New Directions in Training Designs

Aaron Cochrane and C. Shawn Green

Contents

Abstract Cognitive training is a rapidly expanding domain, both in terms of academic research and commercial enterprise. Accompanying this expansion is a continuing evolution of training design that is driven by advances on various fronts. Foundational learning principles such as spacing and interleaving have always, and continue to, inform the design of training for cognitive improvements, yet advances are constantly made in how to best instantiate these principles in training paradigms. Improvements in hardware have allowed for training to be increasingly immersive (e.g., using virtual reality) and to include multifaceted measurements and dynamics (e.g., using wearable technology and biofeedback). Further, improved training algorithms and gamification have been hallmarks of advances in training software. Alongside the development of these tools, researchers have also increasingly established cognitive training as a more coherent field through an emerging consensus regarding the appropriate methods (e.g., control group selection and tasks to test generalization) for different possible studies of training-related benefits. Hardware, software, and methodological developments have

A. Cochrane (✉) · C. S. Green
Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA
e-mail: akcochrane@wisc.edu; cshawn.green@wisc.edu

T. Strobach, J. Karbach (eds.), *Cognitive Training*,
https://doi.org/10.1007/978-3-030-39292-5_3

quickly made cognitive training an established field, yet many questions remain. Future studies should address the extent and type of generalization induced by training paradigms while taking into account the many possible patterns of improvements from training. Patterns of benefits vary across training types as well as individuals, and understanding individual differences in training benefits will help advance the field. As the field of cognitive training matures, the upcoming years are set to see a proliferation of innovation in training design.

Introduction

Cognitive training has existed, in something like its current form, for only a few decades. It is therefore not surprising that, like many fledgling domains, the field continues to be rife with rapid change and advancement. This is especially true given the fact that, unlike many other areas of psychology, many questions in the cognitive training sphere are not of purely academic or theoretical nature. Instead, the potential for the commercialization of cognitive training has frequently pushed current practices as well (although not always with methodology to demonstrate efficacy to match – see below). Concurrently, advancements in computer hardware as well as training software have facilitated research and applications of training in increasingly diverse and ecologically valid contexts. Here we focus on recent advances (e.g., improvements in hardware and software capabilities), endemic challenges (e.g., as related to methods for controlling for expectation effects or how to best translate from broad principles of effective learning to specific instantiations in cognitive training paradigms), and future directions in the field of cognitive training.

Cognitive Training: Built upon Foundational Principles of Learning and Neuroplasticity

Although the field of cognitive training continues to develop, in most cases these improvements are situated squarely within the existing work in the learning sciences. For instance, one of the best single predictors of the extent to which a new skill will be learned is time on task (e.g., the "total time hypothesis," Ebbinghaus 1913). Simply put, the more time that individuals spend on a given task, the more they will learn. It is thus not surprising that this appears to be the case in perceptual and cognitive training as well (Jaeggi et al. 2008; Stafford and Dewar 2014; Stepankova et al. 2014), with some recent work truly pushing the envelope in terms of length of training (Schmiedek et al. 2010). Next, while the total amount of time spent learning is clearly important, not all time is equally well spent. One of the most replicated findings in the learning literature is that learning is more efficient (i.e., in terms of improvement per unit time) when training sessions are distributed rather than massed in time (Baddeley and Longman 1978). While this general finding is likely due to multiple mechanisms working in concert (e.g., decay of irrelevant learning, homeostatic regulation associated with sleep, etc.), it nonetheless indicates a clear design recommendation for cognitive training: many shorter training sessions are better than fewer longer training sessions. Indeed, the potential importance of both total training time and distribution of practice can be seen in comparing the results of two similar studies utilizing video game training – one that employed 50 total hours of training with each training session generally lasting around 1 hour (Green et al. 2010), and which produced generally positive results, and a second that employed up to 40 fewer hours of training and sessions that lasted up to four times as long, and which produced largely null results (Van Ravenzwaaij et al. 2014).

Another principle of effective learning common across domains is that of adaptivity of the to-be-learned material. In many cases this adaptivity takes the form of increasing difficulty as learner ability increases. That is, as a participant becomes proficient at completing training tasks, those tasks should become more difficult – thus keeping the participant at the edge of what they are able to handle (Deveau et al. 2015; Vygotsky 1981). Feedback during learning is also key. While a full discussion of the topic requires more nuance than is possible here, generally speaking learning is more effective when learners are provided with immediate and informative feedback related to their performance (Seitz and Dinse 2007). Finally, many other principles of effective learning find their empirical roots, at least partially, in the study of neuroplasticity (see also Wenger and Kühn, this volume). For instance, elegant basic science work has delineated the importance of various neuromodulatory systems in activating neuroplastic brain states (e.g., the cholinergic system via the nucleus basalis (Kilgard et al. 1998), and the dopaminergic system via the ventral tegmental area (Bao et al. 2001)). This has, in turn, served to strongly underscore the importance of designing training paradigms so as to induce a certain degree of physiological arousal and to make proper use of reward in order to maximize the potential efficacy of the training (Green and Bavelier 2010).

Other core principles that are foundational to the field of cognitive training focus not on the learning of the training tasks themselves, but on the extent to which the learning that occurs generalizes to untrained tasks (Schmidt and Bjork 1992). In essentially all areas of learning there exists a tension between learning that is highly specific to the trained paradigm and learning that transfers to untrained contexts and situations. A host of core learning task characteristics are known to increase the degree to which learning generalizes. Interestingly, most of these characteristics simultaneously decrease the overall rate of improvement. The goal of most cognitive training paradigms is to maximize the extent to which the learning generalizes broadly, and relevant principles of learning might therefore fall under the category of what have been dubbed *desirable difficulties* (Schmidt and Bjork 1992). For example, increases in both overall training heterogeneity and the extent to which training tasks are intermixed improve the generality of learning. Generalization tends to be increased when training is not homogeneous, but instead includes variation (Deveau et al. 2015; Dunlosky et al. 2013; Xiao et al. 2008); note though that effects may vary across populations of interest, see (Karbach and Kray 2009).

Yet, while the principles above have clearly been influential in the development of the paradigms employed in the cognitive training literature, as we will see later in the chapter, (1) it is not always clear how to best instantiate the principles in practice (e.g., how to engender motivation) and (2) these principles can interact in multiple, and sometimes unexpected ways.

Advances in Hardware for Cognitive Training

Before considering the training paradigms themselves, it is worth briefly considering changes in available hardware, as this represents the first bottleneck of training design. Over the past decade portable technology such as tablets have become increasingly common in cognitive training interventions (e.g., Ge et al. 2018; Oei and Patterson 2013; Shin et al. 2016; Wang et al. 2016). Tablets are relatively inexpensive, easy to use across a wide range of age groups, can be readily available for participants to train at their convenience, and can provide continuous updates of data for researchers. They can also be easily paired with wearable technology able to track heart rate, physical activity, and an increasing number of other variables (Piwek et al. 2016). These benefits though are accompanied by a loss in control over the administration of training and, as such, compliance with training regimens may be impossible to perfectly ensure. Even compliant learners may not adhere strictly to training instructions, and many sources of unwanted variance may be completely out of the control of training designers (e.g., screen viewing distance, device volume, and distracting environments). Although improvements in online psychological studies have addressed and mitigated some issues regarding experimental control, there will inevitably be some compromises when training is completed outside of controlled settings (Yung et al. 2015). The use of tablets, cell phones, or other portable devices thus involves accepting a tradeoff between the amount of data that is collected and the variability in the data.

Virtual reality (VR) headsets are another recently-developed type of hardware that has the potential for cognitive training applications. By using VR headsets, training programs can be more aligned with the field of view, depth, and actions of naturalistic settings. While cognitive training research utilizing VR is in its infancy, there have been some attempts to adapt typical monitor-based tasks to 3-dimensional virtual reality (Nyquist 2019; Nyquist et al. 2016). The immersion and ecological validity promised by VR could have the potential to improve many cognitive training paradigms. Barriers to effective deployment of virtual reality training continue to exist, however. Powerful computers are necessary for rendering virtual environments, and even the best computers for VR cannot yet compete with the spatial and temporal resolutions available on high-end monitors. And even as this technology improves, challenges will remain with respect to the human experience of VR. One clear example is nausea; the subtle mismatches between perceptual-motor predictions and simulated realities in VR can compound into debilitating “simulator sickness” (Allen et al. 2016; Kim et al. 2018).

Like virtual reality, wearable technology is increasingly available and likely to play a major role in future studies of cognitive training. Combined effects of physical training and cognitive training have promised greater improvements than either in isolation (Hertzog et al. 2008). Furthermore, even when implementing cognitive training with minimal physical demands, physiological measurements may nonetheless be informative to researchers regarding mediators or moderators of training outcomes. As examples, physical activity and sleep are each linked to neuroplasticity (Atienza et al. 2004; Bavelier et al. 2010; Tononi and Cirelli 2003). For each of these factors measurement with wearable technologies is simple. Even technology formerly relegated to research such as electroencephalography (EEG) is now available in portable formats and has been used in biofeedback-based training paradigms (Shin et al. 2016). As with EEG, increased interest in transcranial direct current stimulation (tDCS) has led to studies of efficacy of tDCS in concert with behavioral cognitive training (Martin et al. 2014; Martin et al. 2013).

Given the possibilities afforded to cognitive training by advances in hardware, the face of training is rapidly changing. Training in the future will likely be designed to be more immersive (e.g., virtual reality or always-available tablets), will integrate a more diverse set of measurements (e.g., heart rate and sleep tracking), many of which can be fed back directly into adaptive training algorithms, and may utilize methods to put the brain in a more plastic state (Hensch 2004; Seitz and Dinse 2007).

Advances in Software for Cognitive Training

One hardware issue not discussed above is the simple increase in computational power that comes with each passing year. This aspect in turn allows ever more complex training algorithms to be implemented (Deveau et al. 2015). Classic training algorithms in perceptual and cognitive fields have relied on unidimensional measures (e.g., correct/incorrect) aggregated across many training trials to determine performance, which then allowed adjustment of difficulty. In contrast, modern training in educational domains has developed more nuanced methods for understanding performance and correspondingly adapting difficulty (Liu et al. 2019; Ritter et al. 2007). In the latter case, interleaved training of various target skills is a straightforward implementation of another well-established principle of learning (e.g., Schmidt and Bjork 1992). The ability to track performance in each of the target skills, and provide on-the-fly adjustment of training demands in order to balance new content with refreshing old content, is a much more difficult task from a training perspective (Zhang et al. 2019). Indeed, in cognitive training research, targets of training are often homogeneous (e.g., only working memory), or trained processes are simply interleaved in a balanced design. This represents an opportunity for cognitive training research to improve as the field matures; while improved assessments and algorithms are increasingly possible, the efficacy of competing assessments and algorithms is still poorly understood. As with educational apps and intelligent tutoring systems, cognitive training can include many principles from basic learning research. These include interleaving, spacing, and adapting training as learners progress through a program. Additionally, personalization of training is a valuable ability facilitated by sensitive on-the-fly assessments of ability.

Possibly the most obvious design trend in cognitive training has been so-called "gamification" (Jaeggi et al. 2011; Squire 2003). Off-the-shelf recreational video games themselves have been used frequently in the cognitive training domain (for a review see Bediou et al. 2018; see also Bediou, Bavelier, and Green, Strobach and Schubert, this volume). These games provide natural instantiations for many of the learning principles discussed earlier and thus are an obvious source material from which designers may develop more dedicated forms of training (Deveau et al. 2015; Gentile and Gentile 2008; Nyquist et al. 2016). For instance, well-designed games produce both external and internal motivation to play, leading to a great deal of time on task. Video games also induce a great deal of physiological arousal and activation of the neural reward systems, which together create a brain state that is capable of efficient learning. Video games often involve a variety of tasks, types of decisions, and varying load on different attention and memory systems. As such, these games conform to the principle of variety and interleaving of learning. By frequently changing the demands placed on players, fast-paced video games are able to produce benefits in overlapping domains (e.g., attention to a wide visual field of view), while avoiding specificity in learning and maintaining adaptive difficulty that supports efficient learning (Deveau et al. 2015).

The increase in gamification has been supported by improved software for developing games or game-like environments. This is in stark contrast to game production in the past which required a set of highly skilled programmers and designers. Ease of game production does not necessarily mean high-quality games, however, and gamification does not directly imply that cognitive training would have the benefits of video games. Gamification should add rewards, engagement, arousal, and/or variety to cognitive training in order to introduce any benefits above, and should go beyond simply training on a cognitive task (Deveau et al. 2015). As noted early however, this may be easier said than done. While creating games has become easier, designing engaging, enjoyable, and effective training games remains challenging. In one test of motivational game-like features in cognitive training of children, Katz et al. (2014) found that none of the motivational game-like features that were implemented produced improvements on training-task learning. There may be various reasons for this outcome, including a highly stimulating base training (i.e., before adding motivational features), distracting effects of features such as points or levels (i.e., that the motivating features took attention away from the critical to-be-learned skills), or an insufficient timescale to detect differences (3 days of training). However, with limited tests of generalization, it may also have been the case that process-level benefits differed between training groups, and these differences were not apparent in the training data. Indeed, as discussed above, a classic finding is that *desirable difficulties* in learning may inhibit initial learning while boosting generalization (Schmidt and Bjork 1992).

Advances in Methods for Studying the Impact of Cognitive Training

There are clearly many outstanding questions regarding the most appropriate and efficacious interventions for given contexts and populations. Yet, many of the deepest questions in the field today concern studies' structural choices and assumptions and best-practice methodologies (see also Könen and Auerswald, Schmiedek, this volume). As an example, while Boot and colleagues have argued that training results are interpretable only if both intervention and control groups improve from pretest to posttest (Boot et al. 2011), Green and colleagues argue that these test–retest effects are theoretically unnecessary, and in fact, reduce the power to observe training-related benefits (Green et al. 2014). As an important step toward establishing a common methodological framework for diverse training paradigms and populations, over 50 leading researchers in the field recently collaborated in the publication of a consensus regarding methodological standards (Green et al. 2019). This section will briefly discuss the four dimensions of relevant methodological issues: control group choice, blinding, randomization, and tests of generalization.

Control Groups

Studies in experimental psychology are only as good as the contrasts utilized, and cognitive training is no exception. In order to demonstrate effectiveness of a training paradigm, and to identify the relevant processes undergoing change, appropriate experimental controls must be implemented. Control group selection in cognitive training is far from simple, and depending on the questions that are being posed, experimenters may choose to maximize the perceptual similarity of the control training with that completed by the experimental training group, to induce similar expectations and/or affective states, to match levels of engagement and interest, or to implement training grounded in alternative hypotheses regarding mechanism or efficacy (Green et al. 2014). The choice of active control is necessarily linked to the specific aims of a study, and there is no one-size-fits-all approach. Such study-specific control design poses difficulty for comparison of results across studies, however, which in turn hinders the ability for the field to move forward. Simply put, because the effects of interest in the field are usually a difference of differences (i.e., changes from pretest to posttest in the experimental group as compared to the pretest to posttest changes in the control group), massive differences in the characteristics of the control group make it difficult-to-impossible to effectively compare and contrast the impact of the experimental training paradigms. Thus, in order to ensure one-to-one comparisons of training effect sizes across studies with varying active control groups, it has recently been suggested that studies should implement nocontact controls in addition to their active control groups (Green et al. 2019). These business-as-usual comparison groups allow for clear qualitative and quantitative matching between effects of varying training regimes and will facilitate future work (Colzato and Hommel, this volume).

Blinding: Managing and Measuring Expectations

Expectation effects refer to changes in studies' outcomes in response to beliefs regarding the purpose or hypothesis of the studies. One well-known example is the placebo effect, in which positive beliefs regarding the efficacy of an intervention lead to beneficial outcomes even in the absence of the proposed mechanism of benefit (e.g., an inert sugar pill producing a similar analgesic effect as acetaminophen). The reduction of these expectation effects is largely accomplished through effective blinding, or ensuring that learners (and experimenters) are unaware of the expectations regarding their condition. For example, in a pain study, participants could be assigned to receive one of the two outwardly identical pills – one of which is a sugar pill, the other being a true analgesic. Because the participants will not know which of the two pills they are receiving, the expected benefit should be matched across groups, and thus any differences in outcome could not be attributed to expectations alone. In the cognitive training domain, it is not possible to produce two outwardly identical paradigms, where one is "inert" (like the sugar pill) and one is "active" (like the true analgesic). The outward appearance of a behavioral training platform is, after all, intractably linked to the extent to which the training is inert or active. As such, the best that can be done in the domain of cognitive training is to devise control experiences that appear plausible as interventions (Green et al. 2019). This is not necessarily trivial. Indeed, it is not even clear how to best measure the success of such attempted blinding (e.g., how to determine what expectations participants in the various groups hold). Advances in this area will therefore be critical for the field going forward.

We note that while minimizing expectation effects is necessary for demonstrating that any experimental training has true efficacy, expectations themselves may be used for the benefit of training once such a demonstration has been made. By intentionally creating expectations and maximizing their influence through conditioning, these expectations may become tools for increasing the effectiveness of training regimens (Green et al. 2019). Benefits of utilizing expectations may be especially pronounced in young populations due to the possibility of compounding long-term effects of small early-life benefits and attitudes (Stanovich 1986). Even if early benefits are "only" placebo effects (e.g., not true improvements in core cognitive processes), these benefits may still have very real positive downstream effects.

Randomization: Ensuring Interpretability of Results

Conventional wisdom in behavioral research is that study participants should be randomly assigned to experimental groups. However, truly random assignment is liable to create inter-group variation at pretest that reduces the interpretability of postintervention results. Given that the intentions of randomization and of group comparisons are each to reduce noise and clarify study-specific differences in behavior (i.e., learning), targeted efforts to match groups' performance on pretests will increase the interpretability of statistical tests of change from pretest to posttest (Green et al. 2014). Several methods exist to establish this masking, ranging from stratified or grouped random sampling (i.e., randomizing group membership after categorizing by other measures such as age or cognitive performance) to condition-difference minimization (i.e., assigning each new participant to whichever condition minimizes the between-condition pretest differences).

But What Is Learned? The Use of Pretest and Posttest Batteries

The target of cognitive training is often a specific process or set of processes. In order to test for changes to this target, or even to detect baseline individual differences, a variety of tasks loading on the target process can be used. By identifying the common component underlying, for example, complex span working memory tasks, individual variation and possible training-related benefits can be better identified (Engle et al. 1999; Green et al. 2014). Null results are likewise strengthened by process-level tests of generalization. By testing generalization to processes that are explicitly not expected to benefit from training, the contrast between null effects and nonnull effects can be used to clarify mechanisms of learning and falsify competing hypotheses. That is, if the mechanism of improvement was simply an increased effort on all tasks, all tests of generalization should benefit uniformly; to the extent that there are some null effects, any nonnull effects are more interpretable.

Despite the benefits of large numbers of pretest and posttest tasks, there are clear limitations. With continued testing fatigue will eventually diminish the quality of behavioral data. Fatigue is especially problematic in lower-functioning populations such as young children or older adults. While normally-functioning young adults may be expected to complete several hours of testing with a uniformly minimal decrement in performance, lower-functioning populations are likely to have a wider variance in their susceptibility to fatigue. In these populations patterns of performance may be shaped by participants' differential abilities to maintain attention and vigilance throughout demanding tasks. Training-related benefits may then be obscured or confounded by individual differences in the ability to complete long task batteries. As such, the size and scope of pretest and posttest batteries should be as large as feasible given the resources, context of training, and population of interest.

Frontiers: Questions and Practices for the Field

Benefits of Training: General or Specific?

All cognitive training is, justifiably, subject to scrutiny regarding the degree to which benefits observed within the training environment also extend to other behaviors. Robust improvements on trained tasks are often accompanied by little or no benefit to untrained tasks. This fact is far from unique to cognitive training; in areas as disparate as math education and visual contrast sensitivity training, learning can be surprisingly specific to the trained task. The lack of generalized benefits observed after using some common "brain training" apps has led to increased scrutiny of cognitive training from the popular press as well as the United States government, with a highly publicized rebuke and fine of one company occurring in 2016 (Federal Trade Commission 2016).

Even in tightly controlled studies, the generalization of cognitive benefits is sometimes not observed. However, we caution against interpretations of absences of generalization as "failures." Rather, specificity of a given training paradigm provides important information about the limiting cases in which cognitive training may or may not be efficacious. This may be relevant, for example, when matching interventions to appropriate populations. As discussed above, in young populations it may be the intention of training to improve scores on (and, ideally, the lifelong downstream consequences of) these specific cognitive abilities (see de Vries, Kenworthy, Dovis, and Geurtz, Johann and Karbach, Rueda et al., this volume).

Re-framing our understanding of generalization or specificity is only a small part of the larger problem: evidence regarding efficacy of training paradigms has been sparse. This problem is exacerbated by varying methodologies in training which make cross-study comparisons problematic at best; only by developing an aggregated estimate of efficacy can the understanding of generalization be advanced. Attempts have been made at aggregation, often with conflicting results (Au et al. 2015; Melby-Lervåg and Hulme 2013). The inter-study variation that causes these divergences is a key motivation for the push toward methodological consensuses mentioned above. Understanding generalization as a function of training design necessitates more data using common methods.

Multiple Forms of Generalization

There is also ambiguity regarding the expected mechanisms of generalization. While "transfer of learning" has typically been understood as immediate benefits observed in untrained contexts or tasks, there are a variety of ways in which initial training can benefit later performance (Barnett and Ceci 2002). Generalization of learning may also cause multiplicative benefits rather than additive benefits to generalized performance, leading to patterns of transfer that appear as learning to learn rather than immediate improvements. That is, even if performance in a test of generalization is not immediately benefited, performance may improve faster on tests of generalization than they would have prior to training (Kattner et al. 2017).

Delayed benefits in training generalization are a largely under-explored area, yet these effects are mechanistically aligned with the theoretical basis of cognitive training. If the targets of training are core cognitive abilities, it is possible that the benefits of these enhanced abilities would not be evident immediately on novel tasks due to task-specific factors (e.g., idiosyncratic interference from prior experience). Indeed, at different points in learning, separate processes may be constraining performance (Ackerman and Cianciolo 2000). This underscores the need to understand learning and generalization as time-evolving processes; the changes and generalized benefits of learning may be evident at some times and obscured at other times by other limiting processes (Rebok et al. 2014).

A different delayed training benefit may occur due to enhancement of cognitive abilities associated with more rapid learning in novel contexts. The locus of this change could be one of the various possibilities (e.g., faster speed of processing and improved perceptual template; (Bejjanki et al. 2014)). In this case of learning to learn, improvements on tests of generalization would be delayed due to the mechanism of generalization causing a divergence in performance with increased experience on a test of generalization. That is, if training causes an improvement in learning ability, there is little reason to believe that immediate benefits would be observed on novel tasks, but benefits should quickly become apparent with time. This is likely the case, for example, in cognitive benefits observed from action video game playing (Green et al. 2010).

Yet another cause for delayed generalization effects of cognitive training concerns the developmental timescales on which benefits are supposed to emerge. Early in the lifespan, interventions may have downstream effects due to trained children's ability to succeed in early school years, leading to an improved ability to use school resources themselves for improvement (Stanovich 1986). This is, for example, one theoretical motivation behind many early-childhood interventions outside the purely cognitive domain (e.g., Head Start, Ludwig and Phillips 2008). Later in life, too, interventions may have long-lasting effects by mitigating the downward trajectory of cognitive decline (Hertzog et al. 2008; Rebok et al. 2014; Willis et al. 2006).

In each of these cases of delayed generalization effects, the training should be designed appropriately for the observation of training-related benefits. That is, if there are very few observations of potential generalization (such as low trial numbers in cognitive assessments), there would inevitably be insufficient evidence to determine the presence or absence of delayed generalization effects. Likewise, if long-term developmental trajectories may be influenced by training, then assessments on the appropriate timescale must be implemented.

Alongside appropriate training design, evidence regarding generalization should also be considered using methods that allow for detection of delayed effects and dissociation between immediate and delayed generalization. In the case of learning to learn, in particular, it is important to understand the time course of performance on generalization tasks. In this case the mechanism of generalization manifests as a difference in performance that may be evident only after some a priori indeterminate amount of task experience. It is important, then, to approach generalization as a dynamically unfolding process in which training-related benefits may cause a divergence in performance between trained and untrained individuals (Bray and Dziak 2018). Each time (e.g., trial within a task) is therefore an important point at which generalization may be occurring, and generalization performance can be quantitatively modeled as a time-dependent process. By utilizing this by-trial modeling of performance, four possible outcomes can be dissociated: (1) immediate generalization, (2) delayed generalization (e.g., learning to learn), (3) lack of generalization, or (4) both (1) and (2). In the absence of time-dependent models of generalization, superficially unrelated factors such as generalization-task number of observations may obscure the effects of training (Kattner et al. 2017).

Variance in Outcomes: Individual Differences in Training Benefits

In an insightful analysis of learning data from several classic studies, Heathcote et al. (2000) noted that a canonical power-law function of learning did not exist in any individual learner, but the power function was instead an artifact of averaging performance across individuals. A similar possibility has the potential for reducing the accuracy of inferences regarding the efficacy of cognitive training. That is, group-level estimates of training efficacy may obscure individual-level changes in cognitive abilities (Bürki et al. 2014). Certain factors, such as genetics, attitudes toward training, or compliance may even mediate positive effects of training on cognition (Colzato et al. 2014; Jaeggi et al. 2014). Further, group-level estimates of change may hide the possibility that some learners actually perform worse at posttest than at pretest. This pattern is obviously not desirable, but it is a very important addition to the field's understanding of training design and efficacy. That is, in realworld applications, training should ideally benefit each learner. While ubiquitous success is an unlikely outcome, it is possible that the time spent training takes away from the time spent on other beneficial activities (e.g., rehabilitation exercises or classroom exercises). If certain populations are unresponsive to training and are better served by "business-as-usual," then the main effects of training vs. control groups can hide this mechanistic nuance. Thus, as far as what is feasible, researchers should consider individual trajectories of improvement, and should develop tools for identifying individuals who do not benefit from the training intervention. This will be an important aspect of adaptivity algorithms in future applied training contexts. As with any intervention that should be stopped when a lack of efficacy has been demonstrated in a certain patient (e.g., administration of medication), cognitive training must not algorithmically “keep trying” when an individual is not responsive to the intervention.

The power of individual-level data is also an important feature of understanding the results of training. While statistical power to detect the effect of an intervention is often understood in terms of the number of participants in a study, the features of the study itself also influence the power to detect any training-related effects. That is, there is a clear resource allocation trade-off between studying few people trained following the best practices, and studying many people trained using practices with less likelihood to detect any effect. In fact, depending on the timescale on which plasticity in target processes would change, it is possible that training programs of different lengths (e.g., 3 days vs. 25 days) would not simply be quantitatively different in their power to detect training-related benefits, but also be qualitatively different in the types of benefits able to be induced in that timescale. Quantitative reviews of various training studies may exacerbate the problem. That is, if studies in a meta-analysis are weighted according to the number of participants, then studies that have emphasized the participant number over training integrity would be more influential in drawing conclusions. Even if other variables are statistically controlled for (e.g., time training per session, number of different training tasks, or number of sessions), there is little way to know whether the target processes of various studies are qualitatively similar enough to justify quantitative aggregation. Nonetheless, to the degree that methods such as the total time and spacing are qualitatively similar across participants and studies, hierarchical and meta-analytic statistical models provide the ability to simultaneously estimate both individual-level and aggregate parameter estimates that can indicate the efficacy of training paradigms.

The Next Generation of Training Design: Integrated, Informed, and More Powerful than Ever

The direction of cognitive training design is toward increasingly engaging, available, and well-informed programs. Recent consensus statements from scientists in the field provide guidelines for theoretically understanding, and methodologically implementing, studies for the advancement of the field (Green et al. 2019; Max Planck Institute for Human Development and Stanford Center on Longevity 2014). These statements encourage healthy skepticism regarding the results of any single program or study, but they also encourage innovation through the recognition that studies and paradigms have widely differing intentions and populations. Advances may be attempted through the use of novel hardware, software, or even cognitive targets of training, and even null results add to the community’s understanding of training mechanisms and efficacy (Green et al. 2014).

References

Ackerman, P. L., & Cianciolo, A. T. (2000). Cognitive, perceptual-speed, and psychomotor determinants of individual differences during skill acquisition. *Journal of Experimental Psychology: Applied, 6*, 259–290.

Allen, B., Hanley, T., Rokers, B., & Green, C. S. (2016). Visual 3D motion acuity predicts discomfort in 3D stereoscopic environments. *Entertainment Computing, 13*, 1–9.

Atienza, M., Cantero, J. L., & Stickgold, R. (2004). Posttraining sleep enhances automaticity in perceptual discrimination. *Journal of Cognitive Neuroscience, 16*, 53–64.

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. D., & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics, 21*, 627–635.

Bao, S., Chan, V. T., & Merzenich, M. M. (2001). Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature, 412*, 79–83.

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.

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. *Journal of Neuroscience, 30*, 14964–14971.

Bediou, B., Adams, D. M., Mayer, R. E., Tipton, E., Green, C. S., & Bavelier, D. (2018). Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychological Bulletin, 144*, 77–110.

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.

Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video games improve perception and cognition? *Frontiers in Psychology, 2*, 226.

Bray, B. C., & Dziak, J. J. (2018). Commentary on latent class, latent profile, and latent transition analysis for characterizing individual differences in learning. *Learning and Individual Differences, 66*, 105–110.

Bürki, C. N., Ludwig, C., Chicherio, C., & de Ribaupierre, A. (2014). Individual differences in cognitive plasticity: An investigation of training curves in younger and older adults. *Psychological Research, 78*, 821–835.

Colzato, L. S., van den Wildenberg, W. P. M., & Hommel, B. (2014). Cognitive control and the COMT Val^{158}Met polymorphism: Genetic modulation of videogame training and transfer to task-switching efficiency. *Psychological Research, 78*, 670–678.

Deveau, J., Jaeggi, S. M., Zordan, V., Phung, C., & Seitz, A. R. (2015). How to build better memory training games. *Frontiers in Systems Neuroscience, 8*, 243.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*, 4–58.

Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology*. University Microfilms. Available at: https://books.google.com/books?id=oRSMDF6y3l8C

Engle, R. W., Tuholski, S. W., Laughlin, J., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable model approach. *Journal of Experimental Psychology: General, 128*, 309–331.

Federal Trade Commission. (2016). *Lumosity to pay $2 million to settle FTC deceptive advertising charges for Its "Brain Training" program*. Available at: https://www.ftc.gov/system/files/documents/cases/160105lumoslabsstip.pdf

Ge, S., Zhu, Z., Wu, B., & McConnell, E. S. (2018). Technology-based cognitive training and rehabilitation interventions for individuals with mild cognitive impairment: A systematic review. *BMC Geriatrics, 18*, 213.

Gentile, D. A., & Gentile, J. R. (2008). Violent video games as exemplary teachers: A conceptual analysis. *Journal of Youth and Adolescence, 37*, 127–141.

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.

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

Green, C. S., Bavelier, D., Kramer, A. F., Vinogradov, S., Ansorge, U., Ball, K. K., ... & Facoetti, A. (2019). Improving methodological standards in behavioral interventions for cognitive enhancement. *Journal of Cognitive Enhancement, 3*, 2–29.

Heathcote, A., Brown, S., & Mewhort, D. J. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin and Review, 7*, 185–207.

Hensch, T. K. (2004). Critical period regulation. *Annual Review of Neuroscience, 27*, 549–579.

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*, 1–65.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105*, 6829–6833.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences (USA), 108*, 10081–10086.

Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory and Cognition, 42*, 464–480.

Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science, 12*, 978–990.

Kattner, F., Cochrane, A., Cox, C. R., Gorman, T. E., & Green, C. S. (2017). Perceptual learning generalization from sequential perceptual training as a change in learning rate. *Current Biology, 27*, 840–846.

Katz, B., Jaeggi, S., Buschkuehl, M., Stegman, A., & Shah, P. (2014). Differential effect of motivational features on training improvements in school-based cognitive training. *Frontiers in Human Neuroscience, 8*, 242.

Kilgard, M. P., & Merzenich, M. M. (1998). Cortical map reorganization enabled by nucleus basalis activity. *Science, 279*, 1714–1718.

Kim, H. K., Park, J., Choi, Y., & Choe, M. (2018). Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied Ergonomics, 69*, 66–73.

Liu, R., Stamper, J., Davenport, J., Crossley, S., McNamara, D., Nzinga, K., & Sherin, B. (2019). Learning linkages: Integrating data streams of multiple modalities and timescales. *Journal of Computer Assisted Learning, 35*, 99–109.

Ludwig, J., & Phillips, D. A. (2008). Long-term effects of head start on low-income children. *Annals of the New York Academy of Sciences, 1136*, 257–268.

Martin, D. M., Liu, R., Alonzo, A., Green, M., Player, M. J., Sachdev, P., & Loo, C. K. (2013). Can transcranial direct current stimulation enhance outcomes from cognitive training? A randomized controlled trial in healthy participants. *International Journal of Neuropsychopharmacology, 16*, 1927–1936.

Martin, D. M., Liu, R., Alonzo, A., Green, M., & Loo, C. K. (2014). Use of transcranial direct current stimulation (tDCS) to enhance cognitive training: Effect of timing of stimulation. *Experimental Brain Research, 232*, 3345–3351.

Max Planck Institute for Human Development and Stanford Center on Longevity. (2014). *A consensus on the brain training industry from the scientific community (full) – Stanford Center on Longevity*. Available at: http://longevity.stanford.edu/a-consensus-on-the-brain-training-industry-from-the-scientific-community-2/. Accessed 29 July 2019.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*, 270–291.

Nyquist, J.B. (2019). NeuroTrainer labs. *NeuroTrainer labs*. Available at: https://neurotrainer.com. Accessed 27 July 2019.

Nyquist, J. B., Lappin, J. S., Zhang, R., & Tadin, D. (2016). Perceptual training yields rapid improvements in visually impaired youth. *Scientific Reports, 6*, 37431.

Oei, A. C., & Patterson, M. D. (2013). Enhancing cognition with video games: A multiple game training study J. J. Geng (ed). *PLoS One, 8*, e58546.

Piwek, L., Ellis, D. A., Andrews, S., & Joinson, A. (2016). The rise of consumer health wearables: Promises and barriers. *PLoS Medicine, 13*, e1001953.

Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H. Y., King, J. W., ... & Willis, S. L. (2014). Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults. *Journal of the American Geriatrics Society, 62*, 16–24.

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive tutor: Applied research in mathematics education. *Psychonomic Bulletin and Review, 14*, 249–255.

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

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Frontiers in aging neuroscience, 2*, 27.

Seitz, A., & Dinse, H. R. (2007). A common framework for perceptual learning. *Current Opinion in Neurobiology, 17*, 148–153.

Shin, M. S., Jeon, H., Kim, M., Hwang, T., Oh, S. J., Hwangbo, M., & Kim, K. J. (2016). Effects of smart-tablet-based neurofeedback training on cognitive function in children with attention problems. *Journal of Child Neurology, 31*, 750–760.

Squire, K. (2003). Video games in education. *Int. J. Intell. Games and Simulation, 2*, 49–62.

Stafford, T., & Dewar, M. (2014). Tracing the trajectory of skill learning with a very large sample of online game players. *Psychological Science, 25*, 511–518.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407.

Stepankova, H., Lukavsky, J., Buschkuehl, M., Kopecek, M., Ripova, D., & Jaeggi, S. M. (2014). The malleability of working memory and visuospatial skills: A randomized controlled study in older adults. *Developmental Psychology, 50*, 1049–1059.

Tononi, G., & Cirelli, C. (2003). Sleep and synaptic homeostasis: A hypothesis. *Brain Research Bulletin, 62*, 143–150.

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.

Vygotsky, L. S. (1981). *Mind in society: The development of higher psychological processes Nachdr*. Cambridge, MA: Harvard Univ. Press.

Wang, Z., Zhang, H., Wang, J., Wang, X., Yu, X., & Wang, H. (2016). Tablet-based multi-domain cognitive training increases the parahippocampus volume in patients with amnestic mild cognitive impairment. *Alzheimer's & Dementia, 12*, P509–P510.

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., ... & Wright, E. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *JAMA, 296*, 2805–2814.

Xiao, L. Q., Zhang, J. Y., Wang, R., Klein, S. A., Levi, D. M., & Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Current Biology, 18*, 1922–1926.

Yung, A., Cardoso-Leite, P., Dale, G., Bavelier, D., & Green, C. S. (2015). Methods to test visual attention online. *Journal of Visualized Experiments, 96*, 52470.

Zhang, P., Zhao, Y., Dosher, B. A., & Lu, Z.-L. (2019). Evaluating the performance of the staircase and quick change detection methods in measuring perceptual learning. *Journal of Vision, 19*, 14.