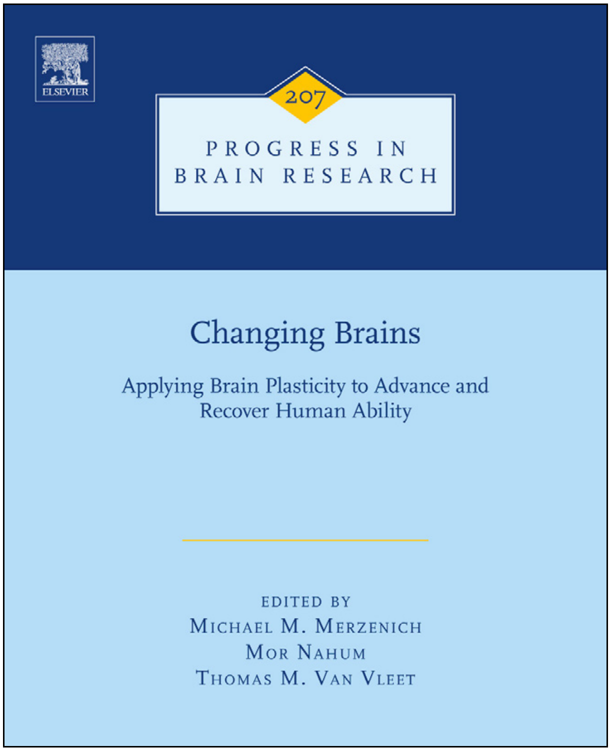

From Nori Jacoby, Merav Ahissar, What does It take to Show that a Cognitive Training Procedure is Useful? A Critical Evaluation. In Michael M. Merzenich, Mor Nahum, Thomas M. Van Vleet editors: *Progress in Brain Research*, Vol. 207, Burlington: Academic Press, 2013, pp. 121-140.
ISBN: 978-0-444-63327-9

CHAPTER 5

What does It take to Show that a Cognitive Training Procedure is Useful?

A Critical Evaluation

Nori Jacoby[*,†,1], **Merav Ahissar**[‡]

[*]*Interdisciplinary Center for Neural Computation, The Hebrew University, Jerusalem, Israel*

[†]*Music Department, Bar Ilan University, Ramat Gan, Israel*

[‡]*ELSC Center for Brain Research and the Department of Psychology, The Hebrew University, Jerusalem, Israel*

[1]*Corresponding author: Tel.: +972-548132872; Fax: +972-25881159, e-mail address: nori.jacoby@alice.nc.huji.ac.il*

Abstract

Individuals substantially improve with training, indicating that a large degree of plasticity is retained across ages. In the past 20 years, many studies explored the ability to boost cognitive skills (reasoning, linguistic abilities, working memory, and attention) by training with other tasks that exploit limited cognitive resources. Indeed, individuals with long-term training on challenging skills (musicians and action video gamers) show impressive behavior on related tasks (linguistic and visual attention, respectively). However, a critical evaluation of training studies that last weeks to months shows typically mild effects, mainly with respect to control groups that either did not practice or practiced with less challenging, rewarding, or exciting conditions. These findings suggest that future training studies should evaluate these factors carefully and assess whether they mainly impact the testing sessions or actual longer-term skills, and whether their impact can be further strengthened. The lack of a comprehensive theory of learning that integrates cognitive, motivational, and alertness aspects poses a bottleneck to improving current training procedures.

Keywords

perceptual learning, working memory, musical education, cognitive training, generalization, action video games, positive affect

Progress in Brain Research, Volume 207, ISSN 0079-6123, http://dx.doi.org/10.1016/B978-0-444-63327-9.00004-7

The potential promise in adult training is exciting. Yet, in spite of impressive evidence for behavioral improvement and for neural plasticity, our understanding of the underlying mechanisms and of the relations between training procedures and their outcomes is extremely limited. Particularly challenging is the extent of transfer from trained to untrained tasks and contexts. Most theories of perceptual learning (e.g., Dosher and Lu, 2007; Seitz and Watanabe, 2005) address specific experimental contexts and hence do not have comprehensive predictions regarding tasks or procedures that are expected to increase generalization. An exception is Reverse hierarchy theory (RHT; though see also a recent version of a theory by Dosher et al., 2013) that addresses the relations between training procedures and their expected outcomes (Ahissar and Hochstein, 2004; Ahissar et al., 2009; Hochstein and Ahissar, 2002). However, RHT's predictions are also quite limited, since it assumes a comprehensive understanding of the nature of hierarchical representations of stimuli, which we still largely lack (Ahissar, 2001). Additionally, it does not integrate motivational, alertness, and related factors, whose impact is probably substantial and perhaps also interacts with the training procedure.

Understanding these general factors, which relate to the state of the participant while performing the task, is important for the formation of a systematic conceptualization of the anticipated gains. These factors include time of day (May and Hasher, 1998), general emotional state (Isen, 2008), motivation and alertness (Boot et al., 2013; Duckworth et al., 2011; Langer et al., 2010), training duration (Censor et al., 2006), limits on response duration (Green et al., 2010), and intervals between training sessions and sleep (Censor et al., 2006; Karni et al., 1994; Korman et al., 2007; Stickgold, 2005; Stickgold et al., 2000). The effect of these parameters may be larger than that of the systematically studied parameters, as we shall show in this chapter.

Given our limited theoretical understanding of the comprehensive learning experience, the aim of this chapter is to define operative guidelines for evaluating the success of past and future training studies, with respect to their goals. Based on our review, we propose further directions for future studies. An important distinction, which was typically ignored in previous studies, is the difference between clinical and theoretical goals. Typically, the design of a given study cannot optimize both. For example, if a procedure improves the performance of a population on a poorly performed range of tasks, it may be a successful rehabilitation procedure even if the mechanisms are not understood, and their control was very crude. However, well-controlled studies that yield a very small improvement compared with careful controls may have a theoretical impact but perhaps no immediate application. The two goals are closely related in longer-term development, where improving training procedures would gain immensely if their underlying mechanisms would be understood. Defining priorities when designing the studies would facilitate our understanding and hence our long-term ability to improve training protocols.

1 WHAT IS A GOOD CONTROL?

Finding the cause of a training-induced difference between groups is almost impossible, since training protocols differ in more than one parameter

1.1 Single-Arm Studies

The main drawback of having only one training group with no control and simply comparing pre- and post-training performance (e.g., Banai and Ahissar, 2010; Gibson et al., 2011; Holmes et al., 2010; Kronenberger et al., 2011; McArthur, 2007; McNab and Klingberg, 2007; Mezzacappa and Buckner, 2010) is that significant differences may reflect the improvement due to prior experience in the pretest itself. Individuals who perform a task for a second time, typically (e.g., Cane and Heim, 1950; Windle, 1954, for a review of the experimental implications see Campbell et al., 1963), though not ubiquitously (e.g., Lemay et al., 2004), perform better, particularly when retested with the very same items. An improvement may be found even when participants are unaware of this repetition, as reported already by Kolers (1976), where participants trained on reading inverted script were faster reading the pages they were trained with compared with novel pages, even when tested more than a year after being trained. Indeed, the scores in standard tests may improve in a retest even if the interval between tests is large (Matarazzo et al., 1980). Still, if test–retest effects for the tested interval are *normed* and improvement in the trained group is clearly larger than these norms, a comparison group is not crucial for attributing some benefit to the training procedure itself.

1.2 Within-Group Control

Another alternative for a control group is a within-group control. This is the prevailing concept in perceptual learning studies. In these studies, one group is trained with a given set of parameters and tested with a different set of parameters that are presumably as difficult as the trained ones (e.g., Ahissar and Hochstein, 1993, 1997; Cohen et al., 2013; Dosher and Lu, 1998, 1999; Karni and Sagi, 1991; Lu et al., 2000; Polat and Sagi, 1994; Polat et al., 2004; Tsushima et al., 2008; Watanabe et al., 2002). If performance is worse in the untrained conditions, then learning is shown to be specific along the tested dimension. However if performance transfers to novel conditions, one cannot delineate the limits of learning. Therefore, this approach was used in studies aimed at showing specificity and was not adopted in studies aimed to show generalization. However, if learning leads to a complex pattern of specificities and transfer, then rehabilitation studies could also use within-group comparisons, testing a well-designed series of tasks. For example, evaluation of generalization of training with a specific working memory task to generally enhanced working memory skills can be obtained by testing (post-training) the trained working memory task with completely novel stimuli, and by administering novel working memory tasks. If there is improvement in the untrained task (i.e., scores are similar to those attained on the trained one), the sources of this transfer would indeed be hard to tap. Yet, most learning is quite specific to the combination of trained task and stimuli (e.g., Ahissar and Hochstein, 1993). Thus, the transfer to novel stimuli, with which participants were not previously tested but whose relative difficulty is expected to be similar to that of trained ones, is likely to be informative. See for example, a recent use of control tasks within the experimental paradigm in Anguera et al. (2013), which would provide a systematic characterization of the learning process. Very few training studies (e.g., Klingberg et al., 2002, 2005; Olesen et al., 2003; Schmiedek et al., 2010) used this within-group control for showing that transfer at least encompassed very different stimuli or even very closely related tasks. Even when such a test battery was used, it mainly served as a collection from which significantly improved tasks could be pulled out rather than as a systematic characterization of the pattern of specificity and generalization that the training induced.

In the typical case, particularly when participants' performance in other tasks is not systematically characterized and there are no suitable test–retest norms, a control test–retest group should be added, in order to verify that the training process contributed to the post-training improvement. If there is indeed a significant difference between the trained and untrained groups, it means that some training-related improvement occurred, but we cannot tell what this improvement should be attributed to, as the differences between groups' procedures encompass many aspects. Still, if the gains of the trained group are large and much larger than those of a "no-contact" control, it means that this procedure is useful for applied purposes, even though the intergroup difference may be due to a very general aspect of practice (e.g., "positive psychology"; see Yang et al., 2013; for a review, see Isen, 2008). An important aspect, which is difficult to determine, is whether the potentially general effects underlying group differences reflect a difference in learning or merely a difference in performance during the testing session itself (for a critical discussion of these issues, see Boot et al., 2011; Duckworth et al., 2011; Shipstead et al., 2012). Since the dissociation between short- and long-term effects of general factors is difficult to determine, one should perhaps equate within-testing factors of "positive psychology" for the two groups. In general, the main limit of having a no-contact control is that we cannot tell how to improve the training procedure. However, if the intergroup effects are large, namely, the protocol as a whole was effective, a useful continuation could be to determine the mechanisms that underlie the intergroup difference. For example, it would be useful to know whether this improvement could have been obtained without any training at all, for example, by encouraging the tested participants. Indeed, a recent meta-analysis found that studies with greater incentive yield significantly (~0.6 standard deviations) larger scores in intelligence tests, previously considered to robustly reflect individual cognitive abilities (Duckworth et al., 2011).

1.3 Active Control Group

A more informative control group is a group that actually goes through some alternative training procedure, termed "active control" (e.g., Colom et al., 2010; Dahlin, 2011; Holmes et al., 2009; Jaeggi et al., 2011; Klingberg et al., 2002, 2005; Richmond et al., 2011; Shavelson et al., 2008; Thorell et al., 2009; Van der Molen et al., 2010; Zhao et al., 2011; for a review see Shipstead et al., 2012). But what should an active control group train with? In the case of working memory training, a common solution is an active control group that is trained with an easy, fixed-stimuli task, whereas the test group is trained with an adaptive, challenging working memory task. However, when post-training performance differs between the trained groups (e.g., Jaeggi et al., 2008; Klingberg et al., 2002, 2005), it is not clear what such improvement should be attributed to. Does it stem from training a working memory task? Would one get similar gains if paid for successful performance instead of training? Would one get similar gains in any challenging task? Perhaps the participants trained with a nonadaptive procedure perform worse because they are bored and have lost interest. In this type of control, the potentially crucial role of motivation, excitement, and challenges that differ between the two training protocols is overlooked when inter group differences in the outcomes are specifically attributed to training with a working memory task. Again, one should note that here too, group differences in the testing session may be related to the different levels of anticipation (how well do we expect to perform) of the two tested groups. Such differences may affect performance, even in perceptual tasks. For example, by affecting the "mind setting" of tested individuals, so that they expect to perform better, participants *actually* perform better, even in visual acuity tasks, considered a robust measure of the visual system (Langer et al., 2010).

1.4 Active Control Equated for Difficulty

An active control whose training procedure is aimed at equating the levels of challenge and engagement (and perhaps anticipation for success) of the trained groups was administered in a novel study of working memory training (Redick et al., 2013). In this case, the control group was trained with a different task (visual search), with no working memory component and with different stimuli. However, the trained task was challenging and engaging. Under these conditions, there was no group difference in post-training working memory or intelligence scores, suggesting that working memory was not the important factor in training. Interestingly, the post-training scores of the actively trained groups did not differ from a third, no-contact control group. Namely, in this study, there was no general effect either. Given that in this group, participants did *not* expect to benefit from the training procedure (based on this groups' previous criticism of the benefit of these training studies; Shipstead et al., 2012), this null result implies the potential impact of the researchers' and hence participants' anticipations. This aspect is difficult to avoid, and it underlies the reasoning of double-blind procedures. Yet, double-blind procedures do not apply when different behavioral procedures are administered. Given that "mind setting" and "positive psychology" (Yang et al., 2013) were shown to affect performance on the tasks that most procedures aim to improve, it is of importance that at least the experimenter who performs the tests will not be aware of the training procedure assigned to the specific participant.

Finally, from a clinical perspective, when a new rehabilitation-aimed training procedure is proposed, the most informative control is the standard of care or the best current treatment or training procedure. For example, if there is a protocol known to have some benefit (e.g., phonological training for reading disability), an assessment of a novel procedure should be conducted with respect to this protocol. This type of comparison had been conducted only to a limited degree (e.g., when training for language impairments; e.g., see Hook et al., 2001). However, from a purely scientific perspective, taking an existing procedure may be informative only to a limited extent since different training procedures typically differ on how they are technically administered (e.g., groups vs. one on one and in person vs. computer games) and in the amount of challenge and excitement they pose (e.g., video games vs. sentence repetition; for a review and discussion of social conditioning effects in the context of video games, see Boot et al. (2011)). These parameters may have a larger impact on the training gains compared with those that uniquely characterize the hypothesis underlying the novel protocol, and therefore, differences in the impact of two different procedures do not necessarily stem from a difference in the validity of their underlying hypotheses.

2 INTERPRETING CORRELATIONS IN GAINS

Performance gains in initially correlated tasks are expected to be correlated even with no transfer

The typical motivation of training studies aimed at boosting cognitive skills is to show that an improvement in one task (the trained task) results in an improvement in the other (the "transfer" task). In this context, it seems natural to examine whether the gain in the trained task (performance in the posttest minus the performance in the pretest) is correlated with the gain in the transfer task/s and interpret such a correlation as reflecting a transfer effect. Indeed, several studies used such a correlation as evidence supporting transfer between tasks (e.g., Anguera et al., 2013; Banai and Ahissar, 2009; Chein and Morrison, 2010; Green and Bavelier, 2003; Jaeggi et al., 2011; Schmiedek et al., 2010). However, a significant correlation between gains in the trained and untrained tasks can be obtained without any transfer. For example, if pre-training scores are initially positively correlated, but the post-training scores are completely independent from pre-training scores and also uncorrelated, then a simple calculation (based on the mathematical fact that the pre-training correlation contributes to the overall correlation in gains) shows that gains are positively correlated, even when scores in the pre- and post-training sessions were independent (i.e. in a simulation—chosen from independent Gaussian distributions of pre- and post-training scores). Therefore, correlations in "gains" (differences between pre- and post-training performance) *per se* cannot be used as evidence for transfer.

When two tasks are initially correlated and only one task is trained so that its performance is consequently improved, we expect that individuals who improve to a greater extent on the trained task will be individuals with lower pre-training scores and higher post-training scores on the trained task (selection bias). Due to the pre-training correlation between the trained and untrained tasks, these individuals will tend to have lower pre-training scores also in the untrained task. If following training, performance is similar for the selected (higher training gain) and unselected (lower training gain) individuals then the calculated gain of the selected individuals will be larger due to this selection bias even when there is no transfer. Such a bias can account for the transfer, that is, correlation in gains, reported in Jaeggi et al. (2011).

Jaeggi et al. (2011) trained two groups of third graders, one (*N*=32) on general knowledge and vocabulary ("active control") and the other (*N*=32) on a spatial working memory task. This initial study had null results. Namely, the two groups did not differ in their general intelligence scores, either before or after training, suggesting no transfer. However, rather than acknowledging null results, Jaeggi et al. (2011) divided the group trained on the working memory task into two subgroups (median split), with high and low gains in working memory, respectively. They found that the subgroup that had larger working memory gains also had larger intelligence gains. The main basis for the "transfer" claim is that individuals who had larger gains in the trained task had larger "gains" in the transfer tasks (e.g., standard tasks for measuring intelligence). Yet the group of individuals who had larger gains in the trained task did not have higher post-training scores in the untrained intelligence tasks. In this study, the reported correlation could be fully accounted for quantitatively without assuming any transfer of learning, that is, solely on the basis of biased statistics (see for details Jacoby and Ahissar, submitted).

The same concept of correlation in gains as supporting evidence for transfer is used in many other training studies. For example, Chein and Morrison (2010) claimed "a strong and statistically significant relationship between enhancement of trained participants' spatial complex working memory span and their improvement in reading comprehension [$r=.49$, $p<0.005$]." Similarly, in an attempt to support a transfer effect between working memory and fluid intelligence in children with ADHD, Klingberg et al. (2002) claimed that "the association between the reasoning task and the working memory tasks is further substantiated by the significant correlation between improvement on the visuo-spatial working memory task and improvement on Raven's progressive matrices" (p. 789). Schmiedek et al. (2010) reported a significant correlation between latent trained factors of episodic memory and untrained ones and used it as supportive evidence for transfer effect. Banai and Ahissar (2009) observed a marginally significant correlation between improvement in a perceptual task (two-tone discrimination) and working memory (Spearman rho 0.59, $p=0.07$) and interpreted it as "support (for) the suggestion that the improvement in working memory scores is specifically related to the 2-tone discrimination training." Green and Bavelier (2003) used the correlation between improvements in game scores and improvements in attentional skills as support for a transfer claim. Most recently, Anguera et al. (2013) trained elderly individuals with a dual-task condition of an action video game and used the correlation between improvement in this trained condition ("multitasking") and working memory gains to support the claim of transfer from training multitasking in action video games to enhanced working memory capacity (see Anguera et al., 2013 figure 3d on page 99).

In summary, the intuitive interpretation of correlation in gains is misleading and statically erroneous. In order to show transfer when studies choose to train one task chosen from a set of correlated tasks, studies should do one of the following analyses suggested below.

The first is to show that the experimental group had larger gains, namely, a significant interaction on a repeated-measure ANOVA (with session as within subject factor and group as between subject factor) using F-statistics. We emphasize that division to groups should not be based on *post hoc* criteria. This procedure asserts that a real and significant effect occurred and that the training group improved more than the control group. If this effect is large, it also has a significant practical value. Some papers used a non repeated *t*-test for comparing the gain scores (post minus pre) between groups. This is in fact the same procedure, as the *t*-score here is the square root of the F-test score of the repeated-measure ANOVA group × session interaction (Knapp and Schafer, 2009). An example for significant interaction using ANOVA can be found in Anguera et al. (2013). Note, however, that their results are only significant when both the no-contact and active control groups were used in order to compute the interaction.

Another similar approach is to use analysis of covariance, with posttest as dependent variable and pretest as covariate. This method uses the same data as the first method, and in fact, it is possible to transform the test score of one into that of the other (Knapp and Schafer, 2009). Even though these two methods are not identical, as it is possible that only one of the tests would be significant (Lord, 1967), in most practical scenarios, they are expected to produce comparable results. For example, Anguera et al. (2013) used both methods and obtained very similar results.

The third alternative is to account for the bias that the correlation structure implies when computing the gains' correlations. Namely, show that the obtained effect is larger than that expected due to the biases that are induced by the implicit correlations (see Jacoby and Ahissar, submitted). In any event, studies should supply the full correlation matrix (which should include all relevant pretests and postests) to enable an evaluation of the dependencies before and after the training process. This matrix is typically not provided.

3 WHAT DO STUDIES OF "NEURAL CORRELATES" TELL US?

For clinical purposes, behavioral changes are the only ones that validate success

Recent advances in imaging brought many novelties to brain research. In principle, imaging studies can be informative with respect to the nature of neural modifications that underlie behavioral improvement and as such provide important insights. However, not all imaging studies provide insights regarding the mechanisms underlying behavioral improvement, even though almost all studies are interpreted as if they do. Moreover, changes in brain responses are often interpreted as reflecting "real" changes compared with behavior-only experiments, even when behavioral modifications are the real goal of the study.

Many studies assessing the neural correlates of behavioral improvement only show that known brain correlates of performance indeed change when performance changes. For example, several studies (e.g., Anguera et al., 2013; François et al., 2012; Moreno et al., 2009) showed that when behavior improves following training, ERP components, known to correlate with performance and measured while participants perform the task, are increased. A recent example is Anguera et al. (2013) who trained older adults with a lab-version dual-task condition of an action video game. They showed that age-related EEG activity (midline–frontal theta power and frontal–posterior theta coherence) was increased after training. However, the EEG was measured on a task that is almost identical to the trained task. They did not show this enhanced EEG activity on a different multitasking activity. For the purpose of rehabilitation, the important aspect is the ability to improve age-related neural mechanisms in general. However, in the Anguera et al. study, the other condition that was used for measuring multitasking—a novel dual-task paradigm—was not significantly improved in the group trained with the dual-task compared with controls trained with one task at a time.

Another example that demonstrates this issue is the study by François et al. (2012), who characterized the effects of music training in children that practiced music for 2 years starting from the age of 8 years old. They showed significant changes in performance in speech segmentation compared with a control group that practiced painting. Following training, they also showed increased activation of frontal electrodes at 450–550 ms post-stimulus onset, when performing this task. In their previous work, Francois and Schön (2011) associated this component with N400-like component, known to be correlated with linguistic skills (Elger et al., 1997). Thus, the observation that improved behavior is coupled with an increase in the expected ERP component is reassuring, but does not point to any specific underlying mechanism. Such observations would be more revealing regarding underlying mechanisms if there were several known ways to attain better performance, accompanied by different brain correlates, or if the same known brain correlate were shown to increase in untrained conditions.

A different type of overinterpretation of imaging and ERP studies stems from their greater susceptibility to posterior statistics and hence to spurious significant effects, because they typically involve many more measures (voxel clusters, voxel combinations, ERP electrodes, or electrode combinations) compared with behavioral studies. An extreme example of this case is a recent study by Neville et al. (2013), who administered a family-based training program designed to improve "brain systems for selective attention" in preschool children of lower socioeconomic status. A main claim for the success of the program to boost cognitive measures compared with an active control group was the group-specific change in ERP indices used for assessing selective attention. Neville et al. used a selective auditory attention paradigm and measured ERP with 29 electrodes. This myriad of electrodes allows for many *post hoc* selections for analysis. In this study, the choice of electrodes for analyses made the difference. Neville's group used this ERP index in a series of previous studies. In all these studies (Coch et al., 2005; Sanders et al., 2006; Stevens et al., 2008, 2009, 2013), they showed that the *anterior and medial electrodes* convey this component, and therefore, they discarded posterior electrodes. But in Neville et al. (2013), the experimental group showed no modifications in anterior electrodes. Based on previous analyses, this means a null ERP result. However, these participants showed changes in posterior electrodes, and Neville et al. conducted an anterior–posterior ANOVA analysis, in which the experimental group "showed a significant increase in this response (time, $P < 0.05$) that was largest *at posterior recording sites*."

4 HOW LARGE ARE THE MAGNITUDES OF GAINS IN UNTRAINED CONDITIONS?

For clinical purposes, training-induced gains should be large with respect to the amount of time dedicated to training

Assuming that transfer to untrained desired tasks is significant, a further question is whether the amount of improvement is substantial with respect to the initial impairment or with respect to the amount of time dedicated to the training process.

In the case of improving working memory, the typically reported gains are in the range reported in studies that do not apply training, but administer positive priming cues or other incentives. In these studies, it was shown that individuals who had greater incentives had higher performance than those with reduced incentives. The magnitude of these effects, as evaluated in the meta-analysis by Duckworth et al. (2011), is about 0.64 standard deviations. For example, Yang et al. (2013) used a priming paradigm (positive affect) where subjects were given an unexpected gift consisting of a small bag containing hard candies and then tested on cognitive tasks. They reported medium to large effects of the positive affect on their subjects' performance: the effect sizes were 0.58, 1.06, and 1.03 standard deviations for remote associate, and two different measures of working memory.

In training studies, typical effect sizes are also in the range of 0.5–1 standard deviations, when positive transfer is reported. For example, Klingberg et al. (2005) trained 53 children with ADHD, either with an adaptive working memory program or with a nonadaptive version. They found a small–medium difference between the groups' gains following training. For example, the size of the transfer effects (Cohen's *d*) for digit span, response time in Stroop paradigm, and Raven's task (measuring fluid intelligence) was 0.59, 0.34, and 0.45, respectively. In addition, they measured these tests on a follow-up session. The results showed that these effects were mostly retained, though slightly decreased (0.57, 0.25, and 0.30 for digit span, Stroop, and Raven's task, respectively). These effects are within the range of social conditioning and positive priming (see Boot et al., 2011). Jaeggi et al. (2008) trained 70 young students: 35 completed a working memory training program of 8–19 days and the other participants composed a no-contact control group. The trained group improved their intelligence scores (Cohen's $d=0.65$) more than the control group (Cohen's $d=0.25$). Lilienthal (2013) studied 52 students who either trained with an adaptive n-back working memory task or trained with a nonadaptive task or were only tested with a similar inter-test interval. They found that participants in the adaptive training group also showed improvement on a span task, which measures the capacity of the focus of attention. The effect sizes computed from their reported data was 0.17, −0.06, and 0.86 for the no-contact, nonadaptive, and adaptive groups, respectively. Other studies that conducted similar training procedures (e.g., Boot et al., 2008; Owen et al., 2010; Redick et al., 2013) did not find any transfer and naturally found only a small or very small effect size (0–0.5).

Other studies trained participants on action video games and measured their subsequent performance on a range of tasks, often measuring the efficiency of spatial attention. This training was motivated by the observation that avid players of action video games, who have thousands of hours of experience, show substantial advantages compared with nonplayers in a range of tasks of visual attention, particularly those that require fast responses to briefly presented stimuli (Achtman et al., 2008; Castel et al., 2005; Chisholm et al., 2010; Dye et al., 2009a,b; Green and Bavelier, 2006a,b, 2007, 2008; Hubert-Wallander et al., 2011; Hubert-Wallander et al., 2011; Li et al., 2009, 2010, 2011; Mishra et al., 2011). For example, Green and Bavelier (2006b) trained individuals who do not play action video games on either an action video game or on a control game (Tetris). They showed that training facilitates the processing of multiple objects in a multiple-object-tracking task. The effect sizes (calculated from their figures) were high—$d=1.01$ and $d=1.39$—compared with the control $d=0.25$ and $d=0.01$ for tracking four and six objects, respectively. Li et al. (2010) trained nonaction video players intensively (50 h over 9 weeks) and found reduced lateral masking with intervals of 60 and 90 ms, with effect sizes (calculated from their figures) of $d=0.77$ and $d=0.80$, respectively. There was almost no effect in the control group, who trained on a nonaction video game (The Sims 2). The same group (Li et al., 2009) also trained nongamers with action video games and showed that contrast sensitivity, which is routinely assessed in clinical evaluations of vision, had improved ($d=1.08$ for frequency of 6 cycles/deg) compared with a much smaller improvement ($d=0.17$) in a control group that trained with a nonaction video game (The Sims 2). A recent study (Anguera et al., 2013) trained elderly individuals on a dual-task condition of a video game (NeuroRacer game). The effect sizes of delayed-recognition working memory task and test of variables of attention was 0.98 and 0.89, respectively, compared to the no-contact group. An additional group trained in a very similar task but did not have a multitasking component. The effect size compared with this group was medium (0.67 and 0.46 to the two earlier-mentioned tasks, respectively). However, this effect was obtained for two tasks of a larger testing battery. One of the additional tasks was another dual-task paradigm for which no significant improvement was found in the main group compared with the other groups.

Though these effect sizes are not small, they would probably not be statistically resilient to corrections for the total number of tasks assessed. Moreover, other groups who trained participants on similar action video games did not obtain transfer to other cognitive tasks (e.g., Boot et al., 2008; Lee et al., 2012). These differences might stem from the use of slightly different experimental (e.g., the duration and color of the masking stimulus in the functional field of view task) and recruitment strategies. This suggests that here too, even though control groups also trained with challenging games, the effects depend on additional components, whose nature should be further studied. In another study, Boot et al. (2013) measured the expectations of subjects ($N=200$) who watched videos of either an action video game or a control game (Tetris or The Sims). They were then asked if they think that training for many hours with the video games presented would improve their performance on a list of other cognitive tasks that were also presented with detailed explanations and videos. Their results show surprisingly high agreement between subjects' expectation and the extent of transfer reported in the literature (Green and Bavelier, 2012). These results suggest that indirect effects, such as subject's expectation, may be crucial for the replication of training effects.

A somewhat similar case is that of musical education. Most people who study music do so for pleasure rather than as a means for attaining other skills, just as is the case for avid players of action video games. Yet, expert musicians were reported to have enhanced verbal and phonological skills (Anvari et al., 2002; Besson et al., 2007; Butzlaff, 2000; Forgeard et al., 2008; François et al., 2012; Moreno et al., 2009) and higher intelligence (Lynn et al., 1989; Schellenberg, 2004, 2006). Although these may reflect a selection bias for people playing music, an exciting hypothesis is that their musical education contributed to their linguistic skills (Kraus and Chandrasekaran, 2010) or intelligence (Schellenberg, 2006).

The causal relations between training and enhanced performance were studied in several longitudinal studies (Costa-Giomi, 1999; Flohr, 1981; Forgeard et al., 2008; François et al., 2012; Hurwitz et al., 1975; Hyde et al., 2009; Moreno et al., 2009; Morrongiello and Roes, 1990; Schellenberg, 2004; Thompson et al., 2004). Most training studies reported significant improvements after medium (few months) to long (about 2 years) trainings. For example, Schellenberg (2004) reported a small to medium effect of increased IQ, which was significantly more prominent in 6-year-old children taking music lessons (for children that received keyboard and voice music lessons, the effect sizes were 0.69 and 0.70, respectively), compared with control groups that studied drama ($d=0.37$) or did not participate in any special training program ($d=0.40$). François et al. (2012) studied the effect of music training in children that practiced music starting at the age 8 years old. They showed significant improvement in speech segmentation in children practicing music, compared with a control group that practiced painting; the effect size of the improvement in the music group compared with the painting was around 1 both for 1 and 2 years of training. However, Rickard et al. (2012) reported results with two large studies ($N=127$ and $N=100$ children) with school-based music instruction that did not show a cognitive enhancement. In a correlation study that we conducted (Banai and Ahissar, 2013) among third-grade children, we found that auditory thresholds were more correlated with verbal memory among children that were not musically trained. Among children with about 1 year of musical education, *only* auditory thresholds were better than those expected when intelligence was statistically controlled for, whereas higher verbal memory scores could be accounted for by higher general intelligence scores. Thompson et al. (2004) showed that music keyboard lessons do not enhance performance in identifying emotional content of semantically neutral speech more than control (drama lessons) even after a year of training. These mixed results leave the question of the impact of musical education (1–2 years) on linguistic skills largely open. The effect sizes on tasks that are not closely related to those directly trained are small. For example, in a meta-analysis of over 30 music intervention studies by Standley (2008), the average effect size on reading skills was small ($d=0.32$).

In a recent study (Hai, Granot, and Ahissar, submitted), we asked whether adult musicians (~24 years old), who had prolonged and intensive musical training, may still have language and reading difficulties. We found that there are indeed poor readers among musicians. Their performance in auditory, sensory, and sensorimotor (tapping) tasks was as good as that of their musician peers and superior to that of their nonmusician peers. But their working memory abilities were impaired compared with their musician peers. Overall, dyslexia seems milder among musicians. Still, these findings indicate that musical education was not sufficient for enhancing their working memory skills to an extent that they would no longer pose a bottleneck to reading.

In summary, the evidences from working memory, action video games, and music training are quite similar. Relatively short periods of targeted training do not provide a substantial widespread cognitive enhancement. The magnitude of these effects is typically in the range of positive priming and enhanced motivation, implying that the sources underlying the transfer effects should be further studied.

5 CONCLUSIONS AND FUTURE DIRECTIONS

Individuals who practice a given task show substantial improvement. This is a very consistent observation across tasks and across populations, which provides a clear indication of adults' brain plasticity, as further shown in brain changes observed in imaging studies. However, the extent of generalization to novel tasks is typically limited, and the means for enhancing generalization are still not understood.

Several studies report mild improvement in general intelligence scores, working memory, or attentional tasks following training. But the larger and more consistent effects are those compared with no training (e.g., Anguera et al., 2013; Jaeggi et al., 2008) or training with less demanding tasks (Jaeggi et al., 2011; Klingberg et al., 2002, 2005) or perhaps less active ones (e.g., Franceschini et al., 2013; Green and

Bavelier, 2003, 2008, 2012). This pattern suggests that perhaps, the more important aspect for generalization resides in participants' and perhaps experimenters' emotional state (such as motivation, arousal, and reward) rather than in the specific choice of cognitive tasks in these studies. An interesting question is whether these emotional effects interact with cognitive states. In order to examine these intricate questions systematically, we must add to the relevant test dimensions the emotional and social aspects (including accurate measures of engagement and challenge) and systematically evaluate their contribution.

Another important yet unresolved question is that of the target population. Rehabilitation is typically oriented toward individuals with specific impairments, compared with the general population. But we still do not know whether individuals with specific deficits are more or less likely to gain from training and whether similar training procedures are likely to be useful to different populations. Here too, we lack a theoretical account for addressing this question. On the one hand, the dynamic range that may be relevant for individuals with specific difficulties is perhaps larger. In the general population, it is often the case that individuals who start with poorer performance improve more than others (e.g., see Duckworth et al., 2011). On the other hand, individuals who did not gain sufficiently with "ecological" training conditions are perhaps more resilient to practice, as suggested by the phenomenon of dyslexic musicians (Hai, Granot, and Ahissar, submitted). For example, Oganian and Ahissar (2012) compared auditory and reading skills of adequate readers and dyslexics with and without musical education. They found that among adequate readers, auditory and reading skills were higher in individuals with musical training. However, among dyslexics, auditory and reading skills did not differ between individuals with and without formal musical education.

Taken together, the limited success of training studies in attaining generalization would perhaps be improved when a broader range of factors will be systematically considered and when the outcome of these studies will be integrated into a comprehensive theory that will guide the design of future procedures.

References

Achtman, R., Green, C., Bavelier, D., 2008. Video games as a tool to train visual skills. Restor. Neurol. Neurosci. 26, 435–446.

Ahissar, M., 2001. Perceptual training: a tool for both modifying the brain and exploring it. Proc. Natl. Acad. Sci. U. S. A. 98, 11842–11843.

Ahissar, M., Hochstein, S., 1993. Attentional control of early perceptual learning. Proc. Natl. Acad. Sci. U. S. A. 90, 5718.

Ahissar, M., Hochstein, S., 1997. Task difficulty and the specificity of perceptual learning. Nature 387, 401–406.

Ahissar, M., Hochstein, S., 2004. The reverse hierarchy theory of visual perceptual learning. Trends Cogn. Sci. 8, 457–464.

Ahissar, M., Nahum, M., Nelken, I., Hochstein, S., 2009. Reverse hierarchies and sensory learning. Philos. Trans. R. Soc. B: Biol. Sci. 364, 285–299.

Anguera, J., Boccanfuso, J., Rintoul, J., Al-Hashimi, O., Faraji, F., Janowich, J., Kong, E., Larraburo, Y., Rolle, C., Johnston, E., 2013. Video game training enhances cognitive control in older adults. Nature 501, 97–101.

Anvari, S.H., Trainor, L.J., Woodside, J., Levy, B.A., 2002. Relations among musical skills, phonological processing, and early reading ability in preschool children. J. Exp. Child Psychol. 83, 111–130.

Banai, K., Ahissar, M., 2009. Perceptual learning as a tool for boosting working memory among individuals with reading and learning disability. Learn. Percept. 1, 115–134.

Banai, K., Ahissar, M., 2010. On the importance of anchoring and the consequences of its impairment in dyslexia. Dyslexia 16, 240–257.

Banai, K., Ahissar, M., 2013. Musical experience, auditory perception and reading-related skills in children. PLoS One 8 (9), e75876.

Besson, M., Schön, D., Moreno, S., Santos, A., Magne, C., 2007. Influence of musical expertise and musical training on pitch processing in music and language. Restor. Neurol. Neurosci. 25, 399–410.

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 Psychol. 129, 387–398.

Boot, W.R., Blakely, D.P., Simons, D.J., 2011. Do action video games improve perception and cognition? Front. Psychol. 2.

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. Perspect. Psychol. Sci. 8, 445–454.

Butzlaff, R., 2000. Can music be used to teach reading? J. Aesthet. Educ. 34, 167–178.

Campbell, D.T., Stanley, J.C., Gage, N.L., 1963. Experimental and Quasi-Experimental Designs for Research. Houghton Mifflin, Boston.

Cane, V., Heim, A.W., 1950. The effects of repeated retesting: III. Further experiments and general conclusions. Q. J. Exp. Psychol. 2, 182–197.

Castel, A.D., 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 Psychol. 119, 217–230.

Censor, N., Karni, A., Sagi, D., 2006. A link between perceptual learning, adaptation and sleep. Vision Res. 46, 4071–4074.

Chein, J.M., Morrison, A.B., 2010. Expanding the mind's workspace: training and transfer effects with a complex working memory span task. Psychon. Bull. Rev. 17, 193–199.

Chisholm, J.D., Hickey, C., Theeuwes, J., Kingstone, A., 2010. Reduced attentional capture in action video game players. Atten. Percept. Psychophys. 72, 667–671.

Coch, D., Sanders, L.D., Neville, H.J., 2005. An event-related potential study of selective auditory attention in children and adults. J. Cogn. Neurosci. 17, 605–622.

Cohen, Y., Daikhin, L., Ahissar, M., 2013. Perceptual learning is specific to the trained structure of information. J. Cogn. Neurosci. 25 (12), 2047–2060. http://dx.doi.org/10.1162/jocn_a_00453.

Colom, R., Quiroga, M.Á., Shih, P.C., Martínez, K., Burgaleta, M., Martínez-Molina, A., Román, F.J., Requena, L., Ramírez, I., 2010. Improvement in working memory is not related to increased intelligence scores. Intelligence 38, 497–505.

Costa-Giomi, E., 1999. The effects of three years of piano instruction on children's cognitive development. J. Res. Music Educ. 47, 198–212.

Dahlin, K.I., 2011. Effects of working memory training on reading in children with special needs. Read. Writ. 24, 479–491.

Dosher, B.A., Lu, Z.-L., 1998. Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. Proc. Natl. Acad. Sci. U. S. A. 95, 13988–13993.

Dosher, B.A., Lu, Z.-L., 1999. Mechanisms of perceptual learning. Vision Res. 39, 3197–3221.

Dosher, B.A., Lu, Z.-L., 2007. The functional form of performance improvements in perceptual learning: learning rates and transfer. Psychol. Sci. 18, 531–539.

Dosher, B.A., Jeter, P., Liu, J., Lu, Z.-L., 2013. An integrated reweighting theory of perceptual learning. Proc. Natl. Acad. Sci. U. S. A. 110, 13678–13683.

Duckworth, A.L., Quinn, P.D., Lynam, D.R., Loeber, R., Stouthamer-Loeber, M., 2011. Role of test motivation in intelligence testing. Proc. Natl. Acad. Sci. U. S. A. 108, 7716–7720.

Dye, M., Green, C., Bavelier, D., 2009a. The development of attention skills in action video game players. Neuropsychologia 47, 1780–1789.

Dye, M.W., Green, C.S., Bavelier, D., 2009b. Increasing speed of processing with action video games. Curr. Dir. Psychol. Sci. 18, 321–326.

Elger, C.E., Grunwald, T., Lehnertz, K., Kutas, M., Helmstaedter, C., Brockhaus, A., Van Roost, D., Heinze, H.J., 1997. Human temporal lobe potentials in verbal learning and memory processes. Neuropsychologia 35, 657–667.

Flohr, J.W., 1981. Short-term music instruction and young children's developmental music aptitude. J. Res. Music Educ. 29, 219–223.

Forgeard, M., Winner, E., Norton, A., Schlaug, G., 2008. Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. PLoS One 3, e3566.

Franceschini, S., Gori, S., Ruffino, M., Viola, S., Molteni, M., Facoetti, A., 2013. Action video games make dyslexic children read better. Curr. Biol. 23 (6), 462–466. http://dx.doi.org/10.1016/j.cub.2013.01.044.

Francois, C., Schön, D., 2011. Musical expertise boosts implicit learning of both musical and linguistic structures. Cereb. Cortex 21, 2357–2365.

François, C., Chobert, J., Besson, M., Schön, D., 2012. Music training for the development of speech segmentation. Cereb. Cortex 23 (9), 2038–2043.

Gibson, B.S., Gondoli, D.M., Johnson, A.C., Steeger, C.M., Dobrzenski, B.A., Morrissey, R.A., 2011. Component analysis of verbal versus spatial working memory training in adolescents with ADHD: a randomized, controlled trial. Child Neuropsychol. 17, 546–563.

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. J. Exp. Psychol. Hum. Percept. Perform. 32, 1465.

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. Psychol. Sci. 18, 88–94.

Green, C., Bavelier, D., 2008. Exercising your brain: a review of human brain plasticity and training-induced learning. Psychol. Aging 23, 692.

Green, C., Bavelier, D., 2012. Learning, attentional control, and action video games. Curr. Biol. 22, R197–R206.

Green, C.S., Li, R., Bavelier, D., 2010. Perceptual learning during action video game playing. Top. Cogn. Sci. 2, 202–216.

Hai, A., Granot, R.Y., Ahissar, M., The enigma of dyslexic musicians. Neuropsychologia, Accepted.

Hochstein, S., Ahissar, M., 2002. View from the top: hierarchies and reverse hierarchies in the visual system. Neuron 36, 791–804.

Holmes, J., Gathercole, S.E., Dunning, D.L., 2009. Adaptive training leads to sustained enhancement of poor working memory in children. Dev. Sci. 12, F9–F15.

Holmes, J., Gathercole, S.E., Place, M., Dunning, D.L., Hilton, K.A., Elliott, J.G., 2010. Working memory deficits can be overcome: impacts of training and medication on working memory in children with ADHD. Appl. Cogn. Psychol. 24, 827–836.

Hook, P.E., Macaruso, P., Jones, S., 2001. Efficacy of Fast ForWord training on facilitating acquisition of reading skills by children with reading difficulties—a longitudinal study. Ann. Dyslexia 51, 73–96.

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. Atten., Percept. Psychophys. 73, 2399–2412.

Hubert-Wallander, B., Green, C.S., Bavelier, D., 2011. Stretching the limits of visual attention: the case of action video games. Wiley Interdiscip. Rev. Cogn. Sci. 2, 222–230.

Hurwitz, I., Wolff, P.H., Bortnick, B.D., Kokas, K., 1975. Nonmusical effects of the kodaly music curriculum in primary grade children. J. Learn. Disabil. 8, 167–174.

Hyde, K.L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A.C., Schlaug, G., 2009. Musical training shapes structural brain development. J. Neurosci. 29, 3019–3025.

Isen, A., 2008. Some ways in which positive affect influences decision making and problem solving. In: Lewis, M., Haviland-Jones, J., Barrett, L. (Eds.), Handbook of Emotions, third ed. Guilford, New York, NY, pp. 548–573.

Jacoby, N., Ahissar, M., "Transfer of training" a misinterpretation of correlation between training gains, submitted.

Jaeggi, S.M., Buschkuehl, M., Jonides, J., Perrig, W.J., 2008. Improving fluid intelligence with training on working memory. Proc. Natl. Acad. Sci. U. S. A. 105, 6829–6833.

Jaeggi, S.M., Buschkuehl, M., Jonides, J., Shah, P., 2011. Short-and long-term benefits of cognitive training. Proc. Natl. Acad. Sci. U. S. A. 108, 10081–10086.

Karni, A., Sagi, D., 1991. Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. Proc. Natl. Acad. Sci. U. S. A. 88, 4966–4970.

Karni, A., Tanne, D., Rubenstein, B.S., Askenasy, J., Sagi, D., 1994. Dependence on REM sleep of overnight improvement of a perceptual skill. Science 265, 679–682.

Klingberg, T., Forssberg, H., Westerberg, H., 2002. Training of working memory in children with ADHD. J. Clin. Exp. Neuropsychol. 24, 781–791.

Klingberg, T., Fernell, E., Olesen, P.J., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C.G., Forssberg, H., Westerberg, H., 2005. Computerized training of working memory in children with ADHD-A randomized, controlled trial. J. Am. Acad. Child Adolesc. Psychiatry 44, 177–186.

Knapp, T.R., Schafer, W.D., 2009. From Gain Score t to ANCOVA F (and vice versa). Pract. Assess. Res. Eval. 14, 2.

Kolers, P.A., 1976. Reading a year later. J. Exp. Psychol. Hum. Learn. Mem. 2, 554.

Korman, M., Doyon, J., Doljansky, J., Carrier, J., Dagan, Y., Karni, A., 2007. Daytime sleep condenses the time course of motor memory consolidation. Nat. Neurosci. 10, 1206–1213.

Kraus, N., Chandrasekaran, B., 2010. Music training for the development of auditory skills. Nat. Rev. Neurosci. 11, 599–605.

Kronenberger, W.G., Pisoni, D.B., Henning, S.C., Colson, B.G., Hazzard, L.M., 2011. Working memory training for children with cochlear implants: a pilot study. J. Speech Lang. Hear. Res. 54, 1182.

Langer, E., Djikic, M., Pirson, M., Madenci, A., Donohue, R., 2010. Believing is seeing using mindlessness (mindfully) to improve visual acuity. Psychol. Sci. 21, 661–666.

Lee, H., Boot, W.R., Basak, C., Voss, M.W., Prakash, R.S., Neider, M., Erickson, K.I., Simons, D.J., Fabiani, M., Gratton, G., 2012. Performance gains from directed training do not transfer to untrained tasks. Acta Psychol. 139, 146–158.

Lemay, S., Bédard, M.-A., Rouleau, I., Tremblay, P.-L., 2004. Practice effect and test-retest reliability of attentional and executive tests in middle-aged to elderly subjects. Clin. Neuropsychol. 18, 284–302.

Li, R., Polat, U., Makous, W., Bavelier, D., 2009. Enhancing the contrast sensitivity function through action video game training. Nat. Neurosci. 12, 549–551.

Li, R., Polat, U., Scalzo, F., Bavelier, D., 2010. Reducing backward masking through action game training. J. Vis. 10 (33), 1–13. http://dx.doi.org/10.1167/10.7.1239.

Li, R.W., Ngo, C., Nguyen, J., Levi, D.M., 2011. Video-game play induces plasticity in the visual system of adults with amblyopia. PLoS Biol. 9, e1001135.

Lilienthal, L., Tamez, E., Shelton, J.T., Myerson, J., Hale, S., 2013. Dual n-back training increases the capacity of the focus of attention. Psychon. Bull. Rev. 20 (1), 135–141.

Lord, F.M., 1967. A paradox in the interpretation of group comparisons. Psychol. Bull. 68, 304–305.

Lu, Z.-L., Liu, C.Q., Dosher, B.A., 2000. Attention mechanisms for multi-location first-and second-order motion perception. Vision Res. 40, 173–186.

Lynn, R., Graham Wilson, R., Gault, A., 1989. Simple musical tests as measures of Spearman's g. Personal. Individ. Differ. 10, 25–28.

Matarazzo, J.D., Carmody, T.P., Jacobs, L.D., 1980. Test-retest reliability and stability of the WAIS: a literature review with implications for clinical practice. J. Clin. Exp. Neuropsychol. 2, 89–105.

May, C.P., Hasher, L., 1998. Synchrony effects in inhibitory control over thought and action. J. Exp. Psychol. Hum. Percept. Perform. 24, 363.

McArthur, G., 2007. Test–retest effects in treatment studies of reading disability: the devil is in the detail. Dyslexia 13, 240–252.

Mcnab, F., Klingberg, T., 2007. Prefrontal cortex and basal ganglia control access to working memory. Nat. Neurosci. 11, 103–107.

Mezzacappa, E., Buckner, J.C., 2010. Working memory training for children with attention problems or hyperactivity: a school-based pilot study. Sch. Ment. Heal. 2, 202–208.

Mishra, J., Zinni, M., Bavelier, D., Hillyard, S.A., 2011. Neural basis of superior performance of action videogame players in an attention-demanding task. J. Neurosci. 31, 992–998.

Moreno, S., Marques, C., Santos, A., Santos, M., Besson, M., 2009. Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. Cereb. Cortex 19, 712–723.

Morrongiello, B.A., Roes, C.L., 1990. Developmental changes in children's perception of musical sequences: effects of musical training. Dev. Psychol. 26, 814.

Neville, H.J., Stevens, C., Pakulak, E., Bell, T.A., Fanning, J., Klein, S., Isbell, E., 2013. Family-based training program improves brain function, cognition, and behavior in lower socioeconomic status preschoolers. Proc. Natl. Acad. Sci. U. S. A. 110, 12138–12143.

Oganian, Y., Ahissar, M., 2012. Poor anchoring limits dyslexics' perceptual, memory, and reading skills. Neuropsychologia 50, 1895–1905.

Olesen, P.J., Westerberg, H., Klingberg, T., 2003. Increased prefrontal and parietal activity after training of working memory. Nat. Neurosci. 7, 75–79.

Owen, A.M., Hampshire, A., Grahn, J.A., Stenton, R., Dajani, S., Burns, A.S., Howard, R.J., Ballard, C.G., 2010. Putting brain training to the test. Nature 465, 775–778.

Polat, U., Sagi, D., 1994. The architecture of perceptual spatial interactions. Vision Res. 34, 73–78.

Polat, U., Ma-Naim, T., Belkin, M., Sagi, D., 2004. Improving vision in adult amblyopia by perceptual learning. Proc. Natl. Acad. Sci. U. S. A. 101, 6692–6697.

Redick, T.S., Shipstead, Z., Harrison, T.L., Hicks, K.L., Fried, D.E., Hambrick, D.Z., Kane, M.J., Engle, R.W., 2013. No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. J. Exp. Psychol. Gen. 142, 359.

Richmond, L.L., Morrison, A.B., Chein, J.M., Olson, I.R., 2011. Working memory training and transfer in older adults. Psychol. Aging 26, 813.

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. Int. J. Music Educ. 30, 57–78.

Sanders, L.D., Stevens, C., Coch, D., Neville, H.J., 2006. Selective auditory attention in 3-to 5-year-old children: an event-related potential study. Neuropsychologia 44, 2126–2138.

Schellenberg, E.G., 2004. Music lessons enhance IQ. Psychol. Sci. 15, 511–514.

Schellenberg, E.G., 2006. Long-term positive associations between music lessons and IQ. J. Educ. Psychol. 98, 457.

Schmiedek, F., Lövdén, M., Lindenberger, U., 2010. Hundred days of cognitive training enhance broad cognitive abilities in adulthood: findings from the COGITO study. Front. Aging Neurosci. 2.

Seitz, A., Watanabe, T., 2005. A unified model for perceptual learning. Trends Cogn. Sci. 9, 329–334.

Shavelson, R.J., Yuan, K., Alonzo, A., 2008. On the impact of computer training on working memory and fluid intelligence. Fostering Change in Institutions, Environments, and People: A Festschrift in Honor of Gavriel Salomon. Routledge, New York, 35-48.

Shipstead, Z., Redick, T.S., Engle, R.W., 2012. Is working memory training effective? Psychol. Bull. 138, 628.

Standley, J.M., 2008. Does music instruction help children learn to read? Evidence of a meta-analysis. Update Appl. Res. Music Educ. 27, 17–32.

Stevens, C., Fanning, J., Coch, D., Sanders, L., Neville, H., 2008. Neural mechanisms of selective auditory attention are enhanced by computerized training: electrophysiological evidence from language-impaired and typically developing children. Brain Res. 1205, 55–69.

Stevens, C., Lauinger, B., Neville, H., 2009. Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: an event-related brain potential study. Dev. Sci. 12, 634–646.

Stevens, C., Harn, B., Chard, D.J., Currin, J., Parisi, D., Neville, H., 2013. Examining the role of attention and instruction in at-risk kindergarteners electrophysiological measures of selective auditory attention before and after an early literacy intervention. J. Learn. Disabil. 46, 73–86.

Stickgold, R., 2005. Sleep-dependent memory consolidation. Nature 437, 1272–1278.

Stickgold, R., James, L., Hobson, J.A., 2000. Visual discrimination learning requires sleep after training. Nat. Neurosci. 3, 1237–1238.

Thompson, W.F., Schellenberg, E.G., Husain, G., 2004. Decoding speech prosody: do music lessons help? Emotion 4, 46.

Thorell, L.B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., Klingberg, T., 2009. Training and transfer effects of executive functions in preschool children. Dev. Sci. 12, 106–113.

Tsushima, Y., Seitz, A.R., Watanabe, T., 2008. Task-irrelevant learning occurs only when the irrelevant feature is weak. Curr. Biol. 18, R516–R517.

Van Der Molen, M., Van Luit, J., Van Der Molen, M., Klugkist, I., Jongmans, M., 2010. Effectiveness of a computerised working memory training in adolescents with mild to borderline intellectual disabilities. J. Intellect. Disabil. Res. 54, 433–447.

Watanabe, T., Náñez, J.E., Koyama, S., Mukai, I., Liederman, J., Sasaki, Y., 2002. Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. Nat. Neurosci. 5, 1003–1009.

Windle, C., 1954. Test-retest effect on personality questionnaires. Educ. Psychol. Meas. 14, 617–633.

Yang, H., Yang, S., Isen, A.M., 2013. Positive affect improves working memory: implications for controlled cognitive processing. Cogn. Emotion 27, 474–482.

Zhao, X., Wang, Y., Liu, D., Zhou, R., 2011. Effect of updating training on fluid intelligence in children. Chin. Sci. Bull. 56, 2202–2205.