results of the Main Model Sorted by Outcome Measure (Meta-Analysis 3)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|------------------------|-----------------|-----------------------|------------------------|--------------------------------|---------------------------------|------------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 131 | 0.09 [0.01; 0.18] | .033 | 3.23 | -0.03 [-0.13; 0.06] | 0.01; -0.11; 0.08; 0.06; | 131 | 0.09 [0.01; 0.18] | .033 | 3.23 |
| Spatial ability | 75 | 0.14 [0.05; 0.22] | .002 | .00 | 0.14 [0.05; 0.22] | 0.07; -0.02; 0.13; 0.10; | 73 | 0.13 [0.04; 0.23] | .004 | .00 |
| Cognitive control | 55 | 0.02 [-0.12; 0.17] | .738 | 27.04 | 0.14 [-0.02; 0.30] | -0.04; -0.16; 0.03; 0.02; | 55 | 0.02 [-0.12; 0.17] | .738 | 27.04 |
| Memory | 67 | 0.13 [0.03; 0.22] | .010 | .00 | 0.22 [0.11; 0.33] | 0.05; -0.06; 0.11; 0.09; | 67 | 0.13 [0.03; 0.22] | .010 | .00 |
| Intelligence/reasoning | 31 | -0.14 [-0.36; 0.08] | .206 | 55.62 | -0.18 [-0.40; 0.04] | -0.15; -0.27; -0.07; -0.06; | 29 | -0.02 [-0.21; 0.16] | .799 | 31.87 |

Note. k = number of the effect sizes; \bar{g} = random effects meta-analytic mean with 95% confidence intervals (in brackets); p -value of the meta-analytic overall effect size; I^2 = ratio of true heterogeneity; T&F = trim-and-fill estimates with 95% confidence intervals (in brackets); SM = moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection models estimates.

7.2.2.1 Type of video game

We analysed this moderator to test the potential differences between types of video game training. A sub-model comprising all the effect sizes referring to action vs. non-action video game players was run. The random-effects meta-analytic overall effect size was $\bar{g} = 0.10$, 95% CI [0.00; 0.21], $k = 94$, $p = .056$. The degree of heterogeneity between effect sizes was $I^2 = 18.47$.

The publication bias analyses once again showed that the effect size was inflated. Trim-and-fill analysis filled 17 studies left of the mean. The estimated overall effect size was $\bar{g} = -0.01$, 95% CI [-0.13; 0.10]. The estimates of the selection model analysis were $\bar{g} = 0.00$, $\bar{g} = -0.11$, $\bar{g} = 0.06$, and $\bar{g} = 0.05$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. No influential effect sizes were detected.

We finally calculated the overall effect sizes sorted by Outcome measure. All the overall effect sizes were small or null. While the influential case analysis detected no outliers, the publication bias analyses estimated moderately smaller overall effect sizes in all the measures. The results are summarized in Table 19.

Table 19

Meta-Analytic and Publication Bias Results Sorted by Outcome Measure, for the Effect Sizes Referring to Action Video Game Players vs. Non-Action Video Game Players (Meta-Analysis 3)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|------------------------|-----------------|-----------------------|-----------------------|--------------------------------|---------------------------------|-----------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 50 | 0.23 [0.05; 0.40] | .011 | 36.71 | 0.14 [-0.04; 0.32] | 0.08; -0.03; 0.14; 0.12; | 50 | 0.23 [0.05; 0.40] | .011 | 36.71 |
| Spatial ability | 16 | 0.02 [-0.20; 0.25] | .833 | 19.19 | - | -0.06; -0.18; 0.01; 0.01; | 16 | 0.02 [-0.20; 0.25] | .833 | 19.19 |
| Cognitive control | 17 | -0.03 [-0.29; 0.23] | .820 | 14.32 | - | -0.13; -0.29; -0.03; -0.03; | 17 | -0.03 [0.05; 0.40] | .820 | 14.32 |
| Memory | 11 | 0.11 [-0.12; 0.34] | .346 | .00 | - | 0.04; -0.07; 0.10; 0.08; | 11 | 0.11 [-0.12; 0.34] | .346 | .00 |
| Intelligence/reasoning | 0 | - | - | - | - | - | 0 | - | - | - |

Note. See Note to Table 18 for abbreviations.

The previous sub-model examined the effects of action video game training compared to non-action video game training. We now consider the comparison action video game players vs. non-video game players. In the sub-model comprising all effect sizes, the random-effects meta-analytic overall effect size was $\bar{g} = -0.12$, 95% CI [-0.24; 0.01], $k = 90$, $p = .062$. The degree of heterogeneity between effect sizes was $I^2 = 46.35$.

Trim-and-fill analysis filled 11 studies left of the mean. The estimated overall effect size was $\bar{g} = -0.26$, 95% CI [-0.38; -0.13]. The estimates of the selection model analysis were $\bar{g} = -0.22$, $\bar{g} = -0.38$, $\bar{g} = -0.13$, and $\bar{g} = -0.11$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. These small negative estimated values were probably statistical artefacts.

Only one influential effect size was detected. The overall effect size without these effect sizes was $\bar{g} = -0.10$, 95% CI [-0.22; 0.02], $k = 89$, $p = .107$, $I^2 = 40.84$.

Finally, the overall effect sizes sorted by Outcome measure were calculated. Four overall effect sizes were small or null. The only exception was the large negative overall effect size referring to intelligence/reasoning measures. However, due to the small number of effect sizes ($k = 8$), this overall effect size is not a reliable estimate. The influential case analysis detected one outlier in the cognitive control and memory models. The adjusted values were significantly closer to zero compared to the negative unadjusted values. The results are summarized in Table 20.

Table 20

Meta-Analytic and Publication Bias Results Sorted by Outcome Measure, for the Effect Sizes Referring to Action Video Game Players vs. Non-Video Game Players (Meta-Analysis 3)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|-------------------------|-----------------|-----------------------|---------------------|--------------------------------|---------------------------------|-------------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 39 | 0.00 [-0.13; 0.14] | .973 | 3.09 | -.12 [-.27; .22] | -0.08; -0.22; -0.01; 0.00; | 39 | 0.00 [-0.13; 0.14] | .973 | 3.09 |
| Spatial ability | 20 | 0.07 [-0.14; 0.28] | .521 | .00 | - | -0.02; -0.16; 0.07; 0.05; | 20 | 0.07 [-0.14; 0.28] | .521 | .00 |
| Cognitive control | 11 | -0.27 [-0.72; 0.17] | .230 | 62.77 | - | -0.33; -0.50; -0.23; -0.19; | 10 | -0.10 [-0.46; 0.25] | .568 | 35.87 |
| Memory | 12 | -0.10 [-0.43; 0.22] | .541 | 38.56 | - | 0.07; 0.07; 0.07; 0.07; | 11 | 0.02 [-0.24; 0.27] | .902 | .00 |
| Intelligence/reasoning | 8 | -1.17 [-1.83; -0.51] | .001 | 69.33 | - | -1.05; -1.17; -1.01; -0.96; | 8 | -1.17 [-1.83; -0.51] | .001 | 69.33 |

Note. See Note to Table 18 for abbreviations.

The third sub-model of this moderator analysis comprised all the effect sizes referring to non-action video game players vs. non-video game players. The random-effects meta-analytic overall effect size was $\bar{g} = 0.13$, 95% CI [0.07; 0.18], $k = 159$, $p < .001$. The degree of heterogeneity between effect sizes was $I^2 = .00$.

Trim-and-fill analysis filled 11 studies right of the mean. The estimated overall effect size was $\bar{g} = 0.17$, 95% CI [0.11; 0.23]. The estimates of the selection model analysis were $\bar{g} = 0.08$, $\bar{g} = -0.01$, $\bar{g} = 0.12$, and $\bar{g} = 0.10$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. The estimates of the publication bias analyses thus did not substantially differ from the unadjusted overall effect size.

Two influential effect sizes were detected. The overall effect size without these effect sizes was $\bar{g} = 0.12$, 95% CI [0.07; 0.17], $k = 157$, $p < .001$, $I^2 = .00$.

Finally, the overall effect sizes sorted by Outcome measure were calculated. All the overall effect sizes were small. The influential case and publication bias analyses had no substantial impact on the estimated values. The results are summarized in Table 21.

Table 21

Meta-Analytic and Publication Bias Results Sorted by Outcome Measure, for the Effect Sizes Referring to Non-Action Video Game Players vs. Non-Video Game Payers (Meta-Analysis 3)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|-----------------------|-----------------|-----------------------|----------------------|-----------------------------|---------------------------------|-----------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 38 | 0.06 [-0.07; 0.18] | .379 | .00 | 0.12 [0.00; 0.24] | 0.02; -0.09; 0.08; 0.06; | 38 | 0.06 [-0.07; 0.18] | .379 | .00 |
| Spatial ability | 39 | 0.20 [0.09; 0.31] | < .001 | .00 | 0.20 [0.09; 0.31] | 0.14; 0.06; 0.18; 0.15; | 37 | 0.21 [0.09; 0.33] | .001 | .00 |
| Cognitive control | 22 | 0.14 [-0.04; 0.32] | .130 | 27.34 | - | 0.09; 0.00; 0.14; 0.11; | 22 | 0.14 [-0.04; 0.32] | .130 | 27.34 |
| Memory | 40 | 0.17 [0.05; 0.29] | .006 | .00 | 0.22 [0.10; 0.34] | 0.16; 0.16; 0.16; 0.16; | 40 | 0.17 [0.05; 0.29] | .006 | .00 |
| Intelligence/reasoning | 20 | 0.08 [-0.09; 0.25] | .362 | .89 | - | 0.01; -0.10; 0.07; 0.05; | 20 | 0.08 [-0.09; 0.25] | .362 | .89 |

Note. See Note to Table 18 for abbreviations.

7.2.2.2 Age

This section investigates the potential moderating role of Age. First, we examined adult video game players. In the sub-model comprising all the effect sizes, the random-effects meta-analytic overall effect size was $\bar{g} = 0.10$, 95% CI [0.05; 0.15], $k = 239$, $p < .001$. The degree of heterogeneity between effect sizes was $I^2 = .00$.

The two publication bias analyses provided slightly smaller estimates. Trim-and-fill analysis filled 21 studies left of the mean. The estimated overall effect size was $\bar{g} = 0.04$, 95% CI [-0.01; 0.10]. The estimates of the selection model analysis were $\bar{g} = 0.03$, $\bar{g} = -0.07$, $\bar{g} = 0.09$, and $\bar{g} = 0.07$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. Five influential effect sizes were detected. The overall effect size without these effect sizes was not different from the unadjusted effect size, $\bar{g} = 0.10$, 95% CI [0.05; 0.15], $k = 234$, $p < .001$, $I^2 = .00$.

Finally, the overall effect sizes sorted by Outcome measure were calculated. All the overall effect sizes were small. The influential case and publication bias analyses had no substantial impact on the estimated values. The results are summarized in Table 22.

Table 22

Meta-Analytic and Publication Bias Results Sorted by Outcome Measure, for the Effect Sizes Referring to Adult Video Game Players (Meta-Analysis 3)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|-----------------------|-----------------|-----------------------|-----------------------|------------------------------|---------------------------------|-----------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 96 | 0.12 [0.03; 0.22] | .013 | .00 | 0.01 [-0.10; 0.12] | 0.04; -0.08; 0.10; 0.08; | 94 | 0.14 [0.05; 0.23] | .002 | .00 |
| Spatial ability | 60 | 0.11 [0.02; 0.21] | .021 | .00 | 0.06 [-0.03; 0.16] | 0.05; -0.05; 0.10; 0.08; | 59 | 0.11 [0.01; 0.21] | .029 | .00 |
| Cognitive control | 36 | 0.02 [-0.16; 0.20] | .842 | 38.19 | 0.12 [-0.07; 0.32] | -0.03; -0.15; 0.03; 0.03; | 36 | 0.02 [-0.16; 0.20] | .842 | 38.19 |
| Memory | 32 | 0.10 [-0.03; 0.23] | .134 | .00 | 0.15 [0.03; 0.28] | 0.03; -0.07; 0.09; 0.07; | 32 | 0.10 [-0.03; 0.23] | .134 | .00 |
| Intelligence/reasoning | 15 | 0.16 [-0.01; 0.33] | .073 | 14.49 | - | 0.10; 0.02; 0.14; 0.11; | 15 | 0.16 [-0.01; 0.33] | .073 | 14.49 |

Note. See Note to Table 18 for abbreviations.

Second, we analysed the results for the old video game players. In the sub-model comprising all effect sizes, the random-effects meta-analytic overall effect size was $\bar{g} = -0.08$, 95% CI [-0.21; 0.04], $k = 92$, $p = .184$. The degree of heterogeneity between effect sizes was $I^2 = 48.90$.

Trim-and-fill analysis filled 21 studies left of the mean. The estimated overall effect size was $\bar{g} = -0.28$, 95% CI [-0.41; -0.16]. The estimates of the selection model analysis were $\bar{g} = -0.18$, $\bar{g} = -0.32$, $\bar{g} = -0.09$, and $\bar{g} = -0.07$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. One influential effect size was detected. The overall effect size without these effect sizes was $\bar{g} = -0.06$, 95% CI [-0.18; 0.05], $k = 91$, $p < .001$, $I^2 = 43.88$.

Finally, the overall effect sizes sorted by Outcome measure were calculated. Four overall effect sizes were small or null. The overall effect size referring to intelligence/reasoning-related measures was significantly negative ($\bar{g} = -0.63$). No influential case was detected. The publication bias analyses estimated values similar to the unadjusted effect sizes. The results are summarized in Table 23.

Table 23

Meta-Analytic and Publication Bias Results of the Effect Sizes Referring to the Old Video Game Players Sorted by Outcome Measure (Meta-Analysis 3)

| Outcome Measure | Model | | | | | | Model without influential cases | | | |
|-----------------------------|----------|-------------------------|-----------------|-----------------------|-----|--------------------------------|---------------------------------|-------------------------|-----------------|-----------------------|
| | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² | T&F | SM | <i>k</i> | \bar{g} | <i>p</i> -value | <i>I</i> ² |
| Visual attention/processing | 26 | 0.00 [-0.21; 0.22] | .965 | 31.74 | - | -0.06; -0.20; 0.02; 0.01; | 26 | 0.00 [-0.21; 0.22] | .965 | 31.74 |
| Spatial ability | 12 | 0.06 [-0.20; 0.33] | .646 | .00 | - | -0.03; -0.17; 0.05; 0.04; | 12 | 0.06 [-0.20; 0.33] | .646 | .00 |
| Cognitive control | 17 | 0.07 [-0.15; 0.29] | .539 | 9.32 | - | -0.03; -0.16; 0.05; 0.04; | 17 | 0.07 [-0.15; 0.29] | .539 | 9.32 |
| Memory | 22 | -0.01 [-0.24; 0.22] | .922 | 18.87 | - | -0.12; -0.28; -0.03; -0.02; | 22 | -0.01 [-0.24; 0.22] | .922 | 18.87 |
| Intelligence/reasoning | 15 | -0.63 [-1.06; -0.20] | .004 | 63.45 | - | -0.66; -0.85; -0.57; -0.49; | 15 | -0.63 [-1.06; -0.20] | .004 | 63.45 |

Note. See Note to Table 18 for abbreviations.

Finally, a sub-model comprising all the effect sizes referring to the young video game players was run. The random-effects meta-analytic overall effect size was $\bar{g} = 0.21$, 95% CI [0.06; 0.36], $k = 28$, $p = .007$. The degree of heterogeneity between effect sizes was $I^2 = 26.99$.

The estimates of the selection model analysis were moderately smaller than the unadjusted effect size, $\bar{g} = 0.15$, $\bar{g} = 0.07$, $\bar{g} = 0.19$, and $\bar{g} = 0.15$ for moderate one-tailed selection, severe one-tailed selection, moderate two-tailed selection, and severe two-tailed selection, respectively. No influential effect size was detected. Due to the scarcity of the effect sizes, no sub-models of the outcome measures were run.

7.3 Discussion

The main model showed a near-zero effect of video-game training on overall cognitive ability ($\bar{g} = 0.07$). Moreover, this effect was found to be a slight overestimation by the publication bias analyses. The same pattern of results occurred in nearly every sub-model, regardless of the type of game – neither action nor non-action video game training exerted any substantial effect on the participants' cognitive ability – and the age of the participants. A significant exception was the negative effect of video game training on intelligence/reasoning-related measures in the sample of old people ($\bar{g} = -0.63$). Given the small number of the effect sizes in that model ($k = 15$) and the high degree of heterogeneity ($I^2 = 63.45$), the overall effect size was probably biased. Finally, duration of training was not a significant moderator. This latter outcome is further evidence against the hypothesis according to which video game training affects cognitive ability: if training were effective, one should expect a positive relationship between duration of training and size of the effects.

8. General Discussion

This paper has addressed the question of the impact of video games on cognitive ability. The three meta-analyses offer a consistent picture: weak correlations between skill and cognitive ability, small differences between video game players and non-players, and no differences between the participants who underwent video game training and the participants in the control groups. In those few cases reporting medium effect sizes, the estimates of publication bias analysis were significantly smaller (e.g., Visual attention/processing overall effect sizes in Table 15). Crucially, most of the models showed a small (or zero) degree of heterogeneity, indicating that most of the variability between studies was due to sampling error (or a few influential cases) rather than some moderating variable. This was particularly evident in the meta-analysis of experimental data ($I^2 = 17.90$ and $I^2 = 10.64$ including and excluding influential cases, respectively). The small degree of heterogeneity often observed in the sub-models also suggests that the categorization of the effect sizes into the five outcome measures is highly reliable.

The findings of the present meta-analytic investigation differed significantly from the more positive results of previous meta-analyses (e.g., Powers et al., 2013). This difference is probably due to our more restrictive inclusion criteria of the meta-analysis of training studies and our more accurate procedure of calculation of the effect sizes and correction of statistical dependence. That said, it is worth mentioning that the present meta-analytic investigation also shows that playing certain types of video games may be related to specific cognitive abilities. For example, there seem to be a reliable, yet small, correlation between video game skill and spatial ability. Also, action video game players appear to outperform non-players in tasks related to visual attention/processing. Thus, the field of video game playing may present characteristics analogous to other domains of expertise, such as chess and music. Playing

video games in general, or some genre in particular, may be associated with specific cognitive abilities predicting, to some extent, a player's skill. However, just like the fields of music and chess, video game training does not generally enhance cognitive ability.

8.1 Theoretical and Practical Implications

Along with substantial research into expertise acquisition and other types of cognitive training, the results of the present meta-analytic investigation point towards a clear direction: while it is evident that training one skill improves that skill, far transfer is extremely unlikely to occur. Video game training is no exception.

The most significant implication of these results is that the lack of generalization across different domains of skills acquired by training appears to be a constant in human cognition. Domain-general cognitive abilities are malleable to training, but the benefits, when there are any, are domain-specific (Chase & Ericsson, 1982; Gobet, 2016). Moreover, as highlighted by Shipstead, Redick, and Engle (2012), such limited benefits, observed after training, probably represent only trainees' improved ability to perform a task. In other words, people may get better at solving cognitive tasks similar to the training task, and yet not show any genuine improvement in cognitive ability. This account also explains why video game training has been sometimes associated with improvements in particular tasks (e.g., UFOV; Feng, Spence, & Pratt, 2007), whereas no effect has been found in broader cognitive constructs (e.g., visual attention/processing; Table 18).

Second, far transfer must be considered a fundamental litmus test for theories of human cognition. The failure of generalization of skills in the field of video gaming represents a further corroboration of those theories of cognition that predict no (or limited) far transfer, such as chunking theory (Chase & Simon, 1973) and template theory (Gobet &

Simon, 1996). More generally, our results support the hypothesis according to which expertise acquisition relies to a large extent on domain-specific, and hence non-transferable, information. By contrast, those theories predicting the occurrence of far transfer after video game training (e.g., “learning to learn;” Bavelier et al., 2012) and cognitive training in general (for a review, see Strobach & Karbach, 2016) are not supported.

Third, given the small or null effects exerted by video gaming on cognitive tests, the neural changes and patterns observed in video game players in several studies (e.g., Colom et al., 2012) probably reflect modifications in domain-specific abilities (e.g., video game skills) rather than domain-general improvements of cognitive ability. Interestingly, the presence of specific neural patterns (functional and anatomical) and absence of significant effects on cognitive tests have also been observed in other domains such as music (e.g., Tierney, Krizman, & Kraus, 2015) and chess (e.g., Hänggi, Brüttsch, Siegel, & Jäncke, 2014). Whether this pattern of results occurs regardless of the domain considered will be a requirement for future research.

Beyond theoretical aspects, the absence of far transfer has important practical implications. If trained skills rarely generalize across different domains, then deliberately training one skill remains the most effective way to acquire that skill. This consideration may appear trivial. However, this conclusion is in contrast with common belief and practice in education and the professions. For instance, considerable emphasis has been given to teaching students transferable skills in recent years (Pellegrino & Hilton, 2012). However, in light of the findings provided by the research on expertise acquisition and cognitive training, the common view that training and possessing transferable skills is one effective way to progress in a particular field appears incorrect. Our conviction is that educational and

professional training should focus on subject-related contents rather than general skills or principles without any explicit reference to any specific discipline.

8.2 Recommendations for Future Research

Given the current evidence, insisting on searching for improbable generalized effects of video game training on cognitive function appears pointless. Rather, the field should focus on investigating the exact cognitive correlates of video game expertise. Specifically, further research is needed to understand whether video game players exhibit general superior cognitive ability or excel only at tasks related to video game expertise. For example, chess masters can recall entire chess positions, even when the material is presented only for a few seconds (e.g., Gobet & Simon, 2000). However, the correlation between chess skill and performance on tests of short-term memory is modest ($r = .22$; Burgoyne et al., 2016). Like chess players, video game players may possess exceptional cognitive abilities only with domain-specific material. A series of experiments testing video game players' performance with both domain-general and domain-specific tasks (e.g., recalling of video game scenarios) would clarify whether and in what contexts video game players show superior cognitive ability.

With regard to the effects of video game training, research in the field should investigate the relationship between the degree of transfer and trainees' baseline cognitive ability. It may be possible that people with below-average and compromised cognitive ability benefit from video game training more than do people with normal (or superior) cognitive function. Moreover, video game training could slow down cognitive decline in older adults and possibly restore impaired cognitive ability. While the meta-analysis of experimental studies reported null effects for video game training on the older adults' cognitive ability, and

thus does not support this hypothesis, the topic probably deserves further investigation, given that no clinical population was included.

8.3 Conclusion

Our comprehensive meta-analytic investigation showed that the relationship between cognitive ability and playing video games is weak. Small or null correlations were obtained in the first meta-analysis. The second meta-analysis reported that video game players' advantage over non-players was modest. Finally, the third meta-analysis found no meaningful effect of video game training on any of the reviewed outcome measures. These findings are in line with substantial research into expertise and cognitive training in domains such as music, chess, WM, and brain training. To date, far transfer remains a chimera.

The generalized absence of far transfer has profound implications. Theories of human cognition predicting (or assuming) the occurrence of far-transfer effects find no support. Conversely, theories predicting no far-transfer effects are corroborated. As for academic and professional education, the lack of far transfer should encourage educators, trainers, and policymakers to implement curricula extensively focused on subject-related material.

Chapter 8: Meta-Analysis of Exergame Training

Rationale for the Meta-Analysis in Chapter 8

Chapter 8 examines a particular variant of video-games: exergames. This field is relatively new compared to the ones discussed in the previous Chapters (the first study included in the meta-analysis was carried out in 2010.) Nonetheless, this meta-analysis – which is a re-analysis of a previous study by Stanmore et al. (2017) – is a worthwhile contribution to the field of cognitive training. In fact, the findings presented in this Chapter represent valuable evidence on which to judge the effects of cognitive-training regimens associated with physical exercises.

The studies included in this meta-analysis are listed in Appendix G.

1. Introduction

Cognitive training has been one of the most influential topics in cognitive science in the last 20 years. The possibility to train global cognitive function would have profound consequences for our understanding of how humans acquire and use knowledge. Moreover, whether and to what degree cognitive ability is malleable has huge practical implications. Consider the educational advantages of enhancing cognitive ability in youth and the benefits, for the public health and global economy, of slowing down cognitive decline in old age.

Due to the potential practical and theoretical implications, the effects of cognitive training have been subject to a lively debate. Whilst some researchers have endorsed cognitive-training activities as tools to enhance general cognition, other scholars in the field have expressed their scepticism about such practices. The controversy is well described by two open letters regarding the presumed benefits of brain-training programs. The first letter, issued by the Stanford Center on Longevity and the Max Planck Institute for Human Development, signed by 75 scholars in the fields of psychology and neuroscience, has expressed serious doubts about the capability of brain games of enhancing cognitive function (“A Consensus on the Brain Training Industry from the Scientific Community,” 2014). The second one, posted on the Cognitive Training Data website (www.cognitivetrainingdata.org) and signed by 133 researchers, has claimed that certain cognitive-training programs can benefit cognitive function.

Crucially, substantial experimental evidence has highlighted the limitations, rather than the benefits, of many different cognitive-training programs. Cognitive-training regimens seem to affect only the trained skills, while no effect is exerted on non-trained tasks. This seems to apply to both those tasks that specifically train cognitive abilities – such as *n*-back tasks in WM training and brain-training programs (Melby-Lervåg, Redick, & Hulme, 2016;

Sala & Gobet, 2017b; Simons et al., 2016) – and intellectually demanding activities such as chess (Sala & Gobet, 2016), video games (Oei & Patterson, 2015), and music (Mosing, Madison, Pedersen, & Ullén, 2016; Sala & Gobet, 2017c).

Given the disappointing results of cognitive-training programs, other pathways to cognitive enhancement have been explored. One of the most appealing is certainly exergaming. Exergames are video games combining cognitive and physical training. Due to this peculiarity, exergaming is thus believed to positively affect overall cognition more than aerobic exercise or cognitive tasks alone. In fact, while aerobic exercise is thought to support oxygenation of the brain and induce neurogenesis (Fabel et al., 2009; Firth et al., 2016; Kempermann et al., 2010), the gamification of cognitive tasks promotes participants' engagement in training programs (Anguera et al., 2013; Stine-Morrow et al., 2014). The presumed benefits of exergaming are thus thought to stem from the combination of the positive effects of engagement, cognitive training, and physical exercise.

A recent meta-analysis (Stanmore, Stubbs, Vancampfort, de Bruin, & Firth, 2017) has thus investigated the effects of exergames on overall cognitive ability. This meta-analysis includes 17 Randomized Control Trials (RCTs) and a total of 926 participants. In most of the studies ($n = 15$), the participants were older people (mean age > 55) with either no or some clinical condition (e.g., Parkinson's disease). The cognitive performance of the exergames-treated participants is compared to the performance of participants involved in a number of different activities (e.g., stretching, reading, and cycling) or no activity at all (passive control groups). The meta-analysis reports a medium overall effect size ($\bar{g} = 0.436$), indicating that exergames can represent an effective tool to improve global cognition and to slow down cognitive decline.

However, due to methodological issues, we think that the results of this meta-analysis are substantially unreliable. First, due to mistakes in the calculations and statistical artefacts, some effect sizes are inflated (for details, see Sections 3. and 4.). Thus, the reported overall effect size is probably an overestimation. Second, the degree of heterogeneity is around 66%. Such degree of heterogeneity indicates, according to a standard categorization (Higgins, Thompson, Deeks, & Altman, 2003), a medium to high between-study variability that, if not explained, does not allow to draw any clear conclusion about the actual effect of the treatment. Third, even though there is some evidence of publication bias (e.g., asymmetrical distribution of the effect sizes in the funnel plot), no proper publication-bias analysis is carried out. Even though Stanmore et al. (2017) run two publication-bias analyses – Begg and Mazumdar’s (1994) rank correlation test and Orwin’s (1983) fail-safe N – neither of these methods provides an adjusted estimate of the overall effect size. In addition, the fail-safe N has been found to provide misleading results (for details, see Becker, 2005; see also Schmidt & Hunter, 2015; pp. 531-534).

Based on these issues, and considering the important theoretical and practical implications, we present a re-analysis of Stanmore et al.’s data (2017). We carry out two models. Model 1 is run by merging all the effect sizes extracted from the same sample into statistically independent effect sizes. Model 2 includes all the effect sizes representing the difference between treatments and controls in every measure of cognitive ability reported in the primary studies. Our analyses focus on the way to calculate effect sizes and multiple publication-bias analyses. Also, we run a sensitivity analysis to control for possible statistical artefacts in the effect sizes.

2. Method

2.1 Included Studies

We used all the studies (RCTs) included in Stanmore et al.'s (2017) meta-analysis except one (i.e., Ackerman, Kanfer, & Calderwood, 2010). This study investigated the effects of the Wii *Big Brain Academy* program that consists of a set of brain-training – rather than exergaming – activities. It is worth noting that Ackerman et al. (2010) reported negative effects of the treatment on the participant cognitive ability. Thus, the exclusion of this study increased – rather than diminished – the overall effect sizes in our models. The number of included studies and independent samples was 16, while the total number of participants was 883.

2.2 Effect Sizes

The effect sizes were calculated for each measure of cognitive ability reported in the studies. When both pre- and post-test data were provided, the standardized means difference (Cohen's d) was calculated with the following formula:

$$d = (M_{g-e} - M_{g-c}) / SD_{pooled-pre} \quad (1)$$

where M_{g-e} and M_{g-c} are the gain of the experimental group and the control group immediately after the end of the training, respectively, and $SD_{pooled-pre}$ is the pooled standard deviation of the two pre-test standard deviations. This formula represents the most appropriate way to calculate the effect size in intervention studies with a repeated-measure design (for details, see Schmidt & Hunter, 2015; pp. 352-353). Also, when the pre-post-test between-group performance was expressed by p -values (Barcelos et al., 2015; Mirelman et al., 2016), the values were converted into Cohen's d s by the Comprehensive Meta-Analysis (CMA; Version

3.0; Biostat, Englewood, NJ) software package. When these data were not reported in the primary studies, we contacted the authors. Three authors positively replied to our e-mails.

When only pre-post-test differences were available (Maillot, Perrot, & Hartley, 2012; Zimmermann et al., 2014), we used the difference between the mean changes and the standard deviations of the changes. This procedure may have inflated the effect sizes extracted from these two studies because standard deviations of pre-post changes are usually significantly smaller than pre-test ones. The influence of these effect sizes on the results was assessed in a sensitivity analysis (see Section 3.3.).

Finally, CMA was used for converting the Cohen's d s into Hedges's g (Hedges & Olkin, 1985), computing the overall effect sizes (\bar{g} s), and conducting statistical analyses.

2.3 Meta-Analytic Models

In the first model (Model 1), the measures from the same sample were meta-analytically averaged into 16 statistically independent effect sizes. In one study (Chan, Ngai, Leung, & Wong, 2010), we used the global score (Cognistat) reported in the text. The whole procedure was analogous to the one used by Stanmore et al. (2017).

In the second model (Model 2), all the effect sizes ($k = 75$) were inserted. As this procedure violates the principle of statistical independence of the samples, Cheung and Chan's (2004) correction was applied. This correction reduces the weight of dependent samples in the analysis by estimating an adjusted (i.e., smaller) N . This method also allows one to build more powerful models without losing any information from the primary studies, biasing the meta-analytic mean, or artificially inflating the degree of heterogeneity.

2.4 Publication Bias Analyses

Publication bias is unanimously acknowledged as a serious problem in meta-analysis and scientific research in general (Begg & Berlin, 1988; Schmidt & Hunter, 2015). For this reason, it has been proposed to use multiple analyses not only to detect the possible publication bias but also to triangulate the true (i.e., unbiased) effect size (e.g., Kepes & McDaniel, 2015).

We thus chose three publication-bias analyses. First, Egger's (Egger, Smith, Schneider, & Minder, 1997) regression test was used to test whether the effect sizes were distributed symmetrically around the meta-analytic mean in the funnel plot. The trim-and-fill analysis (Duval & Tweedie, 2000) was then applied to estimate the number of below-average missing studies and calculate the unbiased effect size. Finally, the PET-PEESE estimators were calculated (Stanley & Doucouliagos, 2014). PET estimator is the intercept of a weighted linear regression where the dependent variable is the effect size, the independent variable is the standard error, and the weight is the inverse of the standard error squared. PEESE estimator is obtained by replacing the standard error with the standard error squared as the independent variable. If PET suggests the presence of a real effect (i.e., intercept different from zero), PEESE estimator must be considered as the corrected overall effect size (Stanley & Doucouliagos, 2014).

3. Results

3.1 Model 1 ($K = 16$)

The random-effects meta-analytic overall effect size was $\bar{g} = 0.286$, 95% CI [0.102; 0.470], $k = 16$, $p = .002$. The test of heterogeneity was marginally significant, $Q(15) = 24.034$, $I^2 = 37.589$, $p = .065$.

The contour-enhanced funnel plot (Peters, Sutton, Jones, Abrams, & Rushton, 2008; Figure 15) representing the relation between effect sizes and standard errors looked asymmetrical, suggesting the presence of publication bias.

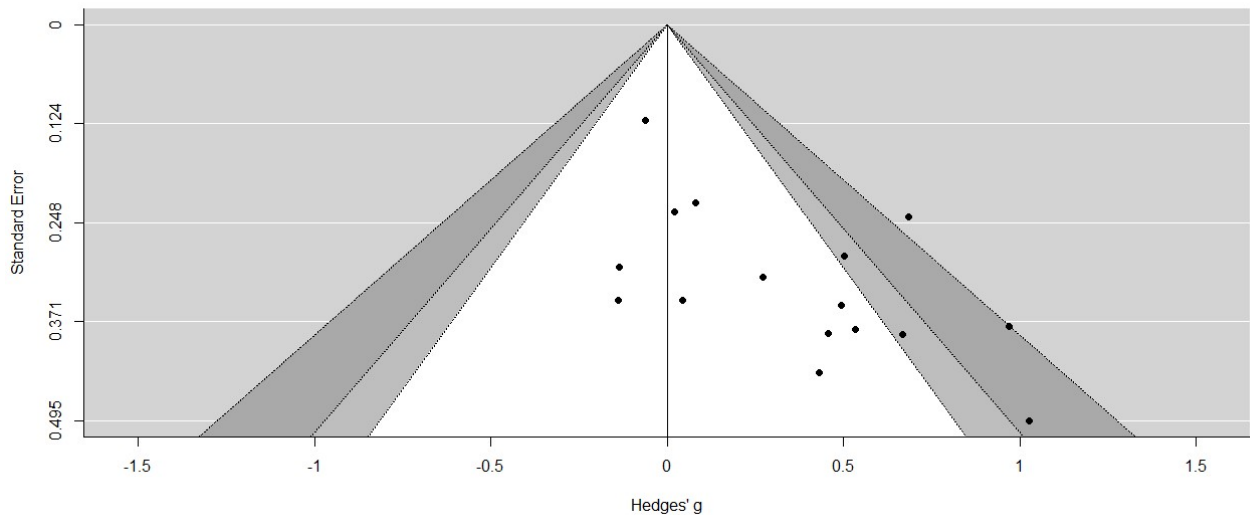


Figure 15. Contour-enhanced funnel plot of standard errors and effect sizes (Hedges's g s) in Model 1. The black circles represent the effect sizes included in the model. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

Egger's regression test confirmed that the effect sizes were asymmetrically distributed ($p = .003$, one-tailed). The trim-and-fill analysis estimated six weaker-than-average missing effect sizes (left of the mean). The adjusted overall effect size was $\bar{g} = 0.079$, 95% CI [-0.125; 0.283]. Finally, the PET and PEESE estimators were $\bar{g} = -0.302$ ($p = .006$) and $\bar{g} = -0.082$ ($p = .006$), respectively. Since PET is different from zero, PEESE estimator must be taken as the adjusted overall effect size.

3.2 Model 2 ($K = 75$)

The random-effects meta-analytic overall effect size was $\bar{g} = 0.240$, 95% CI [0.141; 0.340], $k = 75$, $p < .001$. The test of heterogeneity was significant, $Q(74) = 107.316$, $I^2 = 31.045$, $p = .007$.

The contour-enhanced funnel plot (Figure 16) representing the relation between effect sizes and standard errors looked asymmetrical, suggesting the presence of publication bias.

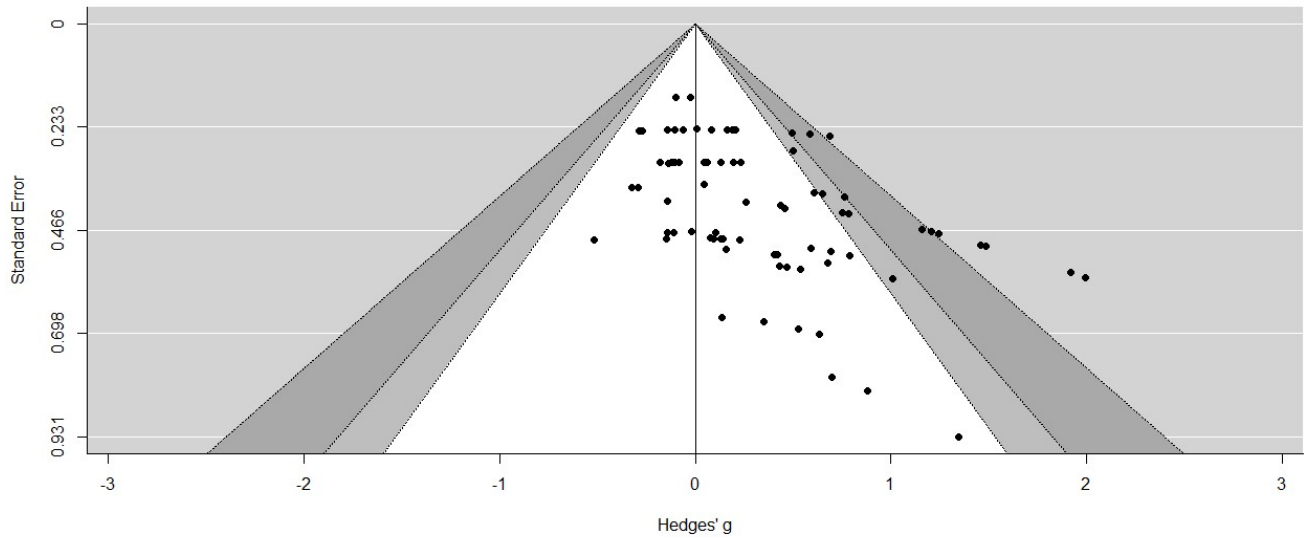


Figure 16. Contour-enhanced funnel plot of standard errors and effect sizes (Hedges's g s) in Model 2. The black circles represent the effect sizes included in the model. Contour lines are at 1%, 5%, and 10% levels of statistical significance.

Egger's regression test confirmed that the effect sizes were asymmetrically distributed ($p < .001$). The trim-and-fill analysis estimated 21 weaker-than-average missing effect sizes (left of the mean). The adjusted overall effect size was $\bar{g} = 0.076$, 95% CI [-0.039; 0.190]. Finally, the PET and PEESE estimators were $\bar{g} = -0.325$ ($p < .001$) and $\bar{g} = -0.040$ ($p < .001$), respectively.

3.3 Sensitivity Analysis

As observed above, the effect sizes extracted from Maillot et al. (2012) and Zimmermann et al. (2014) may be biased. For this reason, two additional models were run by excluding these effect sizes from Models 1 and 2.

In the first model, the random-effects meta-analytic overall effect size was $\bar{g} = 0.245$, 95% CI [0.058; 0.432], $k = 14$, $p = .010$. The test of heterogeneity was non-significant, $Q(13) = 19.637$, $I^2 = 33.799$, $p = .105$. Egger's regression test suggested the presence of publication bias that the effect sizes were asymmetrically distributed ($p = .008$, one-tailed). The trim-and-fill analysis estimated five weaker-than-average missing effect sizes (left of the mean). The adjusted overall effect size was $\bar{g} = 0.063$, 95% CI [-0.145; 0.271]. Finally, the PET and PEESE estimators were $\bar{g} = -0.271$ ($p = .016$) and $\bar{g} = -0.071$ ($p = .016$), respectively.

In the second model, the random-effects meta-analytic overall effect size was $\bar{g} = 0.113$, 95% CI [0.029; 0.197], $k = 58$, $p = .008$. The test of heterogeneity was non-significant, $Q(57) = 44.013$, $I^2 = 0.000$, $p = .896$. Egger's regression test suggested the presence of publication bias that the effect sizes were asymmetrically distributed ($p = .001$, one-tailed). The trim-and-fill analysis estimated 18 weaker-than-average missing effect sizes (left of the mean). The adjusted overall effect size was $\bar{g} = 0.003$, 95% CI [-0.086; 0.092]. Finally, the PET and PEESE estimators were $\bar{g} = -0.162$ ($p = .002$) and $\bar{g} = -0.010$ ($p = .001$), respectively.

4. Discussion

The aim of the present paper was to test the reliability of the findings of a recent meta-analysis about the effects of exergame intervention on overall cognitive ability (Stanmore et al., 2017). Contrary to the claims of that meta-analysis, our re-analyses of the data have shown that the overall effect (if any) of such practices on one's cognitive ability is small.

4.1 Inflation of the Effect Sizes

A simple comparison between Model 1 and Stanmore et al.'s (2017) meta-analytical model of the effects of exergames on global cognition (Fig. 2, p. 39) shows that there are no large differences in most of the effect sizes. However, a few effect sizes are meaningfully greater in Stanmore et al.'s (2017) models than in ours (e.g., Anderson-Hanley et al., 2012). We do not know exactly how Stanmore and colleagues calculated these effect sizes.

The inflation of some effect sizes also depends on some statistical artefacts. In fact, our sensitivity analysis shows that excluding those effect sizes at high risk of upward bias (Maillot et al., 2012; Zimmermann et al., 2014) sensibly reduces the overall effect size, especially in Model 2. Moreover, other studies such as Barcelos et al. (2015), Schättin, Arner, Gennaro, and de Bruin (2016), and Staiano, Abraham, and Calvert (2012) report differences at baseline between groups, while in Hughes et al. (2014) and Schoene et al. (2013) a decrement in the control group's performance is observed in some tests. It is thus likely that the effect sizes from these studies are inflated as well.

These considerations strongly suggest that the overall effect size reported in Stanmore et al. (2017; $\bar{g} = 0.436$) is an overestimation. The same applies, to a lesser extent, to the overall effect sizes calculated in our Models 1 and 2 ($\bar{g} = 0.286$ and $\bar{g} = 0.240$, respectively).

4.2 Heterogeneity

Like the overall effect sizes, the degree of heterogeneity in our models is much smaller than the one reported in Stanmore et al. (2017). This difference is the logical consequence of the corrections in the calculations of the effect sizes. In Model 1, the observed heterogeneity is marginally significant ($I^2 = 37.589$, $p = .065$). In Model 2, the degree of heterogeneity is between small and medium ($I^2 = 31.045$, $p = .007$). As highlighted

by the sensitivity analysis, the presence of some inflated effect sizes – rather than some potential moderator (e.g., cognitive ability assessed, age, clinical condition, and type of control group) – accounts for most of the heterogeneity in Model 1 ($I^2 = 33.799, p = .105$) and, more clearly, in Model 2 ($I^2 = 0.000, p = .896$; see Section 3.3).

4.3 Publication Bias

In both models, the analyses show a textbook case of publication bias. The largest effect sizes are associated with the largest standard errors (i.e., smallest N s) producing an asymmetrical distribution in the funnel plot. Consequently, the three publication bias analyses provide consistent outcomes and estimate that the true effect of exergaming activities on overall cognitive ability is zero.

Like heterogeneity, the presence of publication bias seems to depend on the inflation of some effect sizes. Consistent with this argument, the exclusion of Maillot et al. (2012) and Zimmermann et al.'s (2014) effect sizes from the models reduces the difference between the overall effect sizes and the adjusted estimates.

4.4 Conclusions and Recommendations for Future Research

Our re-analyses show that the correct calculation of the effect sizes from the primary studies diminishes both the overall effect sizes and heterogeneity. Also, the publication-bias analyses suggest that the actual effect on overall cognitive ability of exergaming is around zero. This pattern of results occurs in both Models 1 (statistically independent effect sizes) and 2 (statistically dependent effect sizes). Taken together, these outcomes cast serious doubts about the presumed cognitive benefits of exergames.

Due to the small number of RTCs carried out so far and the between-study variability in the participants' characteristics, further research is recommendable. The current meta-

analysis cannot exclude that exergames may improve (or preserve) some cognitive abilities in specific clinical and non-clinical populations. Moreover, as already mentioned, most of the reviewed studies tested the effects of the treatment on older participants. Thus, the results of the current meta-analysis, while representing robust evidence against the effectiveness of exergames at improving cognitive function in the elderly, are hardly generalizable to individuals of different age.

That said, in line with substantial research into cognitive and educational psychology, our re-analyses corroborate the idea that cognitive-training activities do not benefit overall cognitive function or promote transfer of skills to non-trained tasks (Melby-Lervåg et al., 2016; Sala & Gobet, in press; Shipstead, Redick, & Engle, 2012; Simons et al., 2016). Following these considerations, future studies should investigate the effects of specific training regimens on particular cognitive skills in populations with certain features (e.g., age range and clinical condition) rather than searching for generalized benefits.

Chapter 9: Overall Discussion

This chapter summarizes the results from the five meta-analyses and discusses their theoretical and practical implications.

1. Summary of the Results

The first meta-analysis (Chapter 4) investigated the effects of working memory training on typically developing children's cognitive skills and academic achievement. While the treatment exerted a medium-size effect on measures of near transfer (i.e., working memory and short-term memory tests), no genuine improvement was observed in any measure of far transfer. The same pattern of results remained several months after the end of training. Moreover, the moderator analyses showed that the size of the effect of the treatment was inversely related to the quality of the design in far-transfer measures. Some positive, yet small, effects were observed in non-randomized samples and experimental groups compared to passive control groups. When the samples were randomized, and the treated participants were compared with active control groups, the effect of the treatment was null. Regarding near-transfer measures, no moderator was significant. Finally, publication bias analysis found no evidence of missing studies.

The second meta-analysis (Chapter 5) analysed the impact of chess instruction on children and young adolescents' overall cognitive ability and academic achievement (mathematics and literacy). The results showed small to medium effects on all these three measures. No or little publication bias was found. The moderator analyses highlighted a positive relationship between the duration of training and size of the effects. This outcome suggests that the effects of chess instruction do transfer to other domains as long as the treatment is long enough (25 hours is the proposed threshold). However, it must be noted that

almost no study included in this meta-analytic review included active control groups. Thus, the observed positive effects could have been due to placebos. In line with this hypothesis, a recent study found no significant difference between a chess-treated group, an active control group (playing checkers), and a passive control group in mathematical or meta-cognitive skills (Sala & Gobet, 2017d).

Finally, due to the scarceness of the effects related to performance in cognitive tests, it is not possible to analyse the effect of chess on specific cognitive abilities (e.g., cognitive control and working memory). Therefore, the mechanisms (if any) linking chess training to enhanced cognitive ability and improvement in school disciplines are yet to be tested properly.

The third meta-analysis (Chapter 6) dealt with the effects of music training on cognitive skills and academic achievement in children and young adolescents with no learning disabilities. While the overall effect size was modest, slightly larger effects were observed in measures of memory and intelligence. However, the type of control group used in the experiment and the presence (or absence) of random allocation were found to be significant moderators. The pattern of results was similar to the one observed in the meta-analysis about working memory training. While some positive effects were obtained when the treated groups were compared with passive control groups, the comparisons with active control groups provided near-zero effects. Analogously, the overall effect size in randomized samples was significantly smaller than the one in non-randomized samples. Finally, no evidence of publication bias was found.

The fourth meta-analysis (Chapter 7) was a more extended investigation on the relationship between video-game practice and cognitive skills in the general population.

Three meta-analytical models were presented. The first meta-analytic model included all the studies reporting correlations between video-game skill and cognitive ability. The overall correlation was tiny, and nearly all the sub-models reported small or null correlations. Regarding the different types of video-game, no significant difference was found. The moderator analyses also showed that correlations were slightly larger when video-game skill was measured by the video-game scores than the frequency of play. The same applies to the correlations between video-game skill and spatial ability. Near-zero correlations were found in all the other outcome measures. Finally, trim-and-fill and selection models analyses reported some evidence of publication bias.

The second meta-analytic model included all the studies assessing cognitive performance in video-game players and non-players. Video-game players proved to be better in all the five outcome measures. However, the publication bias analyses calculated reduced estimates of the effect sizes. Thus, the difference between video-game players and non-players is probably significantly smaller than the one usually reported in the primary studies. Finally, the moderator analyses found no effect of the type of video game.

The third meta-analytic model examined the effects of video-game training on non-players' cognitive skills. Negligible effects were found in both the main model and all sub-models. Notably, this pattern of results occurred regardless the type of video game (action or non-action).

The fifth meta-analysis (Chapter 8) was a re-analysis of Stanmore et al.'s (2017) meta-analysis about the impact of exergames on overall cognitive function. Contrary to what Stanmore and colleagues claimed, the effects of exergames were found to be small if not null. In the re-analysis included in this dissertation, the effect sizes were correctly extracted from

the primary studies, and more powerful publication bias analyses were run. The present re-analysis is thus a perspicuous example of how errors in the calculation of the effect sizes and the lack of a proper sensitivity analysis may lead to biased results.

In sum, all the five meta-analyses have provided weak or no support for the hypothesis that practicing cognitive tasks or engaging in cognitively demanding activities enhance overall cognition. Rather, the benefits do not transfer to other domains and thus appear to be task-specific (e.g., working memory training). Furthermore, it is worth noting that most of the meta-analytic models report low or no degree of heterogeneity. This outcome confirms the reliability of the results.

2. Theoretical and Practical Implications

2.1 Main Implications

The five meta-analyses point to an inescapable conclusion: there is no far transfer regardless of the type of training. This state of affairs occurs not only in the cognitive-training programs reviewed in this dissertation (i.e., working memory training, chess, music, video-game training, and exergames) but also in domains studied by other researchers such as brain-training (Simons et al., 2016). Therefore, the most relevant implication of these findings is that the lack of generalization across different domains of skills acquired by training seems to be a constant in human cognition.

This broad meta-analytic investigation and its clear-cut conclusions provide strong support for Thorndike and Woodworth's common elements theory and, more generally, to all those theories of expertise and skill acquisition predicting no (or little) far transfer, such as chunking (Chase & Simon, 1973) and template theories (Gobet & Simon, 1996). Conversely, those theories postulating the possibility of enhancing domain-general cognitive ability are

not supported. Examples of such theories include Halford et al.'s (2007) theory of common capacity constraint, Schellenberg's (2004, 2006) theory of IQ enhancement through music activities, and Bavelier et al.'s (2012) learning to learn theory.

Finally, it must be noted that the lack of far transfer does not imply that human cognition is not malleable to training. In fact, cognitive-training often exerts visible effects. For example, working memory training improves performance in memory-related tests (see Chapter 5). Analogously, spatial training positively affects one's ability to solve spatial tasks. However, these effects are task-specific (e.g., Chase & Ericsson, 1982) and do not necessarily represent improvements in broad cognitive constructs (I will take up this point below).

2.2 Training and Neural Plasticity

Another theoretical implication of these results concerns neural plasticity. The substantial absence of far transfer suggests that the neural patterns observed in people engaged in cognitively demanding activities reflect modifications in domain-specific abilities (e.g., chess skill) rather than enhanced domain-general cognitive ability. The occurrence of specific neural patterns (anatomical and functional) and absence of far-transfer effects on cognitive tests have been reported in domains such as music (e.g., Tierney, Krizman, & Kraus, 2015), chess (e.g., Hänggi, Brüttsch, Siegel, & Jäncke, 2014), video game training (e.g., Colom et al., 2012), and working memory training (Clark, Lawlor-Savage, & Goghari, 2017).

Understanding the actual significance of the observed neural patterns is essential to avoid misinterpretations. The domain of music offers a perspicuous example of this problem. Given the findings reported in Chapter 6, it is implausible that functional changes occurring after a music-training intervention represent domain-general improvements in cognitive

function. Rather, it is likely that such neural patterns underlie the enhancement of music-related skills such as pitch discrimination (e.g., Moreno et al., 2009). It is thus imperative not to erroneously interpret – as sometimes has happened (e.g., Habibi, Cahn, Damasio, & Damasio, 2016; Tierney et al., 2015) – functional neural changes in brain areas involved in domain-general cognitive abilities as evidence of cognitive enhancement. The same applies to anatomical neural changes (e.g., increased density of grey or white matter). Such patterns frequently observed in professional musicians and chess players are most likely neural correlates of their domain-specific expertise rather than superior overall cognitive ability.

2.3 Education and Skill Acquisition

In addition to theoretical aspects, the most obvious practical implications of these findings concern education. If skills rarely generalize across different domains, then the most effective way to acquire a skill is to train that particular skill. There are no shortcuts. Considering the insights provided by the research on expert performance and a variety of cognitive-training programs, educational and professional curricula should focus on discipline-related material rather than general principles without any specific reference to a particular subject (e.g., domain-general problem-solving skills). Moreover, the benefits of such domain-specific training should not be expected to generalize to other domains (e.g., learning Latin to improve logical thinking in mathematics).

Also, in line with the idea that training domain-general cognitive skills leads to benefits in a wide range of real-life skills, the last decade has seen the rise of a multibillion-dollar industry of commercial brain-training programs. Companies such as Posit Science® and Cogmed® claim that their training programs can help people in their daily, professional, and academic lives. However, in light of the results reported in this dissertation, the effectiveness of these programs remains doubtful (see also the discussion in Simons et al.,

2016). In brief, to date, no cognitive-training program has proved to be of any use in educational settings.

3. Methodological Considerations

The present broad meta-analytic investigation highlights the importance of an appropriate experimental design to test the alleged benefits of cognitive-training regimens. In fact, the type of allocation (random or non-random) and the type of control group (active or passive) have been found to be significant moderators in working memory training (Chapter 4) and music training (Chapter 6). In these two meta-analyses, the effect sizes tend to be smaller in the randomized samples than non-randomized samples. Furthermore, when the experimental groups are matched with active control groups, the overall effect sizes are smaller than the ones provided by the comparisons with the passive control groups. Thus, researchers must be extremely cautious when interpreting the results of treatment studies lacking random allocation to the groups or active control groups. Both the conditions are necessary to rule out possible between-group differences at baseline level and placebo effects.

Another important consideration concerns the conceptual difference between cognitive enhancement and performance in cognitive tasks. Cognitive-training programs have often been claimed to foster cognitive function because treated individuals usually show improvements in a few cognitive tests (e.g., Anguera et al., 2013; Diamond, Barnett, Thomas, & Munro, 2007). This inference is simply incorrect. The practice of cognitive tasks usually leads to a better performance in the trained task and similar tasks. However, such progress does not necessarily represent an enhancement in overall cognitive function. Shipstead et al. (2012) have explained this difference in the case of working memory training. The fact that individuals improve in a broad set of working memory and short-term memory tasks may be

due to amelioration in some general ability to perform working memory tasks rather than an enhanced working memory capacity. Thus, multivariate measures of a particular cognitive construct are not enough to test the capacity of a cognitive-training program to foster overall cognitive function. Rather, structural equation modelling is necessary to verify whether the improvements in the single cognitive tasks occur through a common factor that is measurement invariant across treatment and control groups (for details, see 4. Discussion in Chapter 4). Regrettably, such experimental design is nearly absent in cognitive-training studies.

Finally, this dissertation highlights the necessity of powerful and sophisticated meta-analytic models for a proper assessment of the effectiveness of experimental treatments in psychology. In an influential article about false-positive findings in psychology, Simmons, Nelson, and Simonsohn (2011) have shown that researchers are more likely to incur in Type I error than Type II error. This is because the flexibility in data collection and statistical analysis often allows one to falsely obtain significant results ($p < .05$). If primary studies usually report significant effects, basic meta-analytic models are bound to provide positive, yet probably biased, results as well. Also, the systematic suppression of non-significant findings (i.e., publication bias) is a further source of bias. It is thus imperative to include a proper sensitivity analysis to test the reliability of the overall results (Kepes & McDaniel, 2015). Moreover, specific formulas must be used to calculate the effect size (Chapter 3). The use of some alternative formulas risks to artificially inflate effect sizes and increase the overall effect (e.g., Chapter 8). Notably, due to design-related flaws, several previous meta-analyses have failed to find the true effect of particular cognitive-training programs. Examples include Au et al.'s (2015) meta-analysis about working memory training, Powers et

al.'s (2013) meta-analysis about video-game training, and the meta-analysis examining the effects of exergames on cognitive function by Stanmore et al. (2017).

4. Generalizability of the Results

This dissertation has documented the failure of several types of cognitive training to obtain any far-transfer effect. Although a large portion of the field has been covered, a comprehensive review of all the activities aimed to increase cognitive ability (or slow cognitive decline) is beyond the scope of this dissertation. Thus, the findings (and relevant implications) do not extend necessarily to several other training programs or particular sub-populations.

First, almost no clinical sample has been included. The only exception is the meta-analysis about exergames (Chapter 8) and two studies in the meta-analysis about chess instruction (Chapter 5). Thus, the results presented in this dissertation do not generalize necessarily to clinical populations. In fact, cognitive-training regimens that fail to enhance overall cognitive ability in individuals with normal cognitive function may provide some benefits in populations with impaired cognitive ability or learning disabilities.

Second, this dissertation does not deal with interventions aimed to optimise cognitive function. Examples include lifestyle modifications (e.g., Stine-Morrow & Basak, 2011), health and fitness activities (e.g., Voss, Vivar, Kramer, & van Praag, 2013), and nutrition and drugs (e.g., Burkhalter & Hillman, 2011). Thus, the results reported in the meta-analyses do not apply to these types of interventions.

Third, all those educational programs that aim to teach thinking skills have not been reviewed either. However, given the substantial absence of far-transfer effects documented in the present dissertation, it is extremely improbable that such programs may exert any meaningful impact on cognitive abilities or academic achievement. In line with this

conclusion, a recent meta-analysis (Abrami et al., 2015) showed modest effects of critical-thinking programs on a variety of school subjects.

Finally, the only variables examined in the meta-analyses are cognitive skills and academic-achievement measures. It is probably a trivial point. However, it must be understood that even if cognitive-training programs do not enhance cognition or academic attainment, it is still possible that such activities provide other types of benefits (e.g., enhanced prosocial skills, well-being, etc.). Whether cognitive-training programs offer such advantages will be a requirement for future research.

5. Conclusions and Future Research Directions

The meta-analytic reviews presented in this dissertation strongly suggest that the optimism about the far-transfer effects of cognitive training is not justified. Rather, converging evidence supporting Thorndike and Woodworth's (1901) common elements theory comes from the research on a broad range of types of cognitive training (e.g., WM training, chess, music, video games, exergames, spatial training, and brain training).

Future interventions trying to obtain far-transfer effects should strive for an experimental design including pre-tests and at least two control groups (a do-nothing group and an active control group). Such a design is the minimum standard in order to evaluate whether the putative benefits of cognitive training are genuine and not produced by statistical artefacts (e.g., differences at baseline level) and non-specific factors (e.g., placebo effects, expectations). Also, the use of multivariate measures of the same construct is recommended to test whether a particular training is enhancing a cognitive construct or just the ability to perform a class of cognitive tasks. Another central aim is to identify the specific characteristics of the training that might improve one's cognitive ability, which abilities they boost, and why these abilities should foster other non-trained abilities (i.e., far transfer).

Nonetheless, given the scarceness of far transfer in the literature, my hypothesis is that future experiments will show findings in line with those presented in this dissertation, at least with regard to non-clinical populations. For this reason, researchers and policymakers should seriously consider stopping spending resources for this type of research. Rather than searching for a way to improve overall domain-general cognitive ability, the field should focus on clarifying the domain-specific cognitive correlates underpinning expert performance.

References

- Abernethy, B., Neal, R. J., & Konig, P. (1994). Visual-perceptual and cognitive differences between expert, intermediate and novice snooker players. *Applied Cognitive Psychology, 18*, 185-211.
- Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research, 85*, 275-314.
- Aciego, R., Garcia, L., & Betancort, M. (2012). The benefits of chess for the intellectual and socio-emotional enrichment in schoolchildren. *The Spanish Journal of Psychology, 15*, 551-559.
- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131*, 30-60.
- Ackerman, P. L., Kanfer, R., & Calderwood, C. (2010). Use it or lose it? Wii brain exercise practice and reading for domain knowledge. *Psychology and Aging, 25*, 753-766.
- Adelson, B. (1981). Problem solving and the development of abstract categories in programming languages. *Memory & Cognition, 9*, 422-433.
- Allard, F., Graham, S., & Paarsalu, M. E. (1980). Perception in sport: Basketball. *Journal of Sport Psychology, 2*, 22-30.
- Allard, F., & Starkes, J. L. (1980). Perception in sport: Volleyball. *Journal of Sport Psychology, 2*, 22-33.
- Anderson, J. R. (1990). *Cognitive psychology and its implication (3rd Ed.)*. New York: Freeman.

- Anderson-Hanley, C., Arciero, P. J., Brickman, A. M., Nimon, J. P., Okuma, N., Westen, S. C., ... Zimmerman, E. A. (2012). Exergaming and older adult cognition: A cluster randomized clinical trial. *American Journal of Preventive Medicine, 42*, 109-119.
- Ang, S. Y., Lee, K., Cheam, F., Poon, K., & Koh, J. (2015). Updating and working memory training: Immediate improvement, long-term maintenance, and generalisability to non-trained tasks. *Journal of Applied Research in Memory and Cognition, 4*, 121-128.
- 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.
- Anvari, S. H., Trainor, L. G., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology, 83*, 111-130.
- Archibald, L. M. D., & Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language & Communication Disorders, 41*, 675-693.
- 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-Lervag and Hulme (2015). *Psychonomic Bulletin & Review, 23*, 331-337.
- Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: A meta-analysis. *Psychonomic Bulletin & Review, 22*, 366-377.
- Baddeley, A. (1992). Working memory. *Science, 255*, 556-559.
- Barcelos, N., Shah, N., Cohen, K., Hogan, M. J., Mulkerrin, E., Arciero, P. J., ... Anderson-Hanley, C. (2015). Aerobic and Cognitive Exercise (ACE) pilot study for older adults:

- executive function improves with cognitive challenge while exergaming. *Journal of the International Neuropsychological Society*, 21, 768-779.
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128, 612-637.
- Bart, W. M. (2014). On the effect of chess training on scholastic achievement. *Frontiers in Psychology*, 5:762.
- Basak, C., Boot, W. R., Voss, M. W., & Kramer, A. F. (2008). Can training in a real-time strategy videogame attenuate cognitive decline in older adults? *Psychology and Aging*, 23, 765-777.
- Bavelier, D., Green, C. S., Pouget, A., Schrater, P. (2012). Brain plasticity through the life span: learning to learn and action video games. *Annual Review of Neuroscience*, 35, 391-416.
- Becker, B. J. (2005). Failsafe *N* or file drawer number. In H. R. Rothstein, A. J., Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. Chichester, UK: John Wiley.
- Begg, C. B., & Berlin, J. A. (1988). Publication bias: A problem in interpreting medical data. *Journal of the Royal Statistical Society*, 151, 419-463.
- Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank order correlation for publication bias. *Biometrics*, 50, 1088-1101.
- Bejjanki, V. R., Zhang, R., Li, R., Pouget, A., Green, C. S., Lu, Z. L., & Bavelier, D. (2014). Action video game play facilitates the development of better perceptual templates. *Proceedings of the National Academy of Sciences*, 111, 16961-16966.

- Benson, C. (2006). How training chess affects children's scores on academic aptitude tests: A meta-analysis. In T. Redman (Ed.), *Chess and education. Selected essays from the Koltanowski Conference* (pp. 144-160). Dallas, TX: University of Texas at Dallas.
- Bermudez, P., Lerch, J. P., Evans, A. C., & Zatorre, R. J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cerebral Cortex, 19*, 1583-1596.
- Bialystok, E., & Depape, A. M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 565-574.
- Bijmolt, T. H. A., & Pieters, R. G. M. (2001). Meta-analysis in marketing when studies contain multiple measurements. *Marketing Letters, 12*, 157-169.
- Bilalić, M., McLeod, P., & Gobet, F. (2007). Does chess need intelligence? A study with young chess players. *Intelligence, 35*, 457-470.
- Bilalić, M., McLeod, P., & Gobet, F. (2009). Specialization effect and its influence on memory and problem solving in expert chess players. *Cognitive Science, 33*, 1117-1143.
- Binev, S., Attard-Montalto, J., Deva, N., Mauro, M., & Takkula, H. (2011). Declaration of the European Parliament, 0050/2011.
- Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video games improve perception and cognition? *Frontiers in Psychology, 2*, 226.
- Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129*, 387-398.

- Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science, 8*, 445-454.
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. London: Wiley.
- Bowels, S. M. (2003). *Tune up the mind: The effect of orchestrating music as a reading intervention*. Doctoral dissertation, Indiana University of Pennsylvania, PA: Indiana.
- Brousseau, G. (1997). *Theory of didactical situations in mathematics*. Dordrecht: Kluwer Academic Publishers.
- Burgoyne, A. P., Sala, G., Gobet, F., Macnamara, B. N., Campitelli, G., & Hambrick, D. Z. (2016). The relationship between cognitive ability and chess skill: A comprehensive meta-analysis. *Intelligence, 59*, 72-83.
- Burkhalter, T. M., & Hillman, C. H. (2011). A narrative review of physical activity, nutrition, and obesity to cognition and scholastic performance across the human lifespan. *Advances in Nutrition, 2*, 201S-206S.
- Buschkuehl, M., Jaeggi, S. M., & Jonides, J. (2012). Neuronal effects following working memory training. *Developmental Cognitive Neuroscience, 2*(Suppl. 1), S167-S179.
- Bühren, C., & Frank, B. (2010). Chess players' performance beyond 64 squares: A case study on the limitations of cognitive abilities transfer. *MAGKS Joint Discussion Paper Series in Economics*, 19-2010.
- Cardarelli, D. M. (2003). *The effects of music instrumental training on performance on the reading and mathematics portions of the Florida Comprehensive Achievement test for third grade students*. Doctoral dissertation, University of Central Florida, FL: Orlando.

- Castel, A., Pratt, J., & Drummond, E. (2005). The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. *Acta Psychologica, 119*, 217-230.
- Chacko, A., Bedard, A. C., Marks, D. J., Feirsen, N., Uderman, J. Z., Chimiklis, A., . . . Ramon, M. (2014). A randomized clinical trial of Cogmed Working Memory Training in school-age children with ADHD: A replication in a diverse sample using a control condition. *Journal of Child Psychology and Psychiatry, 55*, 247-255.
- Chan, C. L., Ngai, E. K., Leung, P. K. H., & Wong, S. (2010). Effect of the adapted virtual reality cognitive training program among Chinese older adults with chronic schizophrenia: A pilot study. *International Journal Geriatric Psychiatry, 25*, 643-649.
- Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 1-58). New York: Academic.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55-81.
- Cheek, J. M., & Smith, L. R. (1999). Music training and mathematics achievement. *Adolescence, 34*, 759-761.
- Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review, 17*, 193-199.
- Cheng, Y. L., & Mix, K. S. (2014). Spatial training improves children's mathematics ability. *Journal of Cognition and Development, 15*, 2-11.
- Cheung, S. F., & Chan, D. K. (2004). Dependent effect sizes in meta-analysis: Incorporating the degree of interdependence. *Journal of Applied Psychology, 89*, 780-791.
- Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73-96). Hillsdale, NJ: Erlbaum.

- Clark, C. M., Lawlor-Savage, L., & Goghari, V. M. (2017). Functional brain activation associated with working memory training and transfer. *Behavioral Brain Research, 334*, 34-49.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum: Hillsdale; NJ.
- Colom, R., Quiroga, M. Á., Solana, A. B., Burgaleta, M., Román, F. J., Privado, J., . . . Karama, S. (2012). Structural changes after videogame practice related to a brain network associated with intelligence. *Intelligence, 40*, 479-489.
- Comprehensive Meta-Analysis (Version 3.3) [Computer software] (2014r). Englewood, NJ: Biostat. Available from <http://www.meta-analysis.com>
- “A consensus on the brain training industry from the scientific community.” (2014). Retrieved from <http://longevity3.stanford.edu/blog/2014/10/15/the-consensus-on-the-brain-training-industry-from-the-scientific-community-2/>
- Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review, 8*, 331-335.
- Conway, A. R. A., & Engle, R. W. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory, 4*, 577-590.
- Costa, A., & Kallick, B. (2009) (Eds.). *Learning and leading with habits of mind: 16 essential characteristics for success*. Alexandria (VA): Association for Supervision & Curriculum Deve.
- Costandi, M. (2016). *Want to 'train your brain'? Forget apps, learn a musical instrument*. Retrieved from <https://www.theguardian.com/education/2016/oct/24/want-to-train-your-brain-forget-apps-learn-a-musical-instrument>

- De Corte, E., & Verschaffel, L. (1986). Effects of computer experience on children's thinking skills. *Journal of Structural Learning, 9*, 161-174.
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*, 13-21.
- Degé, F., Wehrum, S., Stark, R., & Schwarzer, G. (2011). The influence of two years of school music training in secondary school on visual and auditory memory. *European Journal of Developmental Psychology, 8*, 608-623.
- Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*, 1387-1388.
- Doll, J., & Mayr, U. (1987). Intelligenz und Schachleistung: Eine Untersuchung an Schachexperten [Intelligence and performance in chess: A study of chess expert]. *Psychologische Beiträge, 29*, 270-289.
- Donovan, M. S., Bransford, J. D., & Pellegrino, J. W. (1999). *How people learn: Bridging research and practice*. Washington, DC: National Academies Press.
- Dougherty, M. R., Hamovits, T., & Tidwell, J. W. (2016). Reevaluating the effect of n-back training on transfer through the Bayesian lens: Support for the null. *Psychonomic Bulletin & Review, 23*, 306-316.
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel plot based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*, 276-284.
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629-634.
- Ehrlich, S., Levine, S., & Goldin-Meadow, S. (2006). The importance of gestures in children's spatial reasoning. *Developmental Psychology, 42*, 1259-1268.

- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309-331.
- Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, *49*, 725-747.
- Fabel, K., Wolf, S., Ehninger, D., Babu, H., Galicia, P., & Kempermann, G. (2009). Additive effects of physical exercise and environmental enrichment on adult hippocampal neurogenesis in mice. *Frontiers in Neuroscience*, *3*, 50.
- Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science*, *18*, 850-855.
- Feuerstein, R., Feuerstein, R. S., Falik, L., & Rand, Y. (2006). *The Feuerstein instrumental enrichment program*. Jerusalem: ICELP Publications.
- Firth, J., Stubbs, B., Rosenbaum, S., Vancampfort, D., Malchow, B., Schuch, F., ... Yung, A. R. (2016). Aerobic exercise improves cognitive functioning in people with schizophrenia: A systematic review and meta-analysis. *Schizophrenia Bulletin*, *43*, 546-556.
- Forgeard, M., Schlaug, G., Norton, A., Rosam, C., Iyengar, U., & Winner, E. (2008). The relation between music and phonological processing in normal-reading children and children with dyslexia. *Music Perception*, *25*, 383-390.
- Forrest, D., Davidson, I., Shucksmith, J., & Glendinning, T. (2005). *Chess development in Aberdeen's primary schools: A study of literacy and social capital*. Aberdeen, Scotland: University of Aberdeen. Retrieved from: www.gov.scot/Resource/Doc/930/0009711.pdf
- Frydman, M., & Lynn, R. (1992). The general intelligence and spatial abilities of gifted young Belgian chess players. *British Journal of Psychology*, *83*, 233-235.

Garner, R. (2012). Chess makes a dramatic comeback in primary schools.

<http://www.independent.co.uk/news/education/education-news/chess-makes-a-dramatic-comeback-in-primary-schools-8301313.html>

Geoghegan, N., & Mitchelmore, M. (1996). Possible effects of early childhood music on mathematical achievement. *Journal for Australian Research in Early Childhood Education, 1*, 57-64.

George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia, 49*, 1083-1094.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306-355.

Gobet, F. (2011). *Psychologie du talent et de l'expertise* [Psychology of talent and expertise]. Bruxelles: De Boeck.

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

Gobet, F., & Campitelli, G. (2002). Intelligence and chess. In J. Retschitzki & R. Haddad-Zubel (Eds.), *Step by step. Proceedings of the 4th colloquium board games in academia* (pp. 103-112). Fribourg (Switzerland): Editions Universitaires.

Gobet, F., & Campitelli, G. (2006). Educational benefits of chess instruction. A critical review. In T. Redman (Ed.), *Chess and education. Selected essays from the Koltanowski Conference* (pp. 124-143). Dallas, TX: University of Texas at Dallas.

Gobet, F., de Voogt, A. J., & Retschitzki, J. (2004). *Moves in mind*. Hove, UK: Psychology Press.

- Gobet, F., Johnston, S. J., Ferrufino, G., Johnston, M., Jones, M. B., Molyneux, A., . . . Weeden, L. (2014). "No level up!": No effects of video game specialization and expertise on cognitive performance. *Frontiers in Psychology, 5*, 1337.
- Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31*, 1-40.
- Gobet, F., & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science, 24*, 651-682.
- Gobet, F., & Waters, A. J. (2003). The role of constraints in expert memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 29*, 1082-1094.
- Gong, D., He, H., Ma, W., Liu, D., Huang, M., Dong, L., . . . Yao, D. (2016). Functional integration between salience and central executive networks: A role for action video game experience. *Neural Plasticity, 2016*, 9803165.
- Gordon, R. L., Fehd, H. M., & McCandliss, B. D. (2015). Does music training enhance literacy skills? A meta-analysis. *Frontiers in Psychology, 6*, 1777.
- Grabner, R. H., Stern, E., & Neubauer, A. C. (2007). Individual differences in chess expertise: A psychometric investigation. *Acta Psychologica, 124*, 398-420.
- Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6*, 316-322.
- Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*, 534-537.
- Green, C. S., Gorman, T., & Bavelier, D. (2016). Action video-game and its effects on perception and attentional control. In T. Strobach, & J. Karbach (Eds.), *Cognitive training: An overview of features and applications* (pp. 107-116). New York: Springer.

- Green, C. S., Kattner, F., Eichenbaum, A., Bediou, B., Adams, D. M., Mayer, R. E., & Bavelier, D. (2017). Playing some video games but not others is related to cognitive abilities: A critique of Unsworth et al. (2015). *Psychological Science*, *28*, 679-682.
- Green, C. S., Li, R., & Bavelier, D. (2010). Perceptual learning during action video game playing. *Topics in Cognitive Science*, *2*, 202-216.
- Green, C. S., Pouget, A., & Bavelier, D. (2010). Improved probabilistic inference as a general learning mechanism with action video games. *Current Biology*, *20*, 1573-1579.
- Grek, S. (2009). Governing by numbers: The PISA 'effect' in Europe. *Journal of Education Policy*, *24*, 23-37.
- Guerin, B., & Matthews, A. (1990). The effects of semantic complexity on expert and novice computer program recall and comprehension. *The Journal of General Psychology*, *117*, 379-389.
- Gurtner, J. L., Gex, C., Gobet, F., Núñez, R., & Restchitzki, J. (1990). La récursivité rend-elle l'intelligence artificielle? [Does recursion make intelligence artificial?] *Revue Suisse de Psychologie*, *49*, 17-26.
- Habibi, A., Cahn, B. R., Damasio, A., & Damasio, H. (2016). Neural correlates of accelerated auditory processing in children engaged in music training. *Developmental Cognitive Neuroscience*, *21*, 1-14.
- Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences*, *11*, 236-242.
- Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, *8*, 1-51.
- Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors

- of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology*, 24, 1149-1167.
- Hanushek, E. A., Peterson, P. E., & Woessmann, L. (2012). *Achievement growth: International and US state trends in student performance*. Harvard's Program on Education Policy and Governance.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Hawes, Z., Moss, J., Caswell, B., & Poliszczuk, D. (2015). Effects of mental rotation training on children's spatial and mathematics performance: A randomized controlled study. *Trends in Neuroscience and Education*, 4, 60-68.
- Hänggi, J., Brüttsch, K., Siegel, A. M., & Jäncke, L. (2014). The architecture of the chess player's brain. *Neuropsychologia*, 62, 152-162.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Hetland, L., & Winner, E. (2001). The arts and academic achievement: What the evidence shows. *Arts Education Policy Review*, 102, 3-6.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analysis. *British Medical Journal*, 327, 557-560.
- Hornung, C., Brunner, M., Reuter, R. A. P., & Martin, R. (2011). Children's working memory: Its structure and relationship to fluid intelligence. *Intelligence*, 39, 210-221.
- Houlahan, M., & Tacka, P. (2015). *Kodály Today: A Cognitive Approach to Elementary Music Education Inspired by the Kodály Concept*. New York, NY: Oxford University Press.

- Hsi, S., Linn, M. C., & Bell, J. E. (1997). The role of spatial reasoning in engineering and the design of spatial instruction. *Journal of Engineering Education, 86*, 151-158.
- Hsu, J. (2010). *Navy says video games can boost "fluid intelligence" of warfighters*. Retrieved from <http://www.popsci.com/technology/article/2010-01/navy-says-video-games-can-boost-fluid-intelligence-warfighters>
- Hughes, T. F., Flatt, J. D., Fu, B., Butters, M. A., Chang, C. C. H., & Ganguli, M. (2014). Interactive video gaming compared with health education in older adults with mild cognitive impairment: A feasibility study. *International Journal Geriatric Psychiatry, 29*, 890-898.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72-98.
- Hunter, J. E., Schmidt, F. L., & Le, H. (2006). Implications of direct and indirect range restriction for meta-analysis methods and findings. *Journal of Applied Psychology, 91*, 594-612.
- Irons, J. L., Remington, R. W., & McLean, J. P. (2011). Not so fast: Rethinking the effects of action video games on attentional capacity. *Australian Journal of Psychology, 63*, 224-231.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 6829-6833.
- James, W. (1890). *The principles of psychology*. New York, NY: Henry Holt and Co.
- Jaušovec, N., & Pahor, A. (2017). *Boost your IQ with music*. Retrieved from <http://scitechconnect.elsevier.com/boost-your-iq-with->

[music/?sf67642888=1#utm_source=socialmedia&utm_medium=psych&utm_campaign=Boost%20Your%20IQ%20With%20Music](https://www.educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/EF_Project_Report_Chess_in_Schools.pdf)

Jerrim, J., Macmillan, L., Micklewright, J., Sawtell, M., & Wiggins, M. (2016). Chess in schools. Evaluation report and executive summary. *Education Endowment Foundation*.

Retrieved from

https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/EF_Project_Report_Chess_in_Schools.pdf

Johnson, M. H., Munakata, Y., & Gilmore, R. O. (Eds.). (2002). *Brain development and cognition: A reader* (2nd ed.). Oxford, England: Blackwell.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*, 47-70.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 66-71.

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

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

Kaviani, H., Mirbaha, H., Pournaseh, M., & Sagan, O. (2014). Can music lessons increase the performance of preschool children in IQ tests? *Cognitive Processing*, *15*, 77-84.

- Kazemi, F., Yektayar, M., & Abad, A. M. B. (2012). Investigation the impact of chess play on developing meta-cognitive ability and math problem-solving power of students at different levels of education. *Procedia – Social and Behavioral Sciences*, *32*, 372-379.
- Kempermann, G., Fabel, K., Ehninger, D., Babu, H., Leal-Galicia, P., Garthe, A., & Wolf, S. (2010). Why and how physical activity promotes experience-induced brain plasticity. *Frontiers in Neuroscience*, *4*, 189.
- Kepes, S., & McDaniel, M. A. (2015). The validity of conscientiousness is overestimated in the prediction of job performance. *PLoS ONE*, *10*, e0141468.
- Kim, Y. W., Kang, D. W., Kim, D., Kim, H. J., Sasaki, Y., & Watanabe, T. (2015). Real-time strategy video game experience and visual perceptual learning. *The Journal of Neuroscience*, *35*, 10485-10492.
- Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*, 317-324.
- Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlstrom, K., . . . Westerberg, H. (2005). Computerized training of working memory in children with ADHD – A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*, 177-186.
- Knecht, M. G. (2003). Music expertise and memory: the relationship between music expertise and memory of music patterns, within various degrees of contextual constraint. *Music Education Research*, *5*, 227-242.
- Koelsch, S., Fritz, T., von Cramon, D. Y., Müller, K., & Friederici, A. D. (2006). Investigating emotion with music: An fMRI study. *Human Brain Mapping*, *27*, 239-250.
- Lauer J. E., & Lourenco, S. F. (2016). Spatial processing in infancy predicts both spatial and mathematical aptitude in childhood. *Psychological Science*, *27*, 1-8.

- Lee, S. E. (2014). *The impact of working memory training on third grade students' reading fluency and reading comprehension performance*. Doctoral dissertation, Southern Illinois University, IL: Carbondale.
- Lee, Y., Lu, M., & Ko, H. (2007). Effect of skill training on working memory capacity. *Learning and Instruction, 17*, 336-344.
- Li, X., & Atkins, M. S. (2004). Early childhood computer experience and cognitive and motor development. *Pediatrics, 113*, 1715-1722.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Lynn, R., Wilson, R. G., & Gault, A. (1989). Simple musical tests as measures of Spearman's ρ . *Personality and Individual Differences, 10*, 25-28.
- Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions a meta-analysis. *Psychological Science, 25*, 1608-1618.
- Maillot, P., Perrot, A., & Hartley, A. (2012). Effects of interactive physical-activity video-game training on physical and cognitive function in older adults. *Psychology and Aging, 27*, 589-600.
- Margulies, S. (1992). *The effect of chess on reading scores: District Nine chess program; Second year report*. New York, NY: The American Chess Foundation.
- Mehr, S. A., Schachner, A., Katz, R. C., & Spelke, E. S. (2013). Two randomized trials provide no consistent evidence for nonmusical cognitive benefits of brief preschool music enrichment. *PLoS ONE, 8*.
- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*, 270-291.

- 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.
- 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.
- Mestre, J. P. (2005). *Transfer of learning from a modern multidisciplinary perspective*. Greenwich, CT: Information Age.
- Miendlarzewska, E. A., & Trost, W. J. (2013). How musical training affects cognitive development: Rhythm, reward and other modulating variables. *Frontiers in Neuroscience*, *7*, 279.
- Miller, D. I., & Halpern, D. F. (2013). Can spatial training improve long-term outcomes for gifted STEM undergraduates? *Learning and Individual Differences*, *26*, 141-152.
- Minear, M., Brasher, F., Guerrero, C. B., Brasher, M., Moore, A., & Sukeena, J. (2016). A simultaneous examination of two forms of working memory training: Evidence for near transfer only. *Memory & Cognition*, *44*, 1014-1037.
- Mirelman, A., Rochester, L., Maidan, I., Del Din, S., Alcock, L., Nieuwhof, F., ... Hausdorff, J. M. (2016). Addition of a non-immersive virtual reality component to treadmill training to reduce fall risk in older adults (V-TIME): A randomised controlled trial. *Lancet*, *388*, 1170-1182.
- Mix, K. S., Levine, S. C., Cheng, Y. L., Young, C., Hambrick, D. Z., Ping, R., & Konstantopoulos, S. (2016). Separate but correlated: The latent structure of space and mathematics across development. *Journal of Experimental Psychology: General*, *145*, 1206-1227.

- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*, 264-269.
- Moreau, D., Kirk, I. J., & Waldie, K. E. (2016). Seven pervasive statistical flaws in cognitive training interventions. *Frontiers in Human Neuroscience, 10*, 153.
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex, 19*, 712-723.
- Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review, 18*, 46-60.
- Mosing, M. A., Madison, G., Pedersen, N. L., & Ullén F. (2016). Investigating cognitive transfer within the framework of music practice: Genetic pleiotropy rather than causality. *Developmental Science, 19*, 504-512.
- Murphy, K., & Spencer, A. (2009). Playing video games does not make for better visual attention skills. *Journal of Articles in Support of the Null Hypothesis, 6*, 1-20.
- Neumann, N., Lotze, M., & Eickhoff, S. B. (2016). Cognitive expertise: An ALE meta-analysis. *Human Brain Mapping, 37*, 262-272.
- Nicotera, A., & Stuit, D. (2014). *Literature review of chess studies*. Denver, CO: Basis Policy Research.
- Oei, A. C., & Patterson, M. D. (2013). Enhancing cognition with video games: A multiple game training study. *PLoS ONE, 8*, e58546.
- Oei, A. C., & Patterson, M. D. (2014). Playing a puzzle video game with changing requirements improves executive functions. *Computers in Human Behavior, 37*, 216-238.

- Oei, A. C., & Patterson, M. D. (2015). Enhancing perceptual and attentional skills requires common demands between the action video games and transfer tasks. *Frontiers in Psychology, 6*, 113.
- Okagaki, L., & Frensch, P. A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. *Journal of Applied Developmental Psychology, 15*, 33-58.
- Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8*, 157-159.
- Passolunghi, M. C. (2006). Working memory and arithmetic learning disability. In T. P. Alloway & S. E. Gathercole (Eds.), *Working memory and neurodevelopmental condition* (pp. 113-138). Hove, England: Psychology Press.
- Pellegrino J. W., & Hilton M. L. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: National Academy of Sciences.
- Peng, P., Namkung, J., Barnes, M., & Sun, C. Y. (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology, 108*, 455-473.
- Perkins, D. N., & Grotzer, T. A. (2000). Teaching intelligence. In R. J. Sternberg (Ed.), *Handbook of Intelligence* (pp. 492-515). Cambridge: Cambridge University Press.
- Perkins, D. N., & Salomon, G. (1994). Transfer of learning. In T. N. Postlethwaite & T. Husen (Eds.), *International Encyclopedia of Education* (pp. 6452-6457). Oxford: Elsevier.
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology, 61*, 991-996.

- Portowitz, A., Lichtenstein, O., Egorova, L., & Brand, E. (2009). Underlying mechanisms linking music education and cognitive modifiability. *Research Studies in Music Education, 31*, 107-128.
- Powers, K. L., Brooks, P. J. (2014). Evaluating the specificity of effects of video game training. In F. C. Blumberg (Ed.), *Learning by playing. Video gaming in education* (pp. 302-332). Oxford University Press, Oxford.
- Powers, K. L., Brooks, P. J., Aldrich, N. J., Palladino, M., & Alfieri, L. (2013). Effects of video-game play on information processing: A meta-analytic investigation. *Psychonomic Bulletin & Review, 20*, 1055-1079.
- Rapport, M. D., Orban, S. A., Kofler, M. J., & Friedman, L. M. (2013). Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? A meta-analytic review of cognitive, academic, and behavioral outcomes. *Clinical Psychology Review, 33*, 1237-1252.
- Raven, J. (1960). *Guide to the standard progressive matrices*. London: H. K. Lewis.
- Redick, T. S., Calvo, A., Gay, C. E., & Engle, R. W. (2011). Working memory capacity and go/no-go task performance: Selective effects of updating, maintenance, and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 308-324.
- Redick, T. S., Shipstead, Z., Wiemers, E. A., Melby-Lervag, M., & Hulme, C. (2015). What's working in working memory training? An educational perspective. *Educational Psychology Review, 27*, 617-633.
- Redick, T. S., Unsworth, N., Kane, M. J., Hambrick, D. Z. (2017). Don't shoot the messenger: Still no evidence that video-game experience is related to cognitive abilities – A reply to Green et al. (2017). *Psychological Science, 28*, 683-686.

- Richland, L. E., Stigler, J. W., & Holyoak, K. J. (2012). Teaching the conceptual structure of mathematics. *Educational Psychologist, 47*, 189-203.
- Rickard, N. S., Bambrick, C. J., & Gill, A. (2012). Absence of widespread psychosocial and cognitive effects of school-based music instruction in 10-13-year-old students. *International Journal of Music Education, 30*, 57-78.
- Rikers, R. M. J. P., Schmidt, H. G., & Boshuizen, H. P. A. (2002). On the constraints of encapsulated knowledge: Clinical case representations by medical experts and subexperts. *Cognition and Instruction, 20*, 27-45.
- Rode, C., Robson, R., Purviance, A., Geary, D. C., & Mayr, U. (2014). Is working memory training effective? A study in a school setting. *PLoS ONE, 9*.
- Roden, I., Kreutz, G., & Bongard, S. (2012). Effects of a school-based instrumental music program on verbal and visual memory in primary school children: A longitudinal study. *Frontiers in Psychology, 3*, 572.
- Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence, 35*, 83-92.
- Ruthsatz, J., Detterman, D., Griscom, W. S., & Cirullo, B. A. (2008). Becoming an expert in the musical domain: It takes more than just practice. *Intelligence, 36*, 330-338.
- Sala, G., Bolognese, M., & Gobet, F. (2017). Spatial training and mathematics: The moderating effect of handedness. *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, London, UK.
- Sala, G., Burgoyne, A. P., Macnamara, B. N., Hambrick, D. Z., Campitelli, G., & Gobet, F. (2017). Checking the “Academic Selection” argument. Chess players outperform non-chess players in cognitive skills related to intelligence: A meta-analysis. *Intelligence, 61*, 130-139.

- Sala, G., Foley, J. P., Gobet, F. (2017). The effect of chess instruction on pupils' cognitive and academic skills: State of the art and theoretical challenges. *Frontiers in Psychology*, 8, 238.
- Sala, G., & Gobet, F. (2016). Do the benefits of chess instruction transfer to academic and cognitive skills? A meta-analysis. *Educational Research Review*, 18, 46-57.
- Sala, G., & Gobet, F. (2017a). Experts' memory superiority for domain-specific random material generalizes across fields of expertise: A meta-analysis. *Memory & Cognition*, 45, 183-193.
- Sala, G., & Gobet, F. (2017b). Working memory training in typically developing children: A meta-analysis of the available evidence. *Developmental Psychology*, 53, 671-685.
- Sala, G., & Gobet, F. (2017c). When the music's over. Does music skill transfer to children's and young adolescents' cognitive and academic skills? A meta-analysis. *Educational Research Review*, 20, 55-67.
- Sala, G., & Gobet, F. (2017d). Does chess instruction improve mathematical problem-solving ability? Two experimental studies with an active control group. *Learning & Behavior*. Advanced online publication. <https://link.springer.com/article/10.3758/s13420-017-0280-3>
- Sala, G., & Gobet, F. (in press). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Direction in Psychological Science*.
- Sala, G., Gobet, F., Trincherro, R., & Ventura S. (2016). Does chess instruction enhance mathematical ability in children? A three-group design to control for placebo effects. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, Philadelphia: PA.
- Sala, G., Gorini, A., & Pravettoni, G. (2015). Mathematical problem solving abilities and chess: An experimental study on young pupils. *SAGE Open*, July-September, 5.
- Sala, G., & Tatlidil, K. S., & Gobet, F. (submitted-a). Video game training does not enhance

- cognitive ability: A comprehensive meta-analytic investigation. *Psychological Bulletin*.
- Sala, G., & Tatlidil, K. S., & Gobet, F. (submitted-b). Exergames do not improve cognitive ability: A commentary on Stanmore et al. (2017). *Neuroscience and Biobehavioral Reviews*.
- Sala, G., & Trincherò, R. (in preparation). *Is meta-cognition the link between chess training and improvement in mathematics? A study on primary school children*. Unpublished manuscript.
- Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist, 24*, 113-142.
- Schättin, A., Arner, R., Gennaro, F., & de Bruin, E. D. (2016). Adaptations of prefrontal brain activity, executive functions, and gait in healthy elderly following exergame and balance training: A randomized-controlled study. *Frontiers in Aging Neuroscience, 8*, 278.
- Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science, 15*, 511-514.
- Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology, 98*, 457-468.
- Schellenberg, E. G., Corrigan, K. A., Dys, S. P., & Malti, T. (2015). Group music training and children's prosocial skills. *PLoS ONE, 10*.
- Schellenberg, E. G., & Mankarious, M. (2012). Music training and emotion comprehension in childhood. *Emotion, 12*, 887-891.
- Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Newbury Park, CA: Sage.
- Schneider, W., Gruber, H., Gold, A., & Opwis, K. (1993). Chess expertise and memory for chess positions in children and adults. *Journal of Experimental Child Psychology, 56*, 328-349.

- Schoene, D., Lord, S. R., Delbaere, K., Severino, C., Davies, T. A., & Smith, S. T. (2013). A randomized controlled pilot study of home-based step training in older people using videogame technology. *PLoS One*, *8*, e57734.
- Scholz, M., Niesch, H., Steffen, O., Ernst, B., Loeffler, M., Witruk, E., & Schwarz, H. (2008). Impact of chess training on mathematics performance and concentration ability of children with learning disabilities. *International Journal of Special Education*, *23*, 138-148.
- Schwaighofer, M., Fischer, F., & Buhner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist*, *50*, 138-166.
- Shavelson, R. J., Yuan, K., Alonzo, A. C., Klingberg, T., & Andersson, M. (2008). On the impact of computerized cognitive training on working memory and fluid intelligence. In D. C. Berliner & H. Kuppermintz (Eds.), *Contributions of educational psychology to changing institutions, environments, and people* (pp. 1-11). New York, NY: Routledge.
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica*, *50*, 245-276.
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*, 628-654.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359-1366.
- Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. L. (2016). Do “brain-training” programs work? *Psychological Science in the Public Interest*, *17*, 103-186.

- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). p-Curve and effect size correcting for publication bias using only significant results. *Perspectives on Psychological Science*, *9*, 666-681.
- Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Better *P*-curves: Making *P*-curve analysis more robust to errors, fraud, and ambitious *P*-hacking, a Reply to Ulrich and Miller. *Journal of Experimental Psychology: General*, *144*, 1146-1152.
- Slater, J., Strait, D. L., Skoe, E., O'Connell, S., Thompson, E., & Kraus, N. (2014). Longitudinal effects of group music instruction on literacy skills in low-income children. *PLoS ONE*, *9*.
- Sloboda, J. A. (1976). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology*, *28*, 1-16.
- Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., & Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *NeuroImage*, *17*, 1613-1622.
- Sorby, S. A. (2009). Education research in developing 3-D spatial skills for engineering students. *International Journal of Science Education*, *31*, 459-480.
- Sorby, S. A. (2011). *Developing spatial thinking*. Independence: Cengage.
- Sorby, S. A., Casey, B., Veurink, N., & Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learning and Individual Differences*, *26*, 20-29.
- 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*. Advanced online publication. Retrieved from <https://link.springer.com/article/10.3758/s13423-016-1217-0>

- Söderqvist, S., & Bergman-Nutley, S. (2015). Working memory training is associated with long term attainments in maths and reading. *Frontiers in Psychology, 6*:1711.
- Staiano, A. E., Abraham, A. A., & Calvert, S. L. (2012). Competitive versus cooperative exergame play for African American adolescents' executive function skills: Short-term effects in a long-term training intervention. *Developmental Psychology, 48*, 337-342.
- Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods, 5*, 60-78.
- Stanmore, E., Stubbs, B., Vancampfort, D., de Bruin, E. D., & Firth, J. (2017). The effect of active video games on cognitive functioning in clinical and non-clinical populations: A meta-analysis of randomized controlled trials. *Neuroscience and Biobehavioral Reviews, 78*, 34-43.
- Stieff, M., & Uttal, D. (2015). How much can spatial training improve STEM achievement? *Educational Psychology Review, 27*, 607-615.
- Stine-Morrow, E. A. L., & Basak, C. (2011). Cognitive interventions. In K. W. Schaie & S. L. Willis (Eds.), *Handbook of the psychology of aging* (7th ed., pp. 153-170). New York, NY: Elsevier.
- Stine-Morrow, E. A., Payne, B. R., Roberts, B. W., Kramer, A. F., Morrow, D. G., Payne, L., ... Parisi, J. M. (2014). Training versus engagement as paths to cognitive enrichment with aging. *Psychology & Aging, 29*, 891-906.
- Strobach, T., & Karbach, J. (2016). *Cognitive training: An overview of features and applications*. New York: Springer.
- "Studies on cognitive training benefits." (2014). Retrieved from <http://www.cognitivetrainingdata.org/studies-cognitivetraining-benefits>
- Swanson, H. L. (2006). Working memory and reading disabilities: Both phonological and

- executive processing deficits are important. In T. P. Alloway & S. E. Gathercole (Eds.), *Working memory and neurodevelopmental disorders* (pp. 59-88). Hove, England: Psychology Press.
- Taatgen, N. A. (2016). Theoretical models of training and transfer effects. In T. Strobach, & J. Karbach (Eds.), *Cognitive training: An overview of features and applications* (pp. 19-29). New York: Springer.
- Terlecki, M. S., Newcombe, N. S., & Little, M. (2008). Durable and generalized effects of spatial experience on mental rotation: Gender differences in growth patterns. *Applied Cognitive Psychology, 22*, 996-1013.
- Thorell, L. B., Lindqvist, S., Bergman, S., Bohlin, N. G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science, 12*, 106-113.
- Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. (I). *Psychological Review, 8*, 247-261.
- Tierney, A. T., Krizman, J., & Kraus, N. (2015). Music training alters the course of adolescent auditory development. *Proceedings of the National Academy of Sciences of the United States of America, 112*, 10062-10067.
- 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.
- Tracz, S. M., Elmore, P. B., & Pohlmann, J. T. (1992). Correlational meta-analysis: Independent and nonindependent cases. *Educational and Psychological Measurement, 52*, 879-888.
- Trinchero, R. (2012). *Gli scacchi, un gioco per crescere. Sei anni di sperimentazione nella scuola primaria* [Chess, a game to grow up with. Six year of research in primary school].

- Milan: Franco Angeli.
- Trinchero, R., & Sala, G. (2016). Chess training and mathematical problem-solving: The role of teaching heuristics in transfer of learning. *Eurasia Journal of Mathematics, Science & Technology Education, 12*, 655-668.
- Tukey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics, 33*, 1-67.
- Unsworth, N., Redick, T. S., McMillan, B. D., Hambrick, D. Z., Kane, M. J., & Engle, R. W. (2015). Is playing video games related to cognitive abilities? *Psychological Science, 26*, 759-774.
- Unterrainer, J. M., Kaller, C. P., Leonhart, R., & Rahm, B. (2011). Revising superior planning performance in chess players: The impact of time restriction and motivation aspects. *American Journal of Psychology, 124*, 213-225.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin, 139*, 352-402.
- van Ravenzwaaij, D., Boekel, W., Forstmann, B. U., Ratcliff, R., Wagenmakers, E. J. (2014). Action video games do not improve the speed of information processing in simple perceptual tasks. *Journal of Experimental Psychology: General, 143*, 1794-1805.
- Vaughn, K. (2000). Music and mathematics: Modest support for the oft-claimed relationship. *Journal of Aesthetic Education, 34*, 149-166.
- Vevea J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods, 10*, 428-443.
- Viechtbauer, W. (2010). Conducting meta-analysis in R with the metafor package. *Journal of Statistical Software, 36*, 1-48. Retrieved from

- <http://brieger.esalq.usp.br/CRAN/web/packages/metafor/vignettes/metafor.pdf>
- Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods, 1*, 112-125.
- Voss, M. W., Vivar, C., Kramer, A. F., & van Praag, H. (2013). Bridging animal and human models of exercise-induced brain plasticity. *Trends in Cognitive Sciences, 17*, 525-544.
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*, 817-835.
- Wang, P., Liu, H., Zhu, X., Meng, T., Li, H., Zuo, X. (2016). Action video game training for healthy adults: A meta-analytic study. *Frontiers in Psychology, 7*, 907.
- Waters, A. J., Gobet, F., & Leyden, G. (2002). Visuospatial abilities in chess players. *British Journal of Psychology, 30*, 303-311.
- Weicker, J., Villringer, A., & Thöne-Otto, A. (2016). Can impaired working memory functioning be improved by training? A meta-analysis with a special focus on brain injured patients. *Neuropsychology, 30*, 190-212.
- Wetter, O. E., Koerner, F., & Schwaninger, A. (2009). Does musical training improve music performance? *Instructional Science, 37*, 365-374.
- Williams, M., Davids, K., Burwitz, L., & Williams, J. (1993). Cognitive knowledge and soccer performance. *Perceptual and Motor Skills, 76*, 579-593.
- Witt, M. (2011). School based working memory training: Preliminary finding of improvement in children's mathematical performance. *Advanced in Cognitive Psychology, 7*, 7-15.
- Wu, S., Cheng, C. K., Feng, J., D'Angelo, L., Alain, C., & Spence, I. (2012). Playing a first-person shooter video game induces neuroplastic change. *Journal of Cognitive*

Neuroscience, 24, 1286-1293.

Xu, C., & LeFevre, J. A. (2016). Training young children on sequential relations among numbers and spatial decomposition: Differential transfer to number line and mental transformation tasks. *Developmental Psychology*, 52, 854-866.

Zimmermann, R., Gschwandtner, U., Benz, N., Hatz, F., Schindler, C., Taub, E., & Fuhr, P. (2014). Cognitive training in Parkinson disease: Cognition-specific vs nonspecific computer training. *Neurology*, 82, 1219-1226.

Appendix A. Studies Included in the Meta-Analysis in Chapter 4

- Bergman-Nutley, S. B., Söderqvist, S., Bryde, S., Thorell, L. B., Humphreys, K., & Klingberg, T. (2011). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: A controlled, randomized study. *Developmental Science, 14*, 591-601.
- Henry, L. A., Messer, D. J., & Nash, G. (2014). Testing for near and far transfer effects with a short, face-to-face adaptive working memory training intervention in typical children. *Infant and Child Development, 23*, 84-103.
- Horvat, M. (2014). Vpliv treninga delovnega spomina na kognitivne sposobnosti [The effect of working memory training on cognitive abilities] (Master's thesis). University of Maribor, Slovenia.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 10081-10086.
- Karbach, J., Strobach, T., & Schubert, T. (2015). Adaptive working-memory training benefits reading, but not mathematics in middle childhood. *Child Neuropsychology, 21*, 285-301.
- Kroesbergen, E. H., van't Noordende, J. E., & Kolkman, M. E. (2014). Training working memory in kindergarten children: Effects of on working memory and early numeracy. *Child Neuropsychology, 20*, 23-37.
- Kuhn, J., & Holling, H. (2014). Number sense or working memory? The effect of two computer-based trainings on mathematical skills in elementary school. *Advances in Cognitive Psychology, 10*, 59-67.
- Kun, Y. (2007). *Impact of computerized cognitive training on working memory, fluid intelligence, science achievement*. Doctoral dissertation, Stanford University, CA: Stanford.

- Lee, S. E. (2014). *The impact of working memory training on third grade students' reading fluency and reading comprehension performance*. Doctoral dissertation, Southern Illinois University, IL: Carbondale.
- Lindsay, J. R. (2012). *Everything I need to know I learned before elementary school: The relationships between phoneme awareness, letter knowledge, and working memory*. Doctoral dissertation, University of California, CA: Irvine.
- Loosli, S. V., Buschkuehl, M., Perrig, W. J., & Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychology, 18*, 62-78.
- Mansur-Alves, M., & Flores-Mendoza, C. (2015). Working memory training does not improve intelligence: Evidence from Brazilian children. *Psicologia: Reflexão e Crítica, 28*, 474-482.
- Mansur-Alves, M., Flores-Mendoza, C., & Tierra-Criollo, C. J. (2013). Evidências preliminares da efetividade do treinamento cognitivo para melhorar a inteligência de crianças escolares [Preliminary evidence of effectiveness in cognitive training to improve school children intelligence]. *Psicologia: Reflexão e Crítica, 26*, 423-434.
- Nevo, E., & Breznitz, Z. (2014). Effects of working memory and reading acceleration training on improving working memory abilities and reading skills among third graders. *Child Neuropsychology, 20*, 752-765.
- Passolunghi, M. C., & Costa, H. M. (2016). Working memory and early numeracy training in preschool children. *Child Neuropsychology, 22*, 81-98.
- Pugin, F., Metz, A. J., Stauffer, M., Wolf, M., Jenny, O. G., & Huber, R. (2015). Working memory training shows immediate and long-term effects on cognitive performance in children. *F1000 Research, 3*, 82.

- Rode, C., Robson, R., Purviance, A., Geary, D. C., & Mayr, U. (2014). Is working memory training effective? A study in a school setting. *PLoS ONE*, *9*.
- Shavelson, R. J., Yuan, K., Alonzo, A. C., Klingberg, T., & Andersson, M. (2008). On the impact of computerized cognitive training on working memory and fluid intelligence. In D. C. Berliner & H. Kuppermintz (Eds.), *Contributions of educational psychology to changing institutions, environments, and people* (pp. 1-11). New York, NY: Routledge.
- Söderqvist, S., & Bergman-Nutley, S. (2015). Working memory training is associated with long term attainments in maths and reading. *Frontiers in Psychology*, *6*:1711.
- St Clair-Thompson, H., Stevens R., Hunt, A., & Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology*, *30*, 203-219.
- Studer-Luethi, B., Bauer, C., & Perrig, W. J. (2016). Working memory training in children: Effectiveness depends on temperament. *Memory & Cognition*, *44*, 171-186.
- Thorell, L. B., Lindqvist, S., Bergman, S., Bohlin, N. G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, *12*, 106-113.
- Wang, Z., Zhou, R., & Shah, P. (2014). Spaced cognitive training promotes training transfer. *Frontiers in Human Neuroscience*, *8*:217.
- Witt, M. (2011). School based working memory training: Preliminary finding of improvement in children's mathematical performance. *Advanced in Cognitive Psychology*, *7*, 7-15.
- Zhao, X., Wang, Y., Liu, D., & Zhou, R. (2011). Effect of updating training on fluid intelligence in children. *Chinese Science Bulletin*, *56*, 2202-2205.

Appendix B. Studies Included in the Meta-Analysis in Chapter 5

- Aciego, R., Garcia, L., & Betancort, M. (2012). The benefits of chess for the intellectual and socio-emotional enrichment in schoolchildren. *The Spanish Journal of Psychology, 15*, 551-559.
- Aydin, M. (2015). Examining the impact of chess instruction for the visual impairment on mathematics. *Educational Research and Reviews, 10*, 907-911.
- Barrett, D. C., & Fish, W. W. (2011). Our move. Using chess to improve math achievement for students who receive special education services. *International Journal of Special Education, 26*, 181-193.
- Christiaen, J., & Verhofstadt-Denève, L. (1981). Schaken en cognitieve ontwikkeling [Chess and cognitive development]. *Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden, 36*, 561-582.
- DuCette, J. (2009). *An evaluation of the Chess Challenge Program of ASAP/After School Activities Partnerships*. Philadelphia, PA: After School Activities Partnerships.
- Eberhard, J. W. (2003). *The relationship between chess instruction and verbal, quantitative, and nonverbal reasoning abilities of economically disadvantage students*. (Unpublished doctoral dissertation). Texas A&M University, Corpus Christi, TX.
- Forrest, D., Davidson, I., Shucksmith, J., & Glendinning, T. (2005). *Chess development in Aberdeen's primary schools: A study of literacy and social capital*. Aberdeen, Scotland: University of Aberdeen. Retrieved from: www.gov.scot/Resource/Doc/930/0009711.pdf
- Fried, S., & Ginzburg, N. (n.d.). *The effect of learning to play chess on cognitive, perceptual and emotional development in children*. Unpublished manuscript.
- Garcia, N. V. (2008). *Scholastic chess club participation and the academic achievement of Hispanic fifth grade students in south Texas*. (Unpublished doctoral dissertation).

- University of Houston, Houston, TX.
- Gluga, F., & Flesner, P. I. (2014). Cognitive benefits of chess training in novice children. *Procedia – Social and Behavioral Sciences*, 116, 962-967.
- Hong, S., & Bart, W. M. (2007). Cognitive effects of chess instruction on students at risk for academic failure. *International Journal of Special Education*, 22, 89-96.
- Kazemi, F., Yektayar, M., & Abad, A. M. B. (2012). Investigation the impact of chess play on developing meta-cognitive ability and math problem-solving power of students at different levels of education. *Procedia – Social and Behavioral Sciences*, 32, 372-379.
- Kramer, A. & Filipp, S. (n.d.). *Chess at Tryer-Olewig Primary School: Summary and the evaluation of the outcomes of the German School Chess Foundation (short version)*. Unpublished manuscript.
- Margulies, S. (1992). *The effect of chess on reading scores: District Nine chess program; Second year report*. New York, NY: The American Chess Foundation.
- Rifner, P. J. (1992). *Playing chess: A study of the transfer of problem-solving skills in students with average and above average intelligence*. (Unpublished doctoral dissertation). Purdue University, West Lafayette, IN.
- Romano, B. (2011). *Does playing chess improve math learning? Promising (and inexpensive) results from Italy* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia, PA.
- Sala, G., Gobet, F., Trinchero, R., & Ventura S. (2016). Does chess instruction enhance mathematical ability in children? A three-group design to control for placebo effects. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, Philadelphia: PA.

- Sala, G., Gorini, A., & Pravettoni, G. (2015). Mathematical problem solving abilities and chess: An experimental study on young pupils. *SAGE Open, July-September, 5*.
- Sala, G., & Trincherò, R. (in preparation). *Is meta-cognition the link between chess training and improvement in mathematics? A study on primary school children*. Unpublished manuscript.
- Scholz, M., Niesch, H., Steffen, O., Ernst, B., Loeffler, M., Witruk, E., & Schwarz, H. (2008). Impact of chess training on mathematics performance and concentration ability of children with learning disabilities. *International Journal of Special Education, 23*, 138-148.
- Sigirtmac, A. D. (2012). Does chess training affect conceptual development of six-year-old children in Turkey? *Early Child Development and Care, 182*, 797-806.
- Trincherò, R. & Piscopo, M. (2007). *Gli scacchi, un gioco per crescere*. [Chess, a game to grow up with]. Unpublished manuscript.
- Trincherò, R., & Sala, G. (2016). Chess training and mathematical problem-solving: The role of teaching heuristics in transfer of learning. *Eurasia Journal of Mathematics, Science & Technology Education, 12*, 655-668.
- Yap, K. O. (2006). *Chess for success evaluation: Final report*. Portland, OR: Northwest Regional Educational Laboratory.

Appendix C. Studies Included in the Meta-Analysis in Chapter 6

- Bhide, A., Power, A., & Goswami, U. (2013). A rhythmic musical intervention for poor readers: A comparison of efficacy with a letter-based intervention. *Mind, Brain, and Education, 7*, 113-123.
- Bilhartz, T. D., Bruhn, R. A., & Olson, J. E. (1999). The effect of early music training on child cognitive development. *Journal of Applied Developmental Psychology, 20*, 615-636.
- Bowels, S. M. (2003). *Tune up the mind: The effect of orchestrating music as a reading intervention*. Doctoral dissertation, Indiana University of Pennsylvania, PA: Indiana.
- Cardarelli, D. M. (2003). *The effects of music instrumental training on performance on the reading and mathematics portions of the Florida Comprehensive Achievement test for third grade students*. Doctoral dissertation, University of Central Florida, FL: Orlando.
- Cogo-Moreira, H., de Avila, C. R. B., Ploubidis, G. B., & Mari, J. D. (2013). Effectiveness of music education for the improvement of reading skills and academic achievement in young poor readers: A pragmatic cluster-randomized, controlled clinical trial. *PLoS ONE, 8*.
- Costa-Giomi, E. (2004). Effects of three years of piano instruction on children's academic achievement, school performance and self-esteem. *Psychology of Music, 32*, 139-152.
- Degé, F., & Schwarzer, G. (2011). The effect of a music program on phonological awareness in preschoolers. *Frontiers in Psychology, 2*, 124.
- Degé, F., Wehrum, S., Stark, R., & Schwarzer, G. (2011). The influence of two years of school music training in secondary school on visual and auditory memory. *European Journal of Developmental Psychology, 8*, 608-623.
- Geoghegan, N., & Mitchelmore, M. (1996). Possible effects of early childhood music on mathematical achievement. *Journal for Australian Research in Early Childhood*

- Education, 1, 57-64.*
- Gromko, J. E. (2005). The effect of music instruction on phonemic awareness in beginning readers. *Journal of Research in Music Education, 53*, 199-209.
- Gromko, J. E., & Poorman, A. S. (1998). The effect of music training on preschoolers' spatial-temporal task performance. *Journal of Research in Music Education, 46*, 173-181.
- Hanson, M. K. (2001). *An investigation of the effects of sequenced Kodaly literacy-based music instruction on the spatial reasoning skills of kindergarten students*. Unpublished manuscript.
- Herrera, L., Lorenzo, O., Defior, S., Fernandez-Smith, G., & Costa-Giomi, E. (2011). Effects of phonological and musical training on the reading readiness of native- and foreign-Spanish-speaking children. *Psychology of Music, 39*, 68-81.
- Hole, K. A., (2013). *The impact of an auditory training programme (The Listening Programme®) on the auditory processing and reading skills of mainstream school children*. Unpublished manuscript.
- Hunt, A. (2012). *“Have you brought your singing voice?” An investigation into whether a small group singing intervention can improve phonological discrimination in young children*. Doctoral dissertation, Cardiff University, UK: Cardiff.
- Janus, M., Lee, Y., Moreno, S., & Bialystok, E. (2016). Effects of short-term music and second-language training on executive control. *Journal of Experimental Child Psychology, 144*, 84-97.
- Kaviani, H., Mirbaha, H., Pournaseh, M., & Sagan, O. (2014). Can music lessons increase the performance of preschool children in IQ tests? *Cognitive Processing, 15*, 77-84.
- Legette, R. M. (1993). *The effect of a selected use of music instruction on the self-concept and academic achievement of elementary public school students*. Doctoral dissertation,

The Florida State University School of Music, FL: Tallahassee.

- Lu, D. T. (1986). *The effect of teaching music skills on the development of reading skills among first graders: An experimental study*. Doctoral dissertation, University of Washington, WA: Seattle.
- Mehr, S. A., Schachner, A., Katz, R. C., & Spelke, E. S. (2013). Two randomized trials provide no consistent evidence for nonmusical cognitive benefits of brief preschool music enrichment. *PLoS ONE*, *8*.
- Moreno, S., Friesen, D., & Bialystok, E. (2011). Effect of music training on promoting preliteracy skills: Preliminary causal evidence. *Music Perception*, *29*, 165-172.
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*, 712-723.
- Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J., & Wolf, M. (2013). Links between early rhythm skills, musical training, and phonological awareness. *Reading and Writing*, *26*, 739-769.
- Myant, M., Armstrong, W., & Healy, N. (2008). Can music make a difference? A small scale longitudinal study into the effects of music instruction in nursery on later reading ability. *Educational and Child Psychology*, *25*, 83-100.
- Portowitz, A., Lichtenstein, O., Egorova, L., & Brand, E. (2009). Underlying mechanisms linking music education and cognitive modifiability. *Research Studies in Music Education*, *31*, 107-128.
- Portowitz, A., Peppler, K. A., & Downton, M. (2014). In harmony: A technology-based music education model to enhance musical understanding and general learning skills. *International Journal of Music Education*, *32*, 242-260.

- Rauscher, F. H., & Zupan, M. A. (2000). Classroom keyboard instruction improves kindergarten children's spatial-temporal performance: A field experiment. *Early Childhood Research Quarterly, 15*, 215-228.
- Register, D. (2004). The effects of live music groups versus an educational children's television program on the emergent literacy of young children. *Journal of Music Therapy, 41*, 2-27.
- Rickard, N. S., Bambrick, C. J., & Gill, A. (2012). Absence of widespread psychosocial and cognitive effects of school-based music instruction in 10-13-year-old students. *International Journal of Music Education, 30*, 57-78.
- Roden, I., Grube, D., Bongard, S., & Kreutz, G. (2014). Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study. *Psychology of Music, 42*, 284-298.
- Roden, I., Konen, T., Bongard, S., Frankenberg, E., Friedrich, E. K., & Kreutz, G. (2014). Effects of music training on attention, processing speed and cognitive music abilities – Findings from a longitudinal study. *Applied Cognitive Psychology, 28*, 545-557.
- Roden, I., Kreutz, G., & Bongard, S. (2012). Effects of a school-based instrumental music program on verbal and visual memory in primary school children: A longitudinal study. *Frontiers in Psychology, 3*, 571.
- Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science, 15*, 511-514.
- Schellenberg, E. G., Corrigan, K. A., Dys, S. P., & Malti, T. (2015). Group music training and children's prosocial skills. *PLoS ONE, 10*.
- Slater, J., Strait, D. L., Skoe, E., O'Connell, S., Thompson, E., & Kraus, N. (2014). Longitudinal effects of group music instruction on literacy skills in low-income children. *PLoS ONE, 9*.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion, 4*, 46-64.

Tierney, A. T., Krizman, J., & Kraus, N. (2015). Music training alters the course of adolescent auditory development. *PNAS, 112*, 10062-10067.

Yazejian, N., & Peisner-Feinberg, E. S. Effects of a preschool music and movement curriculum on children's language skills. *NHSA Dialog, 12*, 327-341.

Appendix D. Studies Included in Chapter 7 (Meta-Analysis 1)

- Adams, D. M. (2013). *Exploring the effect of video game playing on static and dynamic spatial cognition skills* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3559770).
- Adams, D. M., Pilegard, C., & Mayer, R. E. (2016). Evaluating the cognitive consequences of playing Portal for a short duration. *Journal of Educational Computing Research, 54*, 173-195.
- Appelbaum, L. G., Cain, M. S., Darling, E. F., & Mitroff, S. R. (2013). Action video game playing is associated with improved visual sensitivity, but not alterations in visual sensory memory. *Attention, Perception, & Psychophysics, 75*, 1161-1167.
- Arthur, W., Strong, M. H., Jordan, J. A., Williamson, J. E., Shebilske, W. L., & Regian, J. W. (1995). Visual attention: Individual differences in training and predicting complex task performance. *Acta Psychologica, 88*, 3-23.
- Bailey, K. (2012). *What would my avatar do? Video games and risky decision making* (Doctoral dissertation). Iowa State University, Ames, Iowa, USA.
- Barlett, C. P., Vowels, C. L., Shanteau, J., Crow, J., & Miller, T. (2009). The effect of violent and non-violent computer games on cognitive performance. *Computers in Human Behavior, 25*, 96-102.
- Blacker, K. J., & Curby, K. M. (2013). Enhanced visual short-term memory in action video game players. *Attention, Perception, & Psychophysics, 75*, 1128-1136.
- Blacker, K. J., Curby, K. M., Klobusicky, E., & Chein, J. M. (2014). Effects of action video game training on visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 1992-2004.
- Bonny, J. W., Castaneda, L. M., & Swanson, T. (2016). *Using an international gaming*

- tournament to study individual differences in MOBA expertise and cognitive skills*. In Proceedings of the SIGCHI conference on human factors in computing systems. San Jose, CA, USA (pp. 3473-3484).
- 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.
- Bragg, H. R., Millard, H. A. T., Millard, R. P., Constable, P. D., & Freeman, L. J. (2016). Association of gender and specialty interest with video-gaming, three-dimensional spatial analysis, and entry-level laparoscopic skills in third-year veterinary students. *Journal of the American Veterinary Medical Association, 248*, 1414-1418.
- Cain, M. S., Landau, A. N., & Shimamura, A. P. (2012). Action video game experience reduces the cost of switching tasks. *Attention, Perception & Psychophysics, 74*, 641-647.
- Cain, M. S., Prinzmetal, W., Shimamura, A. P., & Landau, A. N. (2014). Improved control of exogenous attention in action video game players. *Frontiers in Psychology, 5*, 69.
- Cherney, I. D., Bersted, K., & Smetter, J. (2014). Training spatial skills in men and women. *Perceptual & Motor Skills: Learning & Memory, 119*, 82-99.
- Clemenson, G. D., & Stark, C. E. L. (2015). Virtual environmental enrichment through video games improves hippocampal-associated memory. *The Journal of Neuroscience, 35*, 16116-16125.
- Colom, R., Quiroga, M. Á., Solana, A. B., Burgaleta, M., Román, F. J., Privado, J., . . . Karama, S. (2012). Structural changes after videogame practice related to a brain network associated with intelligence. *Intelligence, 40*, 479-489.
- De Lisi, R., & Wolford, J. L. (2002). Improving children's mental rotation accuracy with computer game playing. *The Journal of Genetic Psychology, 163*, 272-282.

- Dean, D. H. (2009). Spatial visualization and the gender gap in videogame interest among young adults. *Young Consumers, 10*, 225-237.
- Dobrowolski, P., Hanusz, K., Sobczyk, B., Skorko, M., & Wiatrow, A (2015). Cognitive enhancement in video game players: The role of video game genre. *Computers in Human Behavior, 44*, 59-63.
- Donohue, S. E., Woldorf, M. G., & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics, 72*, 1120-1129.
- Ferguson, C. J., Cruz, A. M., & Rueda, S. M. (2008). Gender, video game playing habits and visual memory tasks. *Sex Roles, 58*, 279-286.
- Ferguson, C. J., Garza, A., Jerabeck J., Ramos, R., & Galindo, M. (2013). Not worth the fuss after all? Cross-sectional and prospective data on violent video game influences on aggression, visuospatial cognition and mathematics ability in a sample of youth. *Journal of Youth and Adolescence, 42*, 109-122.
- Foroughi, C. K., Serraino, C., Parasuraman, R., & Boehm-Davis, D. A. (2016). Can we create a measure of fluid intelligence using Puzzle Creator within Portal 2? *Intelligence, 56*, 58-64.
- Gagnon, D. (1986). *Interactive versus observational media: The influence of user control and cognitive styles on spatial learning* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 8616765).
- Gobet, F., Johnston, S. J., Ferrufino, G., Johnston, M., Jones, M. B., Molyneux, A., . . . Weeden, L. (2014). "No level up!": No effects of video game specialization and expertise on cognitive performance. *Frontiers in Psychology, 5*, 1337.
- Greenfield, P. M., Brannon, C., & Lohr, D. (1994). Two-dimensional representation of

- movement through three- dimensional space: The role of video game expertise. *Journal of Applied Developmental Psychology*, 15, 87-103.
- Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology*, 24, 1149-1167.
- Hartanto, A., Toh, W. X., & Yang, H. (2016). Age matters: The effect of onset age of video game play on task-switching abilities. *Attention, Perception, & Psychophysics*, 78, 1125-1136.
- Hollis, S. D. (2014). *Cognitive effects and academic consequences of video game playing* (Master's thesis). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 1566326).
- Irons, J. L., Remington, R. W., & McLean, J. P. (2011). Not so fast: Rethinking the effects of action video games on attentional capacity. *Australian Journal of Psychology*, 63, 224-231.
- Jackson, L. A., von Eye, A., Witt, E. A., Zhao, Y., & Fitzgerald, H. E. (2011). A longitudinal study of the effects of Internet use and videogame playing on academic performance and the roles of gender, race and income in these relationships. *Computers in Human Behavior*, 27, 228-239.
- Jain, P. P. (2012). *Gender, video game playing, and kindergarten children's mental rotation abilities* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3515690).
- Jones, J. R. (2007). *An exploration of multiple-object tracking across various fields of view* (Master's thesis). The University of Guelph, Guelph, Ontario, Canada.
- Keller, S. M. (1990). *Children and the Nintendo: A prediction of eye-hand coordination*,

- lateral awareness, and directionality based on hours of play* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 9117416).
- Kuhlman, J. S., & Beitel, P. A. (1991). Videogame experience: A possible explanation for difference in anticipation of coincidence. *Perceptual and Motor Skills*, 72, 483-488.
- Kuschpel, M. S., Liu, S., Schad, D. J., Heinzl, S., Heinz, A., & Rapp, M. A. (2015). Differential effects of wakeful rest, music and video game playing on working memory performance in the n-back task. *Frontiers in Psychology*, 6, 1683.
- Latham A.J., Patston L.L.M., Westermann C., & Kirk I.J., Tippett L.J. (2013). Earlier visual N1 latencies in expert video-game players: A temporal basis of enhances visuospatial performance? *PLoS ONE*, 8, e75231.
- Lee, H., Boot, W. R., Baniqued, P. L., Voss, M. W., Prakash, R. S., Basak, C., & Kramer, A. F. (2015). The relationship between intelligence and training gains is moderated by training strategy. *PLoS ONE*, 10, e0123259.
- Lee, H., Boot, W. R., Basak, C., Voss, M. W., Prakash, R. S., Neider, M., . . . Kramer, A. F. (2012). Performance gains from directed training do not transfer to untrained tasks. *Acta Psychologica*, 139, 146-158.
- Mack, D. J., & Ilg, U. J. (2014). The effects of video game play on the characteristics of saccadic eye movements. *Vision Research*, 102, 26-32.
- Mack, D. J., Wiesmann, H., & Ilg, U. J. (2016). Video game players show higher performance but no difference in speed of attention shifts. *Acta Psychologica*, 169, 11-19.
- Millard, H. A., Millard, R. P., Constable, P. D., & Freeman, L. J. (2014). Relationships among video gaming proficiency and spatial orientation, laparoscopic, and traditional surgical skills of third-year veterinary students. *Journal of the American Veterinary Medical Association*, 244, 357-362.

- Moreau, D. (2013). Differentiating two- from three-dimensional mental rotation training effects. *The Quarterly Journal of Experimental Psychology*, *66*, 1399-1413.
- Murphy, K., & Spencer, A. (2009). Playing video games does not make for better visual attention skills. *Journal of Articles in Support of the Null Hypothesis*, *6*, 1-20.
- Nelson, R. A., & Strachan, I. (2009). Action and puzzle video games prime different speed/accuracy tradeoffs. *Perception*, *38*, 1678-1687.
- Okagaki L., & Frensch, P. A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. *Journal of Applied Developmental Psychology*, *15*, 33-58.
- Pasinski, A. (2014). *Possible benefits of playing music video games* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3686131).
- Quaiser-Pohl, C., Geiser, C., & Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality and Individual Differences*, *40*, 609-619.
- Rabbitt, P., Banerji, N., & Szymanski, A. (1989). Space fortress as an IQ test? Predictions of learning and of practised performance in a complex interactive video-game. *Acta Psychologica*, *71*, 243-257.
- Richardson, A. E., Powers, M. E., & Bousquet, L. G. (2011). Video game experience predicts virtual, but not real navigation performance. *Computers in Human Behavior*, *27*, 552-560.
- Rilea, S. L. (2002). *Sex differences in spatial ability: A lateralization of function approach* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3067305).
- Rodán A., Contreras M.J., Elosúa M.R., & Gimeno P. (2016). Experimental but not sex differences of a mental rotation training program on adolescents. *Frontiers in Psychology*,

7, 1050.

- Ruiz, J. R., Ortega, F. B., Castillo, R., Martín-Matillas, M., Kwak, L., Vicente-Rodríguez, G., . . . Moreno, L. A. (2010). Physical activity, fitness, weight status, and cognitive performance in adolescents. *The Journal of Pediatrics*, *157*, 917-922.
- Santone, A. L. (2009). *Visuospatial characterization and analysis of spatial abilities of video game players* (Doctoral dissertation). Retrieved from ProQuest Dissertation and Theses database. (UMI No. 3379761).
- Searle, H. (2007). *Cognitive skill and gender in 2D and 3D game play* (Master's thesis). Retrieved from ProQuest Dissertation and Theses database. (UMI No. 1444270).
- Sekuler, R., McLaughlin, C., & Yotsumoto, Y. (2008). Age-related changes in attentional tracking of multiple moving objects. *Perception*, *37*, 867-876.
- Shute, V. J., Ventura, M., & Ke, F. (2015). The power of play: The effects of Portal 2 and Lumosity on cognitive and noncognitive skills. *Computers & Education*, *80*, 58-67.
- Sims, V. K., & Mayer, R. E. (2002). Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology*, *16*, 97-115.
- Steenbergen, L., Sellaro, R., Stock, A. K., Beste, C., & Colzato, L. S. (2015). Action video gaming and cognitive control: Playing first person shooter games is associated with improved action cascading but not inhibition. *PLoS ONE*, *10*, e0144364.
- Subrahmanyam, K., & Greenfield, P. M. (1994). Effect of video game practice on spatial skills in girls and boys. *Journal of Applied Developmental Psychology*, *15*, 13-32.
- Swing, E. L. (2012). *Plugged in: The effects of electronic media use on attention problems, cognitive control, visual attention, and aggression* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3539430).
- Unsworth, N., Redick, T. S., McMillan, B. D., Hambrick, D. Z., Kane, M. J., & Engle, R. W.

- (2015). Is playing video games related to cognitive abilities? *Psychological Science*, *26*, 759-774.
- van Schie, E. G. M., & Wiegman, O. (1997). Children and videogames: Leisure activities, aggression, social integration, and school performance. *Journal of Applied Social Psychology*, *27*, 1175-1194.
- Wilms, I. L., Petersen, A., & Vangkilde, S. (2013). Intensive video gaming improves encoding speed to visual short-term memory in young male adults. *Acta Psychologica*, *142*, 108-118.
- Wu, S., Cheng, C. K., Feng, J., D'Angelo, L., Alain, C., & Spence, I. (2012). Playing a first-person shooter video game induces neuroplastic change. *Journal of Cognitive Neuroscience*, *24*, 1286-1293.
- Wu, S., & Spence, I. (2013). Playing shooter and driving videogames improves top-down guidance in visual search. *Attention, Perception, & Psychophysics*, *75*, 673-686.

Appendix E. Studies Included in Chapter 7 (Meta-Analysis 2)

- Adams, D. M. (2013). *Exploring the effect of video game playing on static and dynamic spatial cognition skills* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3559770).
- Alves, L., & Carvalho, A. M. (2010). Videogame impact in attention test. *Psicologia em Estudo, 15*, 519-525.
- Andrews, G., and Murphy, K. (2006). "Does video-game playing improve executive function?". In M. A. Vanchevsky, (Ed.), *Frontiers in Cognitive Sciences* (pp. 145-161). USA: Nova Science Publishers.
- Appelbaum, L. G., Cain, M. S., Darling, E. F., & Mitroff, S. R. (2013). Action video game playing is associated with improved visual sensitivity, but not alterations in visual sensory memory. *Attention, Perception & Psychophysics, 75*, 1161-1167.
- Bailey, K. M. (2009). *Individual differences in video game experience: Cognitive control, affective processing, and visuospatial processing* (Master's thesis). Iowa State University, Ames, Iowa, USA.
- Bailey, K., West, R., & Anderson, C. A. (2010). A negative association between video game experience and proactive cognitive control. *Psychophysiology, 47*, 34-42.
- Bailey, K. (2012). *What would my avatar do? Video games and risky decision making* (Doctoral dissertation). Iowa State University, Ames, Iowa, USA.
- Bavelier, D., Achtman, R. L., Mani, M., & Föcker, J. (2012). Neural bases of selective attention in action video game players. *Vision Research, 61*, 132-143.
- Bejjanki, V. R., Zhang, R., Li, R., Pouget, A., Green, C. S., Lu, Z. L., & Bavelier, D. (2014). Action video game play facilitates the development of better perceptual templates. *Proceedings of the National Academy of Sciences, 111*, 16961-16966.

- Berard, A. V., Cain, M. S., Watanabe, T., & Sasaki, Y. (2015). Frequent video game players resist perceptual interference. *PLoS ONE*, *10*, e0120011.
- Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon Task. *Canadian Journal of Experimental Psychology*, *60*, 68-79.
- Blacker, K. J., & Curby, K. M. (2013). Enhanced visual short-term memory in action video game players. *Attention, Perception & Psychophysics*, *75*, 1128-1136.
- Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica*, *129*, 387-398.
- Cain, M. S., Landau, A. N., & Shimamura, A. P. (2012). Action video game experience reduces the cost of switching tasks. *Attention, Perception & Psychophysics*, *74*, 641-647.
- Cain, M. S., Prinzmetal, W., Shimamura, A. P., & Landau, A. N. (2014). Improved control of exogenous attention in action video game players. *Frontiers in Psychology*, *5*, 69.
- Castel, A., Pratt, J., & Drummond, E. (2005). The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. *Acta Psychologica*, *119*, 217-230.
- Castell, S., Larios, H., Jenson, J., & Smith, D. H. (2015). The role of video game experience in spatial learning and memory. *Journal of Gaming & Virtual Worlds*, *7*, 21-40.
- Chisholm, J. D., Hickey, C., Theeuwes, J., & Kingstone, A. (2010). Reduced attentional capture in action video game players. *Attention, Perception & Psychophysics*, *72*, 667-671.
- Chisholm, J. D., & Kingstone, A. (2012). Improved top-down control reduces oculomotor capture: The case of action video game players. *Attention, Perception & Psychophysics*, *74*, 257-262.
- Chisholm, J. D., & Kingstone, A. (2015a). Action video game players' visual search

- advantage extends to biologically relevant stimuli. *Acta Psychologica*, 159, 93-99.
- Chisholm, J. D., & Kingstone, A. (2015b). Action video games and improved attentional control: Disentangling selection- and response-based processes. *Psychonomic Bulletin & Review*, 22, 1430-1436.
- Clark, K., Fleck, M. S., & Mitroff, S. R. (2011). Enhanced change detection performance reveals improved strategy use in avid action video game players. *Acta Psychologica*, 136, 67-72.
- Clemenson, G. D., & Stark, C. E. L. (2015). Virtual environmental enrichment through video games improves hippocampal-associated memory. *The Journal of Neuroscience*, 35, 16116-16125.
- Collins, E., & Freeman, J. (2014). Video game use and cognitive performance: Does it vary with the presence of problematic video game use? *Cyberpsychology, Behavior, and Social Networking*, 17, 153-159.
- Colzato, L. S., van Leeuwen, P. J. A., van den Wildenberg, W. P. M., & Hommel, B. (2010). DOOM'd to switch: Superior cognitive flexibility in players of first person shooter games. *Frontiers in Psychology*, 1, 8.
- Colzato, L. S., van den Wildenberg, W. P. M., Zmigrod, S., & Hommel, B. (2013). Action video gaming and cognitive control: Playing first person shooter games is associated with improvement in working memory but not action inhibition. *Psychological Research*, 77, 234-239.
- Dean, D. H. (2009). Spatial visualization and the gender gap in videogame interest among young adults. *Young Consumers*, 10, 225-237.
- Dobrowolski, P., Hanusz, K., Sobczyk, B., Skorko, M., & Wiatrow, A (2015). Cognitive enhancement in video game players: The role of video game genre. *Computers in Human*

- Behavior*, 44, 59-63.
- Donohue, S. E., James, B., Eslick, A. N., & Mitroff, S. R. (2012). Cognitive pitfall! Videogame players are not immune to dual-task costs. *Attention, Perception, & Psychophysics*, 74, 803-809.
- Donohue, S. E., Woldorf, M. G., & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics*, 72, 1120-1129.
- Durlach, P. J., Kring, J. P., & Bowens, L. D. (2009). Effects of action video game experience on change detection. *Military Psychology*, 21, 24-39.
- Dye, M. W. G., & Bavelier, D. (2010). Differential development of visual attention skills in school-age children. *Vision Research*, 50, 452-459.
- Dye, M. W. G., Green, C. S., & Bavelier, D. (2009). The development of attention skills in action video game players. *Neuropsychologica*, 47, 1780-1789.
- Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science*, 18, 850-855.
- Fleck, M. S. (2009). *Effects of expectation, experience, and environment on visual search* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3350479).
- Foroughi, C. K., Serraino, C., Parasuraman, R., & Boehm-Davis, D. A. (2016). Can we create a measure of fluid intelligence using Puzzle Creator within Portal 2? *Intelligence*, 56, 58-64.
- Gaspar, J. G., Neider, M. B., Crowell, J. A., Lutz, A., Kaczmariski, H., & Kramer, A. F. (2014). Are gamers better crossers? An examination of action video game experience and dual task effects in a simulated street crossing task. *Human Factors*, 56, 443-452.

- Gobet, F., Johnston, S. J., Ferrufino, G., Johnston, M., Jones, M. B., Molyneux, A., . . . Weeden, L. (2014). "No level up!": No effects of video game specialization and expertise on cognitive performance. *Frontiers in Psychology, 5*, 1337.
- Gong, D., He, H., Ma, W., Liu, D., Huang, M., Dong, L., . . . Yao, D. (2016). Functional integration between salience and central executive networks: A role for action video game experience. *Neural Plasticity, 2016*, 9803165.
- Gonzales, C. J. (2011). *How expertise can modulate spatial attention within and across sensory modalities: The case of video game players* (Doctoral dissertation). University of Hawaii at Manoa, Honolulu, USA.
- Gozli, D. G., Bavelier, D., & Pratt, J. (2014). The effect of action video game playing on sensorimotor learning: Evidence from a movement tracking task. *Human Movement Science, 38*, 152-162.
- Green, C. S. (2008). *The effects of action video game experience on perceptual decision making* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3314679).
- Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*, 534-537.
- Green, C. S., & Bavelier, D. (2006a). Effect of action video games on the spatial distribution of visuospatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1465-1478.
- Green, C. S., & Bavelier, D. (2006b). Enumeration versus multiple object tracking: The case of action video game players. *Cognition, 101*, 217-245.
- Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science, 18*, 88-94.

- Green, C. S., Sugarman, M. A., Medford, K., Klobusicky, E., & Bavelier, D. (2012). The effect of action video game experience on task-switching. *Computers in Human Behavior*, 28, 984-994.
- Greenfield, P. M., DeWinstanley, P., Kilpatrick, H., & Kaye, D. (1994). Action video games and informal education: Effects on strategies for dividing visual attention. *Journal of Applied Developmental Psychology*, 15, 105-123.
- Greenle, E. T. (2013). *Expert enhancement of spatial perception in the face of uncertainty* (Master's thesis). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 1545811).
- Griffith, J. L., Voloschin, P., Gibb, G. D., & Bailey, J. R. (1983). Differences in eye-hand motor coordination of video-game users and non-users. *Perceptual and Motor Skills*, 57, 155-158.
- Gutierrez, M. (2013). *Strengthening cognitive development in minority populations: A study of the beneficial effects of bilingualism* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3609486).
- Hartanto, A., Toh, W. X., & Yang, H. (2016). Age matters: The effect of onset age of video game play on task-switching abilities. *Attention, Perception, & Psychophysics*, 78, 1125-1136.
- Heimler, B., Pavani, F., Donk, M., & van Zoest, W. (2014). Stimulus- and goal-driven control of eye movements: Action videogame players are faster but not better. *Attention, Perception, & Psychophysics*, 76, 2398-2412.
- Hollis, S. D. (2014). *Cognitive effects and academic consequences of video game playing* (Master's thesis). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 1566326).

- Hubert-Wallander, B., Green, C. S., Sugarman, M., & Bavelier, D. (2011). Changes in search rate but not in the dynamics of exogenous attention in action videogame players. *Attention, Perception, & Psychophysics*, *73*, 2399-2412.
- Hutchinson, C. V., & Stocks, R. (2013). Selectively enhanced motion perception in core video gamers. *Perception*, *42*, 675-677.
- Irons, J. L., Remington, R. W., & McLean, J. P. (2011). Not so fast: Rethinking the effects of action video games on attentional capacity. *Australian Journal of Psychology*, *63*, 224-231.
- Jain, P. P. (2012). *Gender, video game playing, and kindergarten children's mental rotation abilities* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3515690).
- Kandula, M., Hofman, D., & Dijkerman, H. C. (2016). Location estimation of approaching objects is modulated by the observer's inherent and momentary action capabilities. *Experimental Brain Research*, *234*, 2315-2322.
- Karle, J. W., Watter, S., & Shedden, J. M. (2010). Task switching in video game players: Benefits of selective attention but not resistance to proactive interference. *Acta Psychologica*, *134*, 70-78.
- Kennedy, A. M., Boyle, E. M., Traynor, O., Walsh, T., & Hill, A. D. (2011). Video gaming enhances psychomotor skills but not visuospatial and perceptual abilities in surgical trainees. *Journal of Surgical Education*, *68*, 414-420.
- Kim, Y. W., Kang, D. W., Kim, D., Kim, H. J., Sasaki, Y., & Watanabe, T. (2015). Real-time strategy video game experience and visual perceptual learning. *The Journal of Neuroscience*, *35*, 10485-10492.
- Krishnan, L., Kang, A., Sperling, G., & Srinivasan, R. (2013). Neural strategies for selective

- attention distinguish fast-action video game players. *Brain Topography*, 26, 83-97.
- Kuhlman, J. S., & Beitel, P. A. (1991). Videogame experience: A possible explanation for difference in anticipation of coincidence. *Perceptual and Motor Skills*, 72, 483-488.
- Latham A.J., Patston L.L.M., Westermann C., Kirk I.J., & Tippett L.J. (2013). Earlier visual N1 latencies in expert video-game players: A temporal basis of enhances visuospatial performance? *PLoS ONE*, 8, e75231.
- Li, L., Chen, R., & Chen, J. (2016). Playing action video games improves visuomotor control. *Psychological Science*, 27, 1092-1108.
- Li, X., Cheng, X., Li, J., Pan, Y., Hu, Y., & Ku, Y. (2015). Examination of mechanisms underlying enhanced memory performance in action video game players: A pilot study. *Frontiers in Psychology*, 6, 843.
- Li, R., Polat, U., Makous, W., & Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nature Neuroscience*, 12, 549-551.
- Li, R., Polat, U., Scalzo, F., & Bavelier, D. (2010). Reducing backward masking through action game training. *Journal of Vision*, 10, 1-13.
- Mack, D. J., & Ilg, U. J. (2014). The effects of video game play on the characteristics of saccadic eye movements. *Vision Research*, 102, 26-32.
- Mack, D. J., Wiesmann, H., & Ilg, U. J. (2016). Video game players show higher performance but no difference in speed of attention shifts. *Acta Psychologica*, 169, 11-19.
- Mackenzie, B., & Cumming, S. (1986). How fragile is the relationship between inspection time and intelligence: The effects of apparent-motion cues and previous experience. *Personality and Individual Differences*, 7, 721-729.
- McDermott, A. F., Bavelier, D., & Green, C. S. (2014). Memory abilities in action video game players. *Computers in Human Behavior*, 34, 69-78.

- Mishra, J. Zinni, M., Bavelier, D., & Hillyard, S. A. (2011). Neural basis of superior performance of action videogame players in an attention-demanding task. *The Journal of Neuroscience*, *31*, 992-998.
- Murphy, K., & Spencer, A. (2009). Playing video games does not make for better visual attention skills. *Journal of Articles in Support of the Null Hypothesis*, *6*, 1-20.
- Novak, E., & Tassell, J. (2015). Using video game play to improve education-majors' mathematical performance: An experimental study. *Computers in Human Behavior*, *53*, 124-130.
- Pasinski, A. (2014). *Possible benefits of playing music video games* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 3686131).
- Pavan, A., Boyce, M., & Ghin, F. (2016). Action video games improve direction discrimination of parafoveal translational global motion but not reaction times. *Perception*, *45*, 1193-1202.
- Pohl, C., Kunde, W., Ganz, T., Conzelmann, A., Pauli, P., & Kiesel, A. (2014). Gaming to see: Action video gaming is associated with enhanced processing of masked stimuli. *Frontiers in Psychology*, *5*, 70.
- Quaiser-Pohl, C., Geiser, C., & Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality and Individual Differences*, *40*, 609-619.
- Romano Bergstrom, J. C., Howard, J. H., & Howard, D. V. (2012). Enhanced implicit sequence learning in college-age video game players and musicians. *Applied Cognitive Psychology*, *26*, 91-96.
- Schubert, T., Finke, K., Redel, P., Kluckow, S., Müller, H., & Strobach, T. (2015). Video game experience and its influence on visual attention parameters: An investigation using

- the framework of the Theory of Visual Attention (TVA). *Acta Psychologica*, *157*, 200-214.
- Sekuler, R., McLaughlin, C., & Yotsumoto, Y. (2008). Age-related changes in attentional tracking of multiple moving objects. *Perception*, *37*, 867-876.
- Sims, V. K., & Mayer, R. E. (2002). Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology*, *16*, 97-115.
- Smith, R. P. (2007). *Contrast effects produced by computer gaming: A study of learning and attention* (Master's thesis). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 1455789).
- Steenbergen, L., Sellaro, R., Stock, A. K., Beste, C., & Colzato, L. S. (2015). Action video gaming and cognitive control: Playing first person shooter games is associated with improved action cascading but not inhibition. *PLoS ONE*, *10*, e0144364.
- Strobach, T., Frensch, P. A., & Schubert, T. (2012). Video game practice optimizes executive control skills in dual-task and task switching situations. *Acta Psychologica*, *140*, 13-24.
- Sungur, H., & Boduroglu, A. (2012). Action video game players form more detailed representation of objects. *Acta Psychologica*, *139*, 327-334.
- Tanaka, S., Ikeda, H., Kasahara, K., Kato, R., Tsubomi, H., Sugawara, S. K., . . . Watanabe, K. (2013). Larger right posterior parietal volume in action video game experts: A behavioral and voxel-based morphometry (VBM) study. *PLoS ONE*, *8*, e66998.
- Trick, L. M., Jaspers-Fayer, F., & Sethi, N. (2005). Multiple-object tracking in children: The "Catch the Spies" task. *Cognitive Development*, *20*, 373-387.
- Unsworth, N., Redick, T. S., McMillan, B. D., Hambrick, D. Z., Kane, M. J., & Engle, R. W. (2015). Is playing video games related to cognitive abilities? *Psychological Science*, *26*, 759-774.
- Vallett, D. B., Lamb, R. L., & Annetta, L. A. (2013). The gorilla in the room: The impacts of

- video-game play on visual attention. *Computers in Human Behavior*, 29, 2183-2187.
- West, G. L., Stevens, S. A., Pun, C., & Pratt, J. (2008). Visuospatial experience modulates attentional capture: Evidence from action video game players. *Journal of Vision*, 8, 1-9.
- Wilms, I. L., Petersen, A., & Vangkilde, S. (2013). Intensive video gaming improves encoding speed to visual short-term memory in young male adults. *Acta Psychologica*, 142, 108-118.
- Wu, S., & Spence, I. (2013). Playing shooter and driving videogames improves top-down guidance in visual search. *Attention, Perception, & Psychophysics*, 75, 673-686.
- Zaparyniuk, N. E. (2006). *The exploration of video games as a tool for problem solving and cognitive skills development* (Master's thesis). University of Alberta, Edmonton, Alberta, Canada.
- Zetu, A. G. (2007). Neurophysiological differences in visual selective attention in action video-game players (Master's thesis). University of Toronto, Toronto, Canada.
- Zhang, Y., Du, G., Yang, Y., Qin, W., Li, X., & Zhang, Q. (2015). Higher integrity of the motor and visual pathways in long-term video game players. *Frontiers in Human Neuroscience*, 9, 98.

Appendix F. Studies Included in Chapter 7 (Meta-Analysis 3)

- Adams, D. M. (2013). *Exploring the effect of video game playing on static and dynamic spatial cognition skills* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3559770).
- Adams, D. M., Pilegard, C., & Mayer, R. E. (2016). Evaluating the cognitive consequences of playing Portal for a short duration. *Journal of Educational Computing Research, 54*, 173-195.
- Barlett, C. P., Vowels, C. L., Shanteau, J., Crow, J., & Miller, T. (2009). The effect of violent and non-violent computer games on cognitive performance. *Computers in Human Behavior, 25*, 96-102.
- Basak, C., Boot, W. R., Voss, M. W., & Kramer, A. F. (2008). Can training in a real-time strategy videogame attenuate cognitive decline in older adults? *Psychology and Aging, 23*, 765-777.
- Bejjanki, V. R., Zhang, R., Li, R., Pouget, A., Green, C. S., Lu, Z. L., & Bavelier, D. (2014). Action video game play facilitates the development of better perceptual templates. *Proceedings of the National Academy of Sciences, 111*, 16961-16966.
- Belchior, P., Marsiske, M., Sisco, S. M., Yam, A., Bavelier, D., Ball, K., & Mann, W. C. (2013). Video game training to improve selective visual attention in older adults. *Computers in Human Behavior, 29*, 1318-1324.
- Blacker, K. J., Curby, K. M., Klobusicky, E., & Chein, J. M. (2014). Effects of action video game training on visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 1992-2004.
- Boot, W. R. (2007). *The effects of video game playing on perceptual and cognitive abilities* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI

No. 3269846).

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.

Buelow, M. T., Okdie, B. M., & Cooper, A. B. (2015). The influence of video games on executive functions in college students. *Computers in Human Behavior, 45*, 228-234.

Cherney, I. D. (2008). Mom, let me play more computer games: They improve my mental rotation skills. *Sex Roles, 59*, 776-786.

Cherney, I. D., Bersted, K., & Smetter, J. (2014). Training spatial skills in men and women. *Perceptual & Motor Skills: Learning & Memory, 119*, 82-99.

Clark, J. E., Lanphear, A. K., & Riddick, C. C. (1987). The effects of videogame playing on the response selection processing of elderly adults. *Journal of Gerontology, 42*, 82-85.

Clemenson, G. D., & Stark, C. E. L. (2015). Virtual environmental enrichment through video games improves hippocampal-associated memory. *The Journal of Neuroscience, 35*, 16116-16125.

Cohen, J. E., Green, C. S., & Bavelier, D. (2008). Training visual attention with video games: Not all games are created equal. In H. O'Neil & R. Perez (Eds.), *Computer games and adult learning* (pp. 205-227). Amsterdam: Elsevier.

Colom, R., Quiroga, M. Á., Solana, A. B., Burgaleta, M., Román, F. J., Privado, J., . . .

Karama, S. (2012). Structural changes after videogame practice related to a brain network associated with intelligence. *Intelligence, 40*, 479-489.

De Lisi, R., & Cammarano, D. M. (1996). Computer experience and gender differences in undergraduate mental rotation mental performance. *Computers in Human Behavior, 12*, 351-361.

- De Lisi, R., & Wolford, J. L. (2002). Improving children's mental rotation accuracy with computer game playing. *The Journal of Genetic Psychology, 163*, 272-282.
- Dorval, M., & Pépin, M. (1986). Effect of playing a video game on a measure of spatial visualization. *Perceptual and Motor Skills, 62*, 159-162.
- Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*, 850-855.
- Fortman, J. (2012). *Computer-based cognitive training for age-related cognitive decline and mild cognitive impairment* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3534589).
- Fulgham, L. M. (2003). *Effects of videogame play in boys: An examination of adrenocortical activity and behavior* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3079297).
- Gagnon, D. (1986). *Interactive versus observational media: The influence of user control and cognitive styles on spatial learning* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 8616765).
- Glass, B. D., Maddow, W. T., & Love, B. C. (2013). Real-time strategy game training: Emergence of a cognitive flexibility trait. *PLoS ONE, 8*, e70350.
- Goldstein, J., Cajko, L., Oosterbroek, M., Michielsen, M., van Houten, O., & Salverda, F. (1997). Video games and the elderly. *Social Behavior and Personality, 25*, 345-352.
- Gonzales, C. J. (2011). *How expertise can modulate spatial attention within and across sensory modalities: The case of video game players* (Doctoral dissertation). University of Hawaii at Manoa, Honolulu, USA.
- Green, C. S. (2008). *The effects of action video game experience on perceptual decision making* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses

- database. (UMI No. 3314679).
- Green, C. S., & Bavelier, D. (2006a). Effect of action video games on the spatial distribution of visuospatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1465-1478.
- Green, C. S., & Bavelier, D. (2006b). Enumeration versus multiple object tracking: The case of action video game players. *Cognition*, 101, 217-245.
- Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science*, 18, 88-94.
- Green, C. S., Sugarman, M. A., Medford, K., Klobusicky, E., & Bavelier, D. (2012). The effect of action video game experience on task-switching. *Computers in Human Behavior*, 28, 984-994.
- Hutchinson, C.V., Barrett, D. J. K., Nitka, A., & Raynes, K. (2016). Action video game training reduces the Simon Effect. *Psychonomic Bulletin & Review*, 23, 587-592.
- Kearney, P. (2005). *Cognitive callisthenics: Do fps computer games enhance the player's cognitive abilities?* Paper presented at the DiGRA 2005 Changing Views: Worlds in Play International Conference, Vancouver, Canada.
- Lee, H., Boot, W. R., Basak, C., Voss, M. W., Prakash, R. S., Neider, M., . . . Kramer, A. F. (2012). Performance gains from directed training do not transfer to untrained tasks. *Acta Psychologica*, 139, 146-158.
- Li, R. (2010). *The effects of action video game playing on low level vision* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3410597).
- Li, L., Chen, R., & Chen, J. (2016). Playing action video games improves visuomotor control. *Psychological Science*, 27, 1092-1108.

- Lorant-Royer, S., Munch, C., Mesclé, H., & Lieury, A. (2010). Kawashima vs “Super Mario”! Should a game be serious in order to stimulate cognitive aptitudes? *Revue Européenne de Psychologie Appliquée*, *60*, 221-232.
- Lucia, L. C. (2013). *An electrophysiological study of the effects of visuospatial training on mental rotation, object identification, and enumeration performance* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3556627).
- Marra, D. (2016). *Video game interventions to improve cognition in older adults* (Master’s thesis). Marquette University, Milwaukee, Wisconsin, USA.
- Minear, M., Brasher, F., Guerrero, C. B., Brasher, M., Moore, A., & Sukeena, J. (2016). A simultaneous examination of two forms of working memory training: Evidence for near transfer only. *Memory & Cognition*, *44*, 1014-1037.
- Nelson, R. A., & Strachan, I. (2009). Action and puzzle video games prime different speed/accuracy tradeoffs. *Perception*, *38*, 1678-1687.
- Novak, E., & Tassell, J. (2015). Using video game play to improve education-majors’ mathematical performance: An experimental study. *Computers in Human Behavior*, *53*, 124-130.
- Oei, A. C., & Patterson, M. D. (2014). Playing a puzzle video game with changing requirements improves executive functions. *Computers in Human Behavior*, *37*, 216-228.
- Okagaki, L., & Frensch, P. A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. *Journal of Applied Developmental Psychology*, *15*, 33-58.
- Orosy-Fildes, C., & Allan, R. W. (1989). Psychology of computer use: XII. Videogame play: Human reaction time to visual stimuli. *Perceptual and Motor Skills*, *69*, 243-247.

- Patton, R. D. (2014). *The effects of videogames on clinical measures of attention, processing speed, and working memory* (Specialist thesis). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 1570531).
- Pepin, M., & Dorval, M. (1986, April). *Effect of playing a video game on adults' and adolescents' spatial visualization*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Sanchez, C. A. (2012). Enhancing visuospatial performance through video game training to increase learning in visuospatial science domains. *Psychonomic Bulletin & Review*, *19*, 58-65.
- Schubert, T., Finke, K., Redel, P., Kluckow, S., Müller, H., & Strobach, T. (2015). Video game experience and its influence on visual attention parameters: An investigation using the framework of the Theory of Visual Attention (TVA). *Acta Psychologica*, *157*, 200-214.
- Shute, V. J., Ventura, M., & Ke, F. (2015). The power of play: The effects of Portal 2 and Lumosity on cognitive and noncognitive skills. *Computers & Education*, *80*, 58-67.
- Sims, V. K., & Mayer, R. E. (2002). Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology*, *16*, 97-115.
- 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.
- Strobach, T., Frensch, P. A., & Schubert, T. (2012). Video game practice optimizes executive control skills in dual-task and task switching situations. *Acta Psychologica*, *140*, 13-24.
- Subrahmanyam, K., & Greenfield, P. M. (1994). Effect of video game practice on spatial skills in girls and boys. *Journal of Applied Developmental Psychology*, *15*, 13-32.
- Swing, E. L. (2012). *Plugged in: The effects of electronic media use on attention problems*,

- cognitive control, visual attention, and aggression* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3539430).
- Terlecki, M. S., Newcombe, N. S., & Little, M. (2008). Durable and generalized effects of spatial experience on mental rotation: Gender differences in growth patterns. *Applied Cognitive Psychology, 22*, 996-1013.
- Valadez, J. J., & Ferguson, C. J. (2012). Just a game after all: Violent video game exposure and time spent playing effects on hostile feelings, depression, and visuospatial cognition. *Computers in Human Behavior, 28*, 608-616.
- van Ravenzwaaij, D., Boekel, W., Forstmann, B. U., Ratcliff, R., Wagenmakers, E. J. (2014). Action video games do not improve the speed of information processing in simple perceptual tasks. *Journal of Experimental Psychology: General, 143*, 1794-1805.
- Vaupel, C. (2002). *The effects of video game playing on academic task performance and brain wave activity* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3054150).
- Whitlock, L. A., McLaughlin, A. C., & Allaire, J. C. (2012). Individual differences in response to cognitive training: Using a multi-modal, attentionally demanding game-based intervention for older adults. *Computers in Human Behavior, 28*, 1091-1096.
- Wu, S., Cheng, C. K., Feng, J., D'Angelo, L., Alain, C., & Spence, I. (2012). Playing a first-person shooter video game induces neuroplastic change. *Journal of Cognitive Neuroscience, 24*, 1286-1293.
- Wu, S., & Spence, I. (2013). Playing shooter and driving videogames improves top-down guidance in visual search. *Attention, Perception, & Psychophysics, 75*, 673-686.
- Yam, A. (2014). *The effects of spatially-demanding videogame training on mental rotation performance and confidence in older adults* (Doctoral dissertation). Retrieved from

ProQuest Dissertations and Theses database. (UMI No. 3691386).

Appendix G. Studies Included in the Meta-Analysis in Chapter 8

- Anderson-Hanley, C., Arciero, P. J., Brickman, A. M., Nimon, J. P., Okuma, N., Westen, S. C., ... Zimmerman, E. A. (2012). Exergaming and older adult cognition: A cluster randomized clinical trial. *American Journal of Preventive Medicine, 42*, 109-119.
- Barcelos, N., Shah, N., Cohen, K., Hogan, M. J., Mulkerrin, E., Arciero, P. J., ... Anderson-Hanley, C. (2015). Aerobic and Cognitive Exercise (ACE) pilot study for older adults: executive function improves with cognitive challenge while exergaming. *Journal of the International Neuropsychological Society, 21*, 768-779.
- Chan, C. L., Ngai, E. K., Leung, P. K. H., & Wong, S. (2010). Effect of the adapted virtual reality cognitive training program among Chinese older adults with chronic schizophrenia: A pilot study. *International Journal Geriatric Psychiatry, 25*, 643-649.
- Eggenberger, P. S., Schumacher, V., Angst, M., Theill, N., & de Bruin, E. D. (2015). Does multicomponent physical exercise with simultaneous cognitive training boost cognitive performance in older adults? A 6-month randomized controlled trial with a 1-year follow-up. *Clinical Intervention in Aging, 10*, 1335-1349.
- Eggenberger, P. S., Wolf, M., Schumann, M., & de Bruin, E. D. (2016). Exergame and balance training modulate prefrontal brain activity during walking and enhance executive function in older adults. *Frontiers in Aging Neuroscience, 8*, 66.
- Hughes, T. F., Flatt, J. D., Fu, B., Butters, M. A., Chang, C. C. H., & Ganguli, M. (2014). Interactive video gaming compared with health education in older adults with mild cognitive impairment: A feasibility study. *International Journal Geriatric Psychiatry, 29*, 890-898.
- Kimhy, D., Vakhrusheva, J., Bartels, M. N., Armstrong, H. F., Ballon, J. S., Khan, S., ... Sloan, R. P. (2015). The impact of aerobic exercise on brain-derived neurotrophic factor

- and neurocognition in individuals with schizophrenia: A single-blind, randomized clinical trial. *Schizophrenia Bulletin*, 41, 859-868.
- Maillot, P., Perrot, A., & Hartley, A. (2012). Effects of interactive physical-activity video-game training on physical and cognitive function in older adults. *Psychology and Aging*, 27, 589-600.
- Mirelman, A., Rochester, L., Maidan, I., Del Din, S., Alcock, L., Nieuwhof, F., ... Hausdorff, J. M. (2016). Addition of a non-immersive virtual reality component to treadmill training to reduce fall risk in older adults (V-TIME): A randomised controlled trial. *Lancet*, 388, 1170-1182.
- Park, J., & Yim, J. E. (2016). A new approach to improve cognition, muscle strength, and postural balance in community-dwelling elderly with a 3-D virtual reality Kayak program. *The Tohoku Journal of Experimental Medicine*, 238, 1-8.
- Pompeu, J. E, Mendes, F., da Silva, K. G, Lobo, A. M., Oliveira, T., Zomignani, P. A., & Piemonte, M. E. P. (2012). Effect of Nintendo Wii™-based motor and cognitive training on activities of daily living in patients with Parkinson's disease: A randomised clinical trial. *Physiotherapy*, 98, 196-204.
- Schättin, A., Arner, R., Gennaro, F., & de Bruin, E. D. (2016). Adaptations of prefrontal brain activity, executive functions, and gait in healthy elderly following exergame and balance training: A randomized-controlled study. *Frontiers in Aging Neuroscience*, 8, 278.
- Schoene, D., Lord, S. R., Delbaere, K., Severino, C., Davies, T. A., & Smith, S. T. (2013). A randomized controlled pilot study of home-based step training in older people using videogame technology. *PLoS One*, 8, e57734.
- Şimşek, T. T., & Çekok, K. (2016). The effects of Nintendo Wii™-based balance and upper extremity training on activities of daily living and quality of life in patients with subacute

stroke: A randomized controlled study. *International Journal of Neuroscience*, 126, 1061-1070.

Staiano, A. E., Abraham, A. A., & Calvert, S. L. (2012). Competitive versus cooperative exergame play for African American adolescents' executive function skills: Short-term effects in a long-term training intervention. *Developmental Psychology*, 48, 337-342.

Zimmermann, R., Gschwandtner, U., Benz, N., Hatz, F., Schindler, C., Taub, E., & Fuhr, P. (2014). Cognitive training in Parkinson disease: Cognition-specific vs nonspecific computer training. *Neurology*, 82, 1219-1226.