Tilo Strobach
Julia Karbach
Editors

Cognitive Training

An Overview of Features and Applications

Springer

Cognitive Training

Tilo Strobach • Julia Karbach
Editors

Cognitive Training

An Overview of Features and Applications

Springer

Editors
Tilo Strobach
Department of Psychology
Medical School Hamburg
Hamburg, Germany

Julia Karbach
Department of Psychology
Goethe-University Frankfurt
Frankfurt/Main, Germany

ISBN 978-3-319-42660-0 ISBN 978-3-319-42662-4 (eBook)
DOI 10.1007/978-3-319-42662-4

Library of Congress Control Number: 2016953131

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Contents

About the Editors

Tilo Strobach finished his doctoral degree at Humboldt University of Berlin. After that, he held postdoc positions at this university and at Ludwig Maximilians University Munich. Currently, Tilo Strobach is full professor for cognitive psychology at the Medical School Hamburg. His research focuses on the analysis of cognitive plasticity as a result of training (e.g., video-game, dual-task, working memory, and task-switching training) and cognitive aging. Furthermore, he aims at specifying the cognitive processing architecture in situations that demand executive functions as well as the perception of complex objects.

Julia Karbach is a developmental cognitive psychologist holding a Ph.D. in Psychology from Saarland University. She currently serves as full professor at Goethe-University Frankfurt. Her research interests include neurocognitive development and plasticity across the life span, psychocardiology, and the prediction of academic achievement from primary to tertiary education. Her work has been published in numerous leading peer-reviewed international journals.

Contributors

Tracy Packiam Alloway Department of Psychology, University of North Florida, Jacksonville, FL, USA

Claudia C. von Bastian Cognition and Cognitive Neuroscience Research Centre, Department of Psychology, Bournemouth University, Bournemouth, England

Daphne Bavelier Department of Psychology, University of Geneva, Geneva, Switzerland

Sylvie Belleville Research Center, Institut Universitaire de Gériatrie de Montréal, Université de Montréal, Montréal, QC, Canada

Louis Bherer PERFORM Centre and Psychology Department, Concordia University, Montreal, QC, Canada

Benjamin Boller Research Center, Institut Universitaire de Gériatrie de Montréal, Université de Montréal, Montréal, QC, Canada

Martin Buschkuehl MIND Research Institute, Irvine, CA, USA

Lorenza S. Colzato Cognitive Psychology Unit, Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands

Lina M. Cómbita Department of Experimental Psychology, Center for Research on Mind, Brain, and Behavior (CIMCYC), University of Granada, Granada, Spain

Michael Falkenstein Leibniz Research Centre for Working Environment and Human Factors at TU Dortmund (IfADo), Dortmund, Germany

Institute for Ageing Learning and Occupation (ALA), Bochum, Germany

Andrea N. Frankenstein Department of Psychology, University of North Florida, Jacksonville, FL, USA

Patrick D. Gajewski Leibniz Research Centre for Working Environment and Human Factors (IfADo), Dortmund, Germany

Thomas Gorman Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA

C. Shawn Green Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA

Sabrina Guye University Research Priority Program (URPP) "Dynamics of Healthy Aging", University of Zurich, Zurich, Switzerland

Bernhard Hommel Cognitive Psychology Unit, Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands

Susanne Jaeggi School of Education, University of California, Irvine, CA, USA

Masha R. Jones School of Education, University of California, Irvine, CA, USA

Julia Karbach Department of Psychology, Goethe-University Frankfurt, Frankfurt/Main, Germany

Benjamin Katz Department of Psychology, University of Michigan, Ann Arbor, MI, USA

Tanja Könen Department of Psychology, Goethe-University Frankfurt, Frankfurt, Germany

Jutta Kray Department of Psychology, Saarland University, Saarbrücken, Germany

Mike Martin University Research Priority Program (URPP) "Dynamics of Healthy Aging", University of Zurich, Zurich, Switzerland

Division of Gerontopsychology, Department of Psychology, University of Zurich, Zurich, Switzerland

Mark A. McDaniel Department of Psychological and Brain Sciences, Washington University in St. Louis, St. Louis, MO, USA

Susan Mérillat University Research Priority Program (URPP) "Dynamics of Healthy Aging," University of Zurich, Zurich, Switzerland

Kristell Pothier PERFORM Centre and Psychology Department, Concordia University, Montreal, QC, Canada

Joan P. Pozuelos Department of Experimental Psychology, Center for Research on Mind, Brain, and Behavior (CIMCYC), University of Granada, Granada, Spain

Tyler Robinson Department of Psychology, University of North Florida, Jacksonville, FL, USA

Christina Röcke University Research Priority Program (URPP) "Dynamics of Healthy Aging", University of Zurich, Zurich, Switzerland

M. Rosario Rueda Department of Experimental Psychology, Center for Research on Mind, Brain, and Behavior (CIMCYC), University of Granada, Granada, Spain

E. Glenn Schellenberg Department of Psychology, University of Toronto Mississauga, Mississauga, ON, Canada

Florian Schmiedek German Institute for International Educational Research (DIPF), Frankfurt am Main, Germany

Torsten Schubert Department of Psychology, Martin Luther University Halle, Halle, Germany

Priti Shah Department of Psychology, University of Michigan, Ann Arbor, MI, USA

Yee Lee Shing Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany

Division of Psychology, University of Stirling, Scotland, UK

Tilo Strobach Department of Psychology, Medical School Hamburg, Hamburg, Germany

Swathi Swaminathan Department of Psychology, University of Toronto Mississauga, Mississauga, ON, Canada

Niels A. Taatgen University of Groningen, Groningen, The Netherlands

Joan Toglia School of Health and Natural Sciences, Mercy College, Dobbs Ferry, New York, NY, USA

Sharda Umanath Department of Psychology, Claremont McKenna College, Claremont, CA, USA

Laura Prieto del Val Department of Physiology, Anatomy, and Cellular Biology, Pablo de Olavide University, Sevilla, Spain

Paul Verhaeghen School of Psychology, Georgia Institute of Technology, Atlanta, GA, USA

Elisabeth Wenger Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany

Introduction

Tilo Strobach and Julia Karbach

Throughout the entire life span, individuals are required to adapt to the demands of changing contexts and dynamic social environments. The potential modifiability of a person's cognitive and neural system resulting from these adaptations has been referred to as cognitive and neural plasticity. One way to understand this plasticity is to apply training interventions and to measure the scope of their effects in order to identify the mechanisms underlying plastic changes in mind and brain (see Hertzog et al. 2008; Karbach and Schubert 2013; Lustig et al. 2009; Schubert et al. 2014, for reviews).

Over the last decades, the literature on cognitive training interventions has been growing rapidly, demonstrating that cognitive training is a timely issue of high academic as well as societal relevance. For instance, a literature search for "cognitive training" in the abstracting and indexing databases PsychINFO and PubMed by February 19, 2016, demonstrated a total of about 1.407 peer-reviewed contributions since 1966/1.217 peer-reviewed contributions since 1973, respectively. Both databases demonstrate that more contributions on cognitive training were published between 2010 and 2016 than in the more than 40 years before that (1966–2010). These impressive numbers raise the question why research on cognitive training became so popular in the second decade of this century?

Several factors may have contributed to this development. First—and this certainly influenced many psychological disciplines, including cognitive and experimental psychology—recent decades were characterized by tremendous technical

T. Strobach (✉)
Department of Psychology, Medical School Hamburg,
Hamburg, Germany
e-mail: tilo.strobach@medicalschool-hamburg.de

J. Karbach
Department of Psychology, Goethe-University Frankfurt, Frankfurt/Main, Germany
e-mail: karbach@psych.uni-frankfurt.de

T. Strobach, J. Karbach (eds.), *Cognitive Training*,
DOI 10.1007/978-3-319-42662-4_1

advances. These technical advances also had a large impact on cognitive training research. For instance, they led to computerized experimental setups allowing to precisely assess changes in both behavior and neural processing; such precise assessments may be particularly relevant for demonstrating the sometimes rather small effects sizes in cognitive training studies. Further, technical advances also allow the efficient handling and sophisticated analyses of large longitudinal data sets that are very common in studies with extensive training protocols including numerous experimental sessions and groups of participants. With state-of-the-art software, data cannot only be analyzed at the group level but training-induced changes can also be modeled at the individual and latent level.

Second, interest in cognitive training has been spurred tremendously by findings showing that cognitive and neural plasticity are present up to very old age. Earlier accounts assumed that the brain is capable of significant plastic changes only early in life and becomes impervious to change afterwards (e.g., Wiesel and Hubel 1965), suggesting that basic processing capacities cannot be improved by training after early adulthood. However, recent work has clearly established that plasticity is not only present up to very old age (Bavelier et al. 2010; Green et al. 2014; Hertzog et al. 2008) but that the mind and brain of older adults can be as plastic as in young adults (Karbach and Verhaeghen 2014, for a meta-analysis).

Another widely discussed issue in the field of cognitive training research is the "curse of specificity," that is, the transferability of training-induced performance gains to untrained tasks and abilities (Green and Bavelier 2008; Strobach et al. 2014). Early cognitive training studies showed that while individuals improved their performance on a task given appropriate training, little to no benefits of this training were seen on new tasks (even if they were seemingly similar to the trained task). Such task-specific learning has been shown in nearly all fields of psychology from motor control, to problem-solving, reasoning, general cognition, and education (Ball and Sekuler 1982; Barnett and Ceci 2002; Detterman and Sternberg 1993; Fahle 2004; Schmidt and Bjork 1992). Nevertheless, recent work suggests that training can indeed produce broad, generalizable effects. In fact, examples of general transfer effects are frequently reported in the literature, especially after cognitive training interventions focusing on basic processing capacities, such as working memory or executive functions (see Au et al. 2015; Hindin and Zelinski 2012; Karbach and Verhaeghen 2014; Karr et al. 2014; Schwaighofer et al. 2015, for meta-analyses). Moreover, playing video games of the "action video game" genre has been shown to improve a variety of cognitive skills (e.g., Li et al. 2009; Strobach et al. 2012; see Toril et al. 2014, for a meta-analysis). Aside from these cognitive trainings, there is also ample evidence for positive effects of musical training (Schellenberg 2004) and particularly physical training (see Bherer et al. 2013, for a review and Colcombe and Kramer 2003, for a meta-analysis) on cognitive abilities across a wide range of ages.

At this point in cognitive training research (50 years after the first publications in the domain of "cognitive training" according to PsycINFO), we aimed at summing up the current state of findings of this first era of cognitive training research. From our perspective, this era is generally characterized as a rather heterogeneous phase in which (1) many studies on cognitive research were published that included a variety of designs, methods, and training protocols which unsurprisingly yielded very mixed findings and (2) studies were often less theory driven and theoretical models describing the mechanisms underlying training and effects are mostly missing. Therefore, we aimed at integrating the state-of-the-art of different domains in the field of cognitive training research accompanied by theoretical models describing the mechanisms underlying training and transfer effects.

The first section of this book covers basic concepts, theory, and methodological issues from a very general perspective (i.e., relevant for different populations, age groups, and cognitive domains). Hence, Taatgen (this volume) presents and elaborates general theoretical models of training and transfer effects. Researchers who investigate these effects can draw on a well-established methodology for the evaluation of psychological interventions. Doing so, they face the equally well-established long list of critical issues, reducing the validity of findings in studies on cognitive training. Therefore, Schmiedek (this volume) discusses the most common and relevant issues as well as possible methodological solutions.

Cognitive training is relevant throughout the entire life span. Thus, the second section of this book elaborates on the cognitive and neural plasticity in different age groups from a developmental perspective. Since effective cognitive skills are key to learning, socialization, and success to a wide range of real-world outcomes, Rueda et al. (this volume) present the great body of literature on the extent to which cognitive skills can be enhanced through training interventions during childhood and adolescence. Furthermore, probably the most prominent way of applying cognitive training is to use it as a tool against age-related decline in cognitive abilities. Guye et al. (this volume) illustrate promising avenues in this domain.

After starting with general perspectives on theory, methodology, and age groups of cognitive training, the third section provides details regarding specific cognitive domains targeted during training. Several prominent types of domain-specific training focused on memory training. Therefore, training and transfer effects are reviewed in the domain of working memory (Könen et al. this volume), episodic memory (Wenger and Shing this volume), and prospective memory (Umanath et al. this volume). The other training domain targeting higher-cognitive processes is executive functions (Karbach and Kray this volume).

Similar to the third section, the fourth section is structured by the type of training. However, in contrast to the theoretically well-defined training domains presented in section three, the chapters of this section are structured by more superficial characteristics. While trainings discussed in these chapters may look very similar, they often tap different cognitive domains (multi-domain training). For instance, video game training—more specifically "action video games"—is characterized by complex visual displays, fast-paced speed, as well as motivational elements. Therefore, Green et al. (this volume) discuss the effects of playing these games on perception and attentional control, while Strobach and Schubert (this volume) rather focus on potential influences of action video game playing on executive functions. The following chapters cover the effects of mindfulness training (Verhaeghen this volume), music training (Swaminathan and Schellenberg this volume), and physical training (Pothier and Bherer this volume), respectively.

The focus of the fifth book section is on the applied perspective. Katz et al. (this volume) present the state-of-the-art regarding individual differences in the effectiveness of cognitive training and the role of motivational processes. Promising ways to apply cognitive training, especially working memory training, in the educational context are discussed by Alloway et al. (this volume). Focusing on cognitive training as a tool against age-related decline in cognitive abilities, Falkenstein and Gajewski (this volume) summarize training-related neurophysiological changes in older adults and relate them to a discussion of data from EEG training studies with elderly workers. Also with a focus on older adults, Belleville et al. (this volume) present different types of cognitive training and show their training and transfer effects in patients with mild cognitive impairments (MCI).

While the previous sections largely focus on past findings in cognitive training research with a strong theoretical perspective, the final section draws conclusions for future research. That is, Colzato and Hommel (this volume) discuss future developments in this area. For instance, they emphasize the need to develop more specific theories guiding cognitive training programs. With this emphasis, they conclude the theoretical perspective of this book and pave the way for future studies on the effects of cognitive training.

The area of cognitive training is a dynamically and fast-growing research area that is increasingly incorporated into scientific teaching and education. The sections of this book should provide comprehensive overviews of state-of-the-art research in cognitive training. They address students and researchers of all academic levels (i.e., from undergraduates to professors) as well as professionals in applied contexts (e.g., teachers, clinicians) by outlining empirical findings and methodological approaches of cognitive training research in different populations, age groups, and cognitive domains. We hope that this volume not only serves to summarize the current state of research but also inspires new exiting, well-designed, and informative studies in this fast-growing scientific field. One of the largest potentials in this area of research lies in the fact that it is very multidisciplinary, integrating research from cognitive, neuropsychological, developmental, educational, and medical science on a theoretical, methodological, and applied level. We believe that this potential may be used in future studies to uncover the cognitive and neural mechanisms underlying training-induced performance benefits and to design adaptive, individually tailored training interventions that can by applied in various contexts, including scientific, educational, and clinical settings.

References

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: A meta-analysis. *Psychonomic Bulletin & Review, 22*(2), 366–377.

Ball, K., & Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science, 218*(4573), 697–698.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612–637.

Bavelier, D., Levi, D. M., Li, R. W., Dan, Y., & Hensch, T. K. (2010). Removing brakes on adult brain plasticity: From molecular to behavioral interventions. *The Journal of Neuroscience, 30*(45), 14964–14971.

Bherer, L., Erickson, K. I., & Liu-Ambrose, T. (2013). A review of the effects of physical activity and exercise on cognitive and brain functions in older adults. *Journal of Aging Research, 2013*, Article ID 657508. doi:10.1155/2013/657508

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults a meta-analytic study. *Psychological Science, 14*(2), 125–130.

Detterman, D. K., & Sternberg, R. J. (1993). *Transfer on trial: Intelligence, cognition, and instruction*. Westport: Ablex.

Fahle, M. (2004). Perceptual learning: A case for early selection. *Journal of Vision, 4*(10), 879–890.

Green, C. S., & Bavelier, D. (2008). Exercising your brain: A review of human brain plasticity and training-induced learning. *Psychology and Aging, 23*(4), 692–701.

Green, C. S., Strobach, T., & Schubert, T. (2014). On methodological standards in training and transfer experiments. *Psychological Research, 78*(6), 756–772.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2008). Enrichment effects on adult cognitive development can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest, 9*(1), 1–65.

Hindin, S. B., & Zelinski, E. M. (2012). Extended practice and aerobic exercise interventions benefit untrained cognitive outcomes in older adults: A meta-analysis. *Journal of the American Geriatrics Society, 60*(1), 136–141.

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 older adults. *Psychological Science, 25*(11), 2027–2037.

Karr, J. E., Areshenkoff, C. N., Rast, P., & Garcia-Barrera, M. A. (2014). An empirical comparison of the therapeutic benefits of physical exercise and cognitive training on the executive functions of older adults: A meta-analysis of controlled trials. *Neuropsychology, 28*(6), 829–845.

Li, R., Polat, U., Makous, W., & Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nature Neuroscience, 12*(5), 549–551.

Lustig, C., Shah, P., Seidler, R., & Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: A review and future directions. *Neuropsychology Review, 19*(4), 504–522.

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science, 15*(8), 511–514.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207–217.

Schubert, T., Strobach, T., & Karbach, J. (2014). New directions in cognitive training: On methods, transfer, and application. *Psychological Research, 78*(6), 749–755.

Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist, 50*(2), 138–166.

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*(1), 13–24.

Strobach, T., Salminen, T., Karbach, J., & Schubert, T. (2014). Practice-related optimization and transfer of executive functions: A general review and a specific realization of their mechanisms in dual tasks. *Psychological Research, 78*(6), 836–851.

Toril, P., Reales, J. M., & Ballesteros, S. (2014). Video game training enhances cognition of older adults: A meta-analytic study. *Psychology and Aging, 29*(3), 706–716.

Wiesel, T. N., & Hubel, D. H. (1965). Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *Journal of Neurophysiology, 28*(6), 1029–1040.

Part I
Basic Concepts and Methodology

Methods and Designs

Florian Schmiedek

Introduction

Researchers who aim to investigate the effectiveness of cognitive trainings can draw on well-established methodology for the evaluation of behavioral interventions in psychology and education (Murnane and Willett 2010; Shadish et al. 2002). Doing so, they face a long list of potential issues that can be characterized as threats to different types of the validity of findings. Here, the most common and relevant threats, as well as possible methodological approaches and study design elements to reduce or rule out these threats in the context of cognitive training studies, will be discussed.

The commonly preferred design for investigating cognitive training interventions is one with random assignment of a sample of participants to training and control groups with pre- and posttest assessments of a selection of tasks chosen to represent one or more cognitive abilities that the training might potentially improve. Significantly larger average improvements on such outcome measures in the training than in a control group are taken as evidence that the training benefits cognition. Such a design indeed clears out a number of potential issues. Certain problems that arise when evaluating cognitive trainings, however, require solutions that go beyond, or modify, commonly used of-the-shelf study design elements. For example, the inclusion of no-treatment control groups for ruling out threats to internal validity and the use of single tasks as outcome measures of transfer effects are associated with certain deficits. In the following, methodological problems and challenges will be discussed along the established typology of statistical conclusion validity, internal and external validity, as well as construct validity (Shadish et al. 2002).

F. Schmiedek (✉)
German Institute for International Educational Research (DIPF),
Frankfurt am Main, Germany
e-mail: schmiedek@dipf.de

T. Strobach, J. Karbach (eds.), *Cognitive Training*,
DOI 10.1007/978-3-319-42662-4_2

Statistical Conclusion Validity

Statistical conclusion validity refers to whether the association between the treatment and the outcome can be reliably demonstrated. Such demonstration is based on inferential statistics, which can provide evidence that observed differences between experimental groups in posttest scores, or in pretest-to-posttest changes, are unlikely to be due to sampling error (i.e., one group having higher scores simply by chance). Given that existing training studies mostly have relatively small sample sizes (with experimental groups of more than 30–40 participants being rare exceptions), the statistical power to do so often is low, and findings are in danger of being difficult to replicate and being unduly influenced by outliers and violations of statistical assumptions.

Furthermore, and in light of recent discussions about the replicability of findings and deficient scientific standards in psychological research (e.g., Maxwell et al. 2015), there is the problem that low power might increase researchers' propensity to lapse into fishing-for-effect strategies. Given that (a) the researchers' desired hypothesis often will be that a training has a positive effect, (b) training studies are resource-intensive, and (c) the non-registered analysis of data allows for a number of choices of how exactly to be conducted (Fiedler 2011), it has to be considered a danger that such choices (like choosing subsamples or subsets of outcome tasks) are made post hoc in favor of "finding" significant effects and thereby invalidate results of inferential test statistics. In combination with publication biases that favor statistically significant over nonsignificant results, such practices in a field with typically low power could lead to a distorted picture of training effectiveness, even in meta-analyses. A general skepticism should therefore be in place regarding all findings that have not been replicated by independent research groups. Regarding the danger of fishing-for-effects practices, preregistration of training studies, including the specific hypotheses and details of data preparation and analysis, are a possible solution, which is well established in the context of clinical trials. In general, effort should be invested to increase statistical power and precision of effect size estimates. Besides large enough sample sizes, this also includes ensuring high reliability of outcome measures and of treatment implementation.

Internal Validity

Internal validity, that is, a study's ability to unambiguously demonstrate that the treatment has a causal effect on the outcome(s), deserves getting a strong weight when judging the quality of intervention studies. It involves ruling out alternative explanations for within-group changes (including, e.g., practice effects, maturation, or statistical regression to the mean from pretest to posttest) and/or between-group differences (e.g., systematic selection effects into the treatment condition). Common reactions to these problems are requests to (a) use a control group that allows to estimate the size of the effects due to alternative explanations and (b) randomly assign participants into the different groups. While intact random assignment assures that the mean differences between groups can be unbiased estimates of the *average causal effect* of the treatment (Holland 1986), several cautionary notes are at place regarding this "gold standard" of intervention studies.

First, the unbiasedness of the estimate refers to the expected value. This does not rule out that single studies (particularly if sample sizes are small) have groups that are not well comparable regarding baseline ability or other person characteristics that might interact with the effectiveness of the training. Therefore, the amount of trust in effect size estimates should only be high for studies with large samples or for replicated (meta-analytic) findings. For single studies with smaller samples, matching techniques based on pretest scores can help to reduce random differences between groups that have an effect on estimates of training effects.

Second, the benefits of randomization get lost if the assignment is not "intact," that is, if participants do not participate in the conditions they are assigned to or do not show up for the posttest. Such lack of treatment integrity or test participation can be associated with selection effects that turn an experiment into a quasi-experiment—with all the potential problems of confounding variables that can affect the estimate of outcome differences.

Third, formal analysis of causal inference based on randomized treatment assignment (Holland 1986) shows that the interpretation of mean group differences as average causal effects is only valid if participants do not interact with each other in ways that make individual outcomes dependent on whether or not particular other participants are assigned to the treatment or the control condition. While this is unlikely to pose a problem if training is applied individually, it could be an issue that has received too little attention in studies with group-based interventions—where interactions among participants might, for example, influence motivation.

Whenever treatment assignment cannot be random, due to practical or ethical considerations, or when randomization breaks down during the course of the study, careful investigation of potential selection effects is required. This necessitates the availability of an as-complete-as-possible battery of potential confounding variables at pretest. If analyses of such variables indicate group differences, findings cannot unambiguously be attributed to the treatment. Attempts to remedy such group differences with statistical control techniques is associated with strong conceptual (i.e., exhaustiveness of the available information regarding selection effects) and statistical assumptions (e.g., linearity of the relation with the outcome) and should therefore be regarded with great caution. An alternative to regression-based control techniques is post hoc matching and subsample selection based on propensity score analyses (Guo and Fraser 2014). This requires sample sizes that are typically not available in cognitive training research, however. Beneficial alternative design approaches for dealing with situations in which randomization is not possible, or likely to not stay intact, are available, like regression discontinuity designs or instrumental variable approaches (Murnane and Willett 2010), but have received little attention in cognitive training research so far.

Construct Validity

While the demonstration of causal effects of the treatment undoubtedly is a necessity when evaluating cognitive trainings, a strong focus on internal validity and randomization should not distract from equally important aspects of construct validity. Addressing the question of whether the investigated variables really represent the theoretical constructs of interest, construct validity is relevant for both, the treatment and the outcome measures.

Regarding the treatment, high internal validity does only assure that one or more aspects that differentiate the treatment from the control condition causally influence the outcome. It does not tell which aspect of the treatment it is, however. Given the complexity of many cognitive training programs and the potential involvement of cognitive processes as well as processes related to motivation, self-concept, test anxiety, and other psychological variables in producing improvements in performance, the comparison to so-called *no-contact control conditions* typically cannot exclude a number of potential alternative explanations of why an effect has occurred. In the extreme case, being in a no-contact control condition and still having to re-do the assessment of outcome variables at posttest is so demotivating that performance in the control group declines from pre- to posttest. Such a pattern has been observed in several cognitive training studies and renders the interpretation of significant interactions of group (training vs. control) and occasion (pretest vs. posttest) as indicating improved cognitive ability very difficult to entertain (Redick 2015). As from a basic science perspective, the main interest is in effects that represent plastic changes of the cognitive system, "active" control conditions therefore need to be designed, which are able to produce the same non-focal effects, but do not contain the cognitive training ingredient of interest. This is a great challenge, however, given the number and complexity of cognitive mechanisms that potentially are involved in processing of, for example, working memory tasks and that can be affected by trainings (von Bastian and Oberauer 2014). For many of these mechanisms, like the use of certain strategies, practice-related improvements are possible, but would have to be considered exploitations of existing behavioral flexibility, rather than extensions of the range of such behavioral flexibility (Lövdén et al. 2010). If motivational effects are partly due to the joy of being challenged by complex tasks, it also will be difficult to invent tasks of comparably joyful complexity but little demand on working memory. In addition to inventive and meticulous creation of control conditions, it is therefore necessary to assess participants' expectations, task-related motivation, and noncognitive outcomes, before, during, and after the intervention.

Regarding the outcome variables, construct validity needs to be discussed in light of the issue of transfer distance and the distinction between skills and abilities. When the desired outcome of a training is the improvement of a specific skill or the acquisition of a strategy tailored to support performing a particular kind of task, the assessment of outcomes is relatively straightforward—it suffices to measure the trained task itself reliably at pre- and posttest. As the goal of cognitive trainings typically is to improve an underlying broad ability, like fluid intelligence or episodic memory, demonstrating improvements on the practiced tasks is not sufficient, however, as those confound potential changes in ability with performance improvements due to the acquisition of task-specific skills or strategies. It is therefore common practice to employ transfer tasks that represent the target ability but are different from the trained tasks. The question of how different such transfer tasks are from the trained ones is often answered using arguments of face validity and classifications as "near" and "far" that are open to criticism and difficult to compare across studies. What seems far transfer to one researcher might be considered near transfer by another one. Particularly if only single tasks are used as outcome measure for a cognitive ability, it is difficult to rule out alternative explanations that explain improvements with a task-specific *skill*, rather than with improvements in the underlying *ability* (see, e.g., Hayes et al. 2015, or Moody 2009).

The likelihood of such potential alternative explanations can be reduced if the abilities that a training is thought to improve are operationalized with several heterogeneous tasks that all have little overlap with the trained tasks and are dissimilar from each other in terms of paradigm and task content. The analysis of effects can then be conducted on the shared variance of these tasks, preferably using confirmatory factor models. This allows to analyze transfer at the level of latent factors that represent the breadth of the ability construct, replacing the arbitrary classification of "near vs. far" with one that defines "narrow" or "broad" abilities by referring to well-established structural models of cognitive abilities (Noack et al. 2009). If transfer effects can be shown for such latent factors, this renders task-specific explanations less likely.

External Validity

External validity encompasses the generalizability of a study's results to other samples, as well as to other contexts, variations of the intervention's setting, and different outcome variables. As few training studies are based on samples that are representative for broad populations, mostly little is known regarding generalizability to different samples. Furthermore, as findings for certain training programs are only rarely replicated by independent research groups, we only have very limited evidence so far regarding the impact of variations of the context, setting, and of the exact implementation of cognitive trainings. As one rare exception, the Cogmed working memory training (http://www.cogmed.com/) has been evaluated in number of studies by different research groups and with diverse samples. This has resulted in a pattern of failed and successful replications of effects that has been reviewed as providing little support for the claims that have been raised for the program (Shipstead et al. 2012a).

Similarly, generalizations of effects for certain transfer tasks to real-life cognitive outcomes, like everyday competencies and educational or occupational achievement, are not warranted, unless shown with direct measures of these outcomes.