

Training Cognition in ADHD: Current Findings, Borrowed Concepts, and Future Directions

Kyle J. Rutledge · Wouter van den Bos ·
Samuel M. McClure · Julie B. Schweitzer

© The American Society for Experimental NeuroTherapeutics, Inc. 2012

Abstract With both its high prevalence and myriad of negative outcomes, Attention-Deficit/Hyperactivity Disorder (ADHD) demands a careful consideration of the efficacy of its treatment options. Although the benefits of medication have a robust empirical background, nonpharmaceutical interventions evoke particular interest, as they are often viewed more favorably by parents. This review pays special attention to the use of working memory and recent cognitive training attempts in ADHD, describing its cognitive, behavioral, and biological effects in relation to current neurological theory of the disorder. While these treatments have demonstrated positive effects on some measures, there are limitations, as studies have failed to demonstrate generalization to critical measures, such as teacher-rated classroom behaviors, and have provided limited but growing evidence of functionally significant improvements in behavior. There is also a clear lack of research on the effects of training on reward systems and self-control. These limitations may be addressed by broadening the scope and procedures of the training and incorporating research concepts from other fields of study. First, it is important to consider the developmental trajectories of brain regions in individuals with the disorder, as they may relate to the effectiveness of cognitive training. Notions from behavioral economics, including delay discounting and framing (i.e., context) manipulations that

influence present orientation, also have applications in the study of cognitive training in ADHD. In considering these other domains, we may find new ways to conceptualize and enhance cognitive training in ADHD and, in turn, address current limitations of interventions that fall in this category.

Keywords Attention deficit hyperactivity disorder · Treatment · Working memory training · Nonpharmacological · Delay discounting

Need for Improved Treatments in ADHD

Attention-deficit/hyperactivity disorder (ADHD) is a critical public health concern, given that it is the most common childhood behavioral disorder, with prevalence rates of at least 5 % [1, 2]. ADHD is associated with negative outcomes in academic achievement, including lower reading and math achievement scores [3–7], lower high school completion rates [8] and an alarmingly high use of special education services [9]. Significantly elevated rates of antisocial behavior, increased arrests, and greater substance abuse exist in adolescents and adults with ADHD [9–11]. This constellation of negative outcomes underscores the need for additional ADHD treatments.

There continues to be a great need to identify more effective and adjunctive interventions for ADHD. This is particularly true, given that current treatments do not appear to have long-term benefits once the intervention has been stopped. Some of the most compelling results on this issue come from the Multimodal Treatment Study of ADHD [12, 13], the most comprehensive treatment study of ADHD to date. The study reported initial positive effects of medication and/or behavioral treatments relative to a community treatment during a 14-month intervention trial. However, when these same participants were evaluated 8 years after the active treatment period concluded, the follow-up

K. J. Rutledge
Department of Human and Community Development,
University of California, Davis,
Davis, CA 95616, USA

W. van den Bos · S. M. McClure
Department of Psychology, Stanford University,
Stanford, CA 94305, USA

K. J. Rutledge · J. B. Schweitzer (✉)
Department of Psychiatry and Behavioral Sciences and MIND
Institute, University of California Davis School of Medicine,
Sacramento, CA 95817, USA
e-mail: julie.schweitzer@ucdmc.ucdavis.edu

measures showed that the treatment effects were not sustained. Still, the results were informative in that the investigators found that there were no differences in effectiveness on ADHD functioning 8 years later between intervention modalities (e.g., medication, behavioral, combined, or community comparison) [13]. Study participants had surprisingly negative outcomes, regardless of intervention modality, with ADHD participants functioning significantly worse on 91 % of outcomes relative to their non-ADHD peers [13].

The need to explore novel treatments for ADHD (for more detail see Epstein and Tsal [14]) is further supported by concerns that arise from current treatments for some persons with ADHD (e.g., side effects, diminished effectiveness). Nonpharmacological treatments are of particular interest to parents of children and adolescents with ADHD due to parental concerns regarding the use of psychoactive medications on brain and body development in children. Parents of children with ADHD prefer behavioral interventions alone in comparison to either medication alone [15] or medication combined with a behavioral treatment [16]. Contrary to public opinion, the data suggest that ADHD is often undertreated [17, 18]. This results, in part, from the fact that concerns regarding medication use frequently stop parents from pursuing treatment. Training caregivers in behavioral procedures for ADHD is well supported [19–21]; however, often the benefits derived from these interventions do not generalize beyond the clinical setting.

We focus here on cognitive training approaches that are appealing precisely because of the promise they may hold for generalization of effects beyond the training setting. This article reviews interventions or strategies devised to target specific cognitive domains; these interventions have the primary goal of improving functioning subsumed under the term cognitive control or executive functioning, but they also have concurrent implications for behavioral impairments associated with ADHD. The hypothesis is that by improving these general, underlying functions, the benefits will generalize to performance regulated by cognitive control and executive functioning. For example, key impairments in ADHD, such as poor attention, working memory, and impulsivity may all be regulated under executive functioning and have the potential to improve as broader cognitive functioning strengthens.

The training typically involves repeated practice on exercises that become increasingly more challenging as the performance improves. Often the training is administered via a computer, but it may also include noncomputerized methods. We recognize that an early use of the term of “cognitive training” referred to the application of self-monitoring and self-reinforcement techniques to improve functioning in ADHD [22]. This article will not review the literature associated with the previous use of the term

“cognitive training” for ADHD. This article also will not review neurofeedback training techniques (see Loo and Makeig [23] and Moriyama et al. [24]).

Support for the consideration of improved generalization of cognitive training effects, beyond the training sessions, derives from evidence suggesting that the training may affect brain regions [25, 26] and neurotransmitter systems [27] directly, thereby theoretically altering the cognitive functions that derive from the function of these neural systems. This intervention style stands in contrast to others that target specific behaviors, therefore making them even more susceptible to low generalizability due to their inherently narrow foci. Furthermore, it has been suggested that individuals with ADHD may be ill-equipped in adopting (across multiple settings) those strategies and improvements garnered within the specific treatment context [28]. Cognitive interventions, however, target processes that are putatively expected to automatically govern behaviors across multiple situations, making this particular type of intervention a hypothetically broad-reaching treatment.

Despite our focus on interventions that impact the prefrontal cortex [25, 26] and dopamine system [27] due to their strong association with ADHD symptomatology, we note that it is important to consider how these systems are couched within broader neural networks. Thus, we will begin with a discussion of the interrelationship between multiple neural systems in the manifestation of ADHD symptoms so that we may demonstrate the utility in considering multiple neural developmental factors when striving to understand treatment effectiveness in the disorder. We strongly recommend that future studies directly test if there is a relationship between changes in neural functioning (i.e., neurotransmitter, brain regions, and connectivity) and generalizability of training effects. A better understanding of neural functioning, training components, and transfer of training effects is crucial to developing more effective treatments for psychiatric and learning disorders in general.

Dual Systems Theory

ADHD Dual-Process Theory

Nearly a decade ago, Sonuga-Barke [29] presented the dual-processing theory of ADHD, positing that the symptom set of individuals with combined type ADHD could be explained by dysfunction in either of two underlying neuropsychological processes: behavioral inhibition or reward dysfunction. In this model, deficits in inhibition and consequent dysregulation of nigro-striatal-prefrontal pathways [30] underlie dysfunctions in attention and

other cognitive processes, including executive functioning [31]. These deficits then affect behavior through limiting state regulation [32], leading to hyperactivity and impulsivity. Dysregulation of the meso-limbic-striatal pathways [30] is associated with reward dysfunction, including problems with self-control or delay aversion, in Sonuga-Barke's conceptualization. More recent versions of the model have evolved, now using slightly different language and emphasizing the contributions of cognitive control ("cool processes") *versus* reward responsivity ("hot processes") to explain the individual variation found in ADHD [30, 33]. We propose that ADHD symptoms are consistent with more narrowly defined deficits arising from: 1) deficits in cognitive control [34] related to improper function of the dorsolateral prefrontal cortex (DLPFC) and/or 2) an overdependence on immediate rewards indicative of mesolimbic dopamine dysfunction [35–38]. Cognitive control is the ability to adapt and regulate behavior to current demands by attending to task-relevant information over distracting, irrelevant stimuli that may interfere with meeting a goal [39]. Individuals with ADHD may experience an equal degree of dysfunction in the cognitive control or reward system, or an imbalance between the two, with greater dysfunction in one of the two systems. By combining these two dissociable mechanisms of cognitive control and reward response into a single theory, the resulting multiple developmental pathway model accounts for the heterogeneity present in the disorder and applies to a greater majority of cases of ADHD. Although our techniques to measure these various behavioral and neural impairments currently lack standardization, this theoretical approach provides guidance in that it implies treatment may be dictated by the type and number of impairments present in the pathway, depending on their expression in the individual. An advantage of many of the interventions we review is that the underlying brain functioning can be quantitatively measured, the treatment strategies are "dose" quantified, and the degree to which the changes in brain functioning and the dose of treatment relate to the degree of generalization as determined by objective measures (e.g., Permanent Product Measure of Performance, a math-based assessment of productivity [40]) and subjective measures (e.g., parent and teacher rating scales).

Dual Systems Theory of Self-Control

A strikingly similar dual systems theory of self-control has emerged in the behavioral economics literature [41]. A major advantage of the overlap between the approaches of these two dual systems theories is that the methodology and research findings from each area can inform one another. Furthermore, findings from the dual systems literature may shed light on how to develop targeted treatments to address

impulsivity in ADHD, as behavioral economists are interested in identifying factors that modify decision-making that may lead to impulsive behavior. Supporting the dual-process theory of ADHD, McClure et al. [42] demonstrated that self-control in healthy adults is determined by two brain systems: 1) a reward system that responds preferentially to immediately available rewards, which includes a number of brain regions targeted by the midbrain dopamine system including the nucleus accumbens, posterior cingulate cortex, and ventromedial prefrontal cortex (for more detail see McClure et al. [43] and O'Doherty et al. [44]); and 2) a cognitive control system composed of the DLPFC and posterior parietal cortex that evaluates choices between waiting for rewards at all delays, whether they be sooner or later in time (see Fig. 1A). Further, the relative activity and interaction between the cognitive control and reward-response systems determine choice in self-control situations [42, 45] (Fig. 1B). Hare et al. [46] elaborated the McClure model by demonstrating that the DLPFC control region modulates the strength of reward regions (i.e., ventromedial prefrontal cortex) by suppressing activity via a two-step pathway that includes the inferior frontal gyrus. A critical barrier to understanding how these systems govern behavior is the absence of studies that directly test how the reward and cognitive control neural systems interact. Nonetheless, we believe (as discussed hereafter in this article) that insights from the behavioral economics literature are ripe for translation to ADHD interventions. Both the ADHD and behavioral economics models suggest that treatments should be tailored for the individual, depending on whether both or one system is impaired and leads to cognitive and behavioral dysfunctions.

Cognitive and Working Memory Training in ADHD

We will begin the review of interventions by discussing evidence for treatments that aim to alter the cognitive control system and are putatively related to DLPFC and posterior parietal cortex functioning. Cognitive training and working memory training, in particular, are nonpharmaceutical options rapidly growing in popularity for ADHD that affect processes related to the cognitive control aspects of the dual hypothesis model. Cognitive training techniques are designed to target different aspects of cognition, such as attention (e.g. [47]) or working memory (e.g. [48]). Working memory training has shown the greatest proliferation of training techniques due to the relationship between working memory capacity and general cognitive functioning [49]. We will review a variety of cognitive interventions that have been tested or have potential for improving functioning of people with ADHD and discuss the neural changes associated with training later in the article.

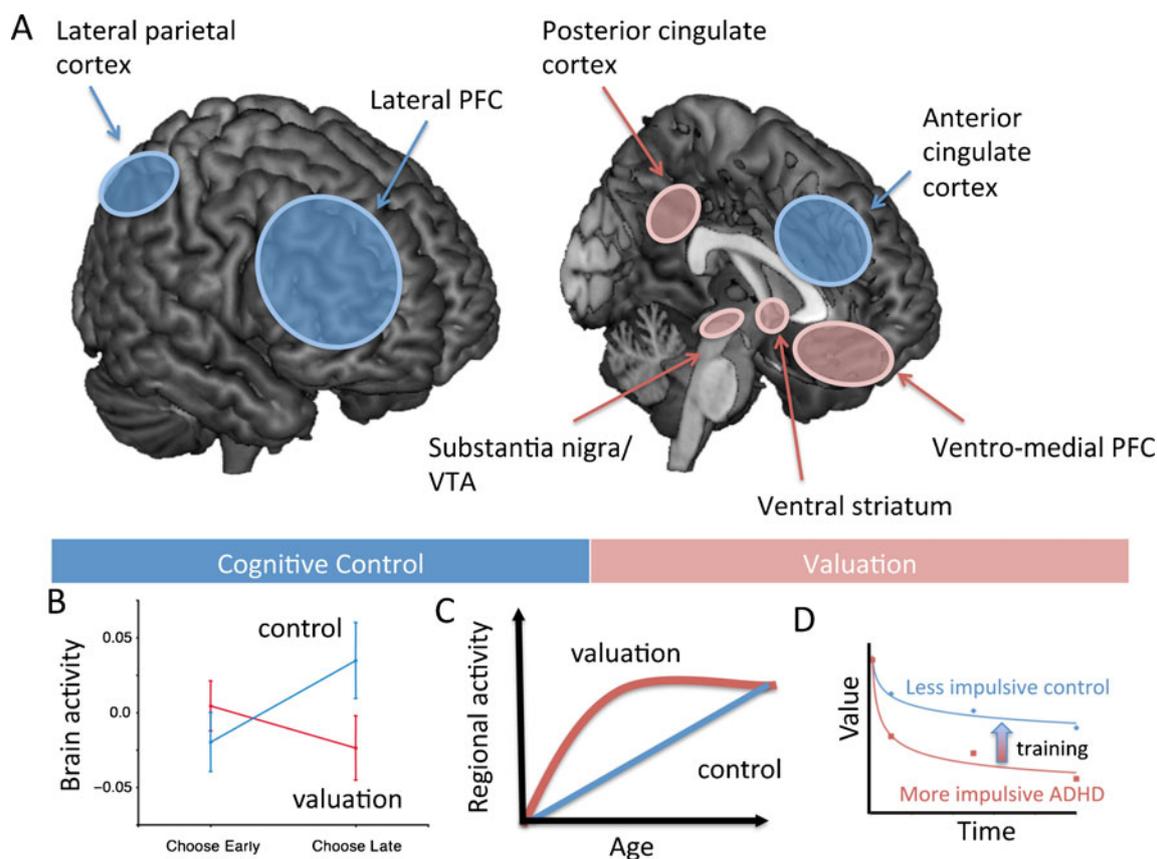


Fig. 1 (A) Visualization of the valuation network (red) and the control network (blue). (B) Results from McClure et al. [42] show that impulsive choices were associated with relatively more activity in the valuation network, whereas future oriented choices were associated with more activity in the control network. (C) Neurobiological model of cognitive and motivational processes showing linear development of

top down prefrontal control network relative to a \cap -shaped function for the development of valuation network (based on Somerville and Casey [84]). (D) Plot showing two discount curves based on four indifference points of a relatively impulsive person with ADHD and a more patient control participant; the arrow indicates the proposed effect of training on the discount curve

Attention Training

The idea of directly training cognitive deficits via a “process-specific approach,” such as a program designed to improve attention or working memory, is borrowed from rehabilitative techniques for individuals suffering from a brain injury [50]. In the first published study to effectively use such a training program for children with formal diagnoses of ADHD, Kerns et al. [47] targeted attention and presumed that improvements would generalize to other domains beyond attention due to the broader function of the neural systems altered by the intervention. The researchers assigned participants to a treatment group ($n=7$) or comparison group ($n=7$), matching them by age, sex, and medication status. The comparison group received an academically oriented, computer-based placebo intervention. Treatment sessions occurred twice a week for 8 weeks, each lasting 30 minutes. During these sessions, participants completed modules taxing sustained, selective, alternating, and divided attention. After treatment, participants exhibited

attentional improvements, as well as improvements in non-trained domains, including academic efficiency. In spite of its limited sample size, which is understandable, given its pioneering nature, this study is commendable for measuring the treatment outcomes via predominantly objective means. Although parent and teacher ratings were included to assess ADHD symptom manifestation in the home and at school, the assessments of visual-spatial abilities, academic efficiency, impulsivity, executive functioning, and attention were completed through directly observing each child’s performance on the respective tasks. Impulsivity was measured with the Matching Familiar Figures Task, which is more specifically a metric of reflection impulsivity, or the tendency to make a decision based on minimal information. This way of conceptualizing impulsivity is quite different from another, operationalized as delay discounting, as we will discuss later; this is useful to note, as it exemplifies how methodological differences between studies may limit the generalizability of the findings, and perhaps even our understanding of the generalizability of treatment effects.

Recent work using the same treatment design replicated these results in addition to recording improvements in inattentive symptoms and executive functioning [51], hyperactive impulsive symptoms, working memory, and fluid reasoning [52] as well as reading comprehension [53]. As noted by Posner and Rothbart [54], the promising results stemming from attentional training research provide encouragement to continue studying the effects of this intervention, both behaviorally and neurologically.

A recent randomized, placebo-controlled study sought to examine the differences in outcome of first grade children with inattention problems completing two different forms of attention training: computerized attention training (CAT) and computer-assisted instruction (CAI) [55]. Children were selected solely based on teacher ratings of inattentive symptoms. Those scoring 1 standard deviation above the sample mean were asked to participate in the training; therefore, children did not need an ADHD diagnosis to participate. CAT incorporated exercises meant to improve sustained attention in visual and auditory tasks, whereas CAI targeted reading and math skills by presenting small pieces of information followed by repetition. Although CAI was described as an “instructional” treatment, it followed a course consistent with cognitive training programs, but differed in that it targeted academic skills rather than a specific cognitive function. Both interventions occurred during school hours 2 days a week for 14 weeks, taking more than an hour per session. In the analyses, treatment groups (CAT: $n=25$; CAI: $n=27$) were compared to each other, as well as a wait list control group ($n=25$). The results demonstrated that both interventions successfully decreased inattentive symptoms, such that 50 % improved significantly and 25 % improved enough for their teacher-rated scores of inattention to fall within the normal range. Children in the CAI intervention also displayed improvements in academic functioning. It appears the instruction training yielded a stronger effect than the attention training did, suggesting that targeting higher order processes may lead to even greater effects. Despite improvements in inattention, regardless of training type, approximately 75 % of participants receiving the treatment continued to exhibit problems with attention after the treatments, such that teacher ratings of attention difficulties still fell at least 1 standard deviation above the normative mean. Furthermore, these effects on attention were not sustained into the following school year, in which all participants, including controls, demonstrated declines in inattention by the second grade.

Working Memory Training

Improvements in working memory functioning for persons with ADHD are particularly relevant, as they may address not only the well-documented behavioral and neural

working memory impairments found in ADHD (e.g., for more detail see Fassbender et al. [56], Schoechlin and Engel [57], Martinussen et al. [58], Willcutt et al. [59], Mills et al. [60], Rapport et al. [61], and Bolden et al. [62]), but also some attentional control issues, as working memory capacity is, at least in part, related to degree of distractibility and cognitive control [63, 64]. Thus, improvements in working memory, or conversely attentional control, may benefit the alternative, related process. A computerized cognitive training option that has the largest evidence base for children with ADHD was developed by Klingberg et al. [65]. The authors based the working memory training protocol on a program designed to improve speech processing in children with language learning impairments [66]. The language intervention involved a daily set of therapy trials presented as computer games that gradually increased in difficulty during the course of 4 weeks. The treatment yielded positive outcomes for the children with respect to their abilities to process speech and language. Klingberg et al. [65] also used a schedule of therapy trials occurring a few times a week for 5 to 6 weeks until approximately 25 sessions were reached. However, the Klingberg et al. [65] program used visuospatial and spatial-verbal working memory (span) tasks. The difficulty level for each task was continuously calibrated based on the individual’s performance with span lengths increasing as performance improved, and difficulty easing when errors occurred. This initial study included a randomized, placebo-controlled design with participants in the control condition performing the same number of sessions at a fixed difficulty level, although the sessions were of shorter duration. Measures only pertained to neuropsychological assessments, including working memory tasks, Raven’s Progressive Matrices, a reaction time task, and a Stroop task. Improvements were seen after the intervention, relative to placebo, in Raven’s Progressive Matrices and Stroop performance, but not reaction time. The results also showed a significant improvement in performance on both trained and nontrained visual-spatial working memory tasks, and a decrease in frequency of head movements measured via an infrared motion detector during a computer-based continuous performance task. This initial study was limited to a small sample size ($n=7$ per group), no follow-up measures, and a lack of specific assessments of change in ADHD symptoms.

A later study using the same randomized, placebo-controlled working memory training in children with ADHD in a larger sample ($n=53$, 44 of whom successfully completed training) demonstrated that the intervention could lead to improvements beyond working memory, including increased executive task performance on the Stroop task and Raven’s Progressive Matrices [48]. Furthermore, inattention and hyperactive/impulsive symptoms, as measured by parent ratings showed significant improvement; such benefits

were not evident, however, from teacher ratings or frequency in the number of head movements. Also, the improvements detected following training persisted for 3 months after the cognitive treatment had ceased. The authors hypothesized that everyday activities allow for more practice of working memory functions, and with the improvements driven by the treatment, activities with high working memory loads become more doable for these individuals, allowing for more practice, and creating a positive feedback loop for these individuals. Thus, treatment effects continue to increase after the active treatment has concluded. Further research has demonstrated effects lasting at least 6 months [67] in children with low working memory performance who were not evaluated for ADHD. The strength of the follow-up effects, however, in the Holmes et al. [67] study must be considered with caution as they did not re-assess the control group at 6 months post-training, but only the active training group; thus, one cannot rule out the effect of maturation on maintaining performance improvements.

A more recent study compared the effect of the Cogmed Working Memory Training program on 25 children diagnosed with ADHD to the effects of stimulant medication on short-term and working memory [68]. The authors of this study did not independently evaluate and confirm the ADHD diagnosis of the participants who were referred to as previously diagnosed and treated persons with ADHD. Furthermore, the dose of stimulant medication given to the participants was one that was prescribed by their treatment provider, but not empirically validated as the optimal treatment dose. Thus, it is difficult to compare the effects of the stimulant medication *versus* the working memory training on cognitive functioning in ADHD. The authors did use objective measures, such as the Automated Working Memory Assessment (AWMA) [69, 70] to assess verbal and visuospatial recall and verbal and visuospatial working memory processing, and the full (i.e., 4 subtest) Wechsler Abbreviated Scales of Intelligence (WASI) [71]. Participants in session 1 were evaluated without medication; in session 2, on their prescribed dose and before working memory training; in session 3, medicated and after training session; and in session 4, medicated and 6 months after the training was completed. No information on the timing of the dose or verification of the medication use was included in the article. The results suggest that the cognitive training had a broader effect on improving working memory and short-term memory in comparison to medication effects. In an attempt to assess for the presence of re-test effects on the memory stimuli, the authors used four new memory subtests from the AWMA out of the eight they tested for session 3 *versus* session 2 comparisons. The pre- *versus* post-training scores on the novel memory tests showed significant improvement, suggesting that treatment gains were not due to

practice effects. Gains in performance persisted at the 6-month follow-up for visuospatial short-term memory, verbal working memory, and visuospatial working memory, but not verbal short-term memory. The WASI scores were unchanged after training during session 3. This study suggests a positive effect of the Cogmed working memory training on objective measures of working memory; however, the design is ultimately limited by not using a randomized, double-blind placebo procedure and experimenter titration of the medication dose.

A recent study of 52 children with ADHD in a wait list controlled trial found significant improvement on parent-rated measures of executive functioning, the Behavior Rating Inventory of Executive Function (BRIEF), immediately after training and at a 4-month follow-up [72]. However, there were no significant changes in teacher ratings in this study, nor were there any in a previous study that included teacher ratings [48]. The wait list control condition used in this study, however, does not control for expectancy effects. Controlling for expectancy effects in a treatment study is particularly concerning in a study in which the primary measure is subjective, such as in the rating scale measures used in this study. These studies demonstrated partial evidence for the generalization of the training on more global functioning on parent (but not teacher) ADHD ratings scales, improved response inhibition functioning, and complex reasoning [48, 65]. There is a need, however, for further use of randomized, placebo-controlled designs utilizing objective measures to assess near and far transfer effects of the working memory training on clinically relevant ADHD behavior.

Our laboratory [73] set out to assess the far transfer effects of working memory training in ADHD on an ecologically valid and objective measure of functioning using a randomized, placebo-controlled design. The study included 26 children ($n=12$ for active training; $n=14$ for placebo training) between the ages of 7- to 14-years-old with children with ADHD either randomized to the active condition in which they performed the Cogmed program for approximately 25 sessions or the “placebo condition” in which the working memory practice did not become increasingly more challenging with success. The Restricted Academic Situations Task (RAST) observational system was used to assess far transfer effects by measuring off-task behavior during a simulated academic exercise, which included the completion of an academic task (e.g., performing simple arithmetic written problems). Working memory training led to significant reductions in “off-task” ADHD-associated behavior on the RAST and improvement on working memory tests. No improvements were detected in the placebo group. Traditional measures of ADHD symptoms, such as the Conners’ Parent Rating Scale, however, did not demonstrate any significant group by time improvements.

This may be due to the small sample size. These preliminary data suggest working memory training may have far transfer effects to a clinically relevant problem (i.e., off-task behavior during an academic task) for children with ADHD.

Neural Effects of Cognitive Training

A precursor to the Klingberg et al. [48] study examined how working memory training might directly affect the neural systems underlying working memory [74] in typically developing adults. The study was the first to examine this effect in humans, building on a previous study that showed experience in a working memory task with degraded images increases both the ability of macaques to discriminate these images and the functionality of their prefrontal cortex neurons during the task [75]. Olesen et al. [74] provided working memory training to a handful of adults without ADHD or any other psychiatric diagnosis and acquired functional magnetic resonance imaging (fMRI) data while they completed a visuospatial working memory task. The first experimental group included a sample size of 3 participants who were scanned before and after the 5-week working memory training program. Compared to a nontreated control group (n=11), the small sample receiving the training showed significantly improved abilities in nontrained tasks, such as faster responses in a Stroop task and improved accuracy on a general measure of spatial intellectual functioning, the Raven's Advanced Progressive Matrices. Furthermore, the treated sample demonstrated greater activation in the middle frontal gyrus (right), the inferior parietal cortex (right), and the intraparietal cortex (bilaterally). The second treatment group contained a sample size of 8 participants who were provided the same working memory training but were scanned via fMRI 5 times across the 5 weeks of training. This second experimental group was also compared to a control group (n=11) that was behaviorally assessed at 5 different time points, verifying that the treatment can yield improved performance on cognitive tasks, even those that were not involved directly in training. Further, this second experiment provided more support for the neuronal plasticity of the systems underlying working memory in that, again, the prefrontal and parietal regions exhibited greater activity, as did thalamic and caudate regions, following working memory training. Methylphenidate also appears to increase thalamic activity during performance on a working memory task when tested in adults with ADHD [76]. Decreases in brain activation were noted in the right cingulate sulcus in the first and second working memory training experiments, with the authors hypothesizing this is due to a reduced need for motor planning. Although these results are promising, care must be taken in

interpreting their findings because the sample sizes included were small, and the nontreated control groups were only used for some behavioral comparisons, not in the imaging analysis.

A more recent study by Hoekzema et al. [25] found additional support for neural adaptation associated with cognitive training; however, this study directly assessed the effect of the training in children with ADHD (n=19, combined subtype). The authors did not use computerized cognitive training, but rather examined the effects on participants assigned to a treatment system using paper and pencil exercises completed under the supervision of a therapist. Sessions lasted for 45 minutes and occurred 5 times per week for 2 weeks. The program included activities designed to involve cognitive functions, such as working memory, attention, planning, and problem solving. The outcomes of this training session were compared to outcomes of age-, IQ-, gender-, symptom- and medication history-matched participants undergoing a social training program in which participants learned about social rules and standards and then engaged in role playing exercises to incorporate these rules into an active scenario. During the pre- and post-treatment fMRI scanning sessions, participants completed a visuospatial discrimination task as well as a go/no-go response inhibition task. The results indicated that the cognitive training program in ADHD led to increases in activity of the left orbitofrontal cortex, right middle temporal gyrus, and bilateral inferior frontal gyri during response inhibition; increased activity occurred in the right superior posterior cerebellum during the attention task. Cognitive training was also associated with decreased activity bilaterally in the precuneus and the right superior parietal cortex during the attention task. Individuals receiving social training did not demonstrate any increases in activity between pre- and post-treatment but did display decreased right superior posterior cerebellar activity.

The Hoekzema et al. [25] study showed similar results to those of Olesen et al. [74], including an increase in frontal brain activity, despite the fact that this study: 1) did not involve computer-based training, and 2) targeted multiple aspects of cognition, not just working memory. In contrast, Olesen et al. [74] demonstrated effects spanning parietal and subcortical regions, such as the striatum, which were not replicated in Hoekzema et al.'s [25] work. Likewise, only the Olesen et al. [74] study resulted in increases in parietal activation after training, whereas Hoekzema et al. [25] found decreases in both parietal and precuneus activation. Ignoring for a moment that differences between the studies may result from limited power due to a limited sample size, interesting differences may also be due to participants' age differences, detailed clinical diagnoses, or the duration, target, and modality of the cognitive training. Regardless, the changes and improvements in brain function are of interest

because the brain regions identified as showing changes in activity with training have all been associated with dysregulation of cognitive control and reward processing in ADHD [77, 78].

The cognitive training regimen used by Hoekzema et al. [25] has also been demonstrated to increase gray matter volume in select regions [79] in a study of 18 children with ADHD combined type, divided into two matched groups to receive either the experimental cognitive training treatment or the comparison social problems intervention. After a 2-week cognitive training treatment, children with ADHD exhibited increases in brain volume within the middle frontal cortex and inferior–posterior cerebellum, demonstrating that the treatment has the capacity to alter the developmental trajectory of these brain regions as compared to a social problem solving training control condition. This added growth may attenuate ADHD symptoms by targeting those regions that have already been shown to be volumetrically deficient in ADHD [77].

In healthy adults, recent research has moved beyond examining change at the level of brain functioning to investigating change in structural connectivity. Takeuchi et al. [80] used voxel-based analysis of fractional anisotropy measures to assess the integrity of fiber tracts through diffusion tensor imaging in 11 typically developing adults. Post-treatment images were only contrasted with pre-treatment images, as there was not a nontreatment control condition used for comparison. The authors demonstrated that working memory training was correlated with improved connectivity of regions surrounding the intraparietal sulcus, as well as the anterior corpus callosum. The associated white matter tracts compose the connections between the lateral prefrontal and parietal cortices that support working memory [81].

A positron emission tomography study evaluated the neurotransmitter underpinnings of the working memory training developed by Klingberg et al., [48], without a placebo condition, in 13 healthy adults [27]. The study showed increased density of dopamine D1-receptor binding in the prefrontal and parietal cortex after treatment. Thus, training may improve behavioral and cognitive impairments of ADHD related to disturbances of the dopamine system [27, 37]. These changes in dopamine functioning may also identify a mechanism by which the cognitive training effects could likely generalize and persist beyond the training period as dopamine is involved in learning.

Developmental Considerations for Cognitive Training

Considerations of brain development are also crucial in choosing the timing of the treatment, as this may impact the effectiveness of training. How actively a brain region is

developing is likely to relate to the effectiveness and generalizability of training. Accordingly, neurodevelopmental age may be a predictor for identifying which types of training may be most effective at targeting different impairments associated with ADHD. Attempting to understand the developmental trajectory of the human brain is complicated by the fact that not all regions develop at the same rate. For example, sensory regions mature about 10 years sooner than regions associated with executive functions (for more detail see Luna [78] for a review). Furthermore, the striatum, which is heavily involved in reward processing and impulsive behavior, plateaus in development more quickly than the prefrontal cortex, which is tied to cognitive control and executive functions. Research has demonstrated that control processes mature rather linearly through childhood into adolescence and adulthood [82], with working memory reaching a plateau at age 23 in typically developing individuals [78] and age 26 in individuals with ADHD [83]. On the other hand, the striatum follows a curvilinear developmental trajectory resembling an inverted-U shape [84], coming online in childhood, reaching peak sensitivity to reward during early adolescence (between 11–13 years of age), and finally declining during late adolescence and early adulthood (Fig. 1C). In the later adult years, the prefrontal cortex and some of the processes it subserves, again declines [85], thus, having implications for the effectiveness and potential limitations of training techniques for the elderly. Functional connectivity between brain regions develops slowly through late childhood into adulthood. Frontostriatal connectivity is likely atypical in ADHD [86]. This pairing of nonparallel development is hypothesized to lead to impulsive and reckless behaviors that peak during adolescent years [87, 88].

Given these developmental trajectories, the effect of cognitive training on executive functions, such as working memory may yield differential effects before the age of 7 and 12, when only visual and auditory processing regions have reached adult-like maturity, respectively [78]. Treatment may accelerate the development of these sensory processing regions, as well as the connections between them and the prefrontal cortex. This may explain why most studies using cognitive training, which typically have samples ranging in age from 7 to 14 years, demonstrate reliable results in attention and executive functions in contrast to inconsistent improvements in measures of impulsivity [47, 48, 51, 52]. The effectiveness of cognitive training exercises that target functions, such as working memory and attention which are regulated by the top-down structures such as the prefrontal cortex, may be even more successful in later years, such as adolescence and young adulthood when the prefrontal cortex is undergoing significant changes. Findings on cognitive-behavioral strategies for ADHD suggest that a certain degree of brain maturation is necessary for

interventions that require higher order functioning and self-management. For example, clinic-delivered cognitive-behavior therapy, with a focus on planning and thinking in regard to how future outcomes would occur depending upon the child's different actions, was tested for children with ADHD and was determined to be ineffective in improving functioning [89]. However, cognitive-behavioral strategies do appear to be effective for adults with ADHD [90, 91].

Other approaches have attempted to be more deliberate in targeting self-control impairments in young children through training. These approaches may be directly targeting both response to reward and cognitive control pathways, as in the dual-systems model [41, 42] or the ADHD dual-pathway model [29]. Training that targets striatal functioning and response to reward is likely feasible earlier in development, as striatal functioning matures at a relatively earlier age in childhood than regions associated with more high order, prefrontal cortical functioning [84, 92]. Schweitzer and Sulzer-Azaroff [93] improved self-control in young children by directly targeting response to the size and delay of a concrete reward. They manipulated the delay that preschool-aged children with ADHD had to wait for a larger reward, over a smaller, more immediate reward. The training program required several weeks; however, the children all demonstrated greater preference for delayed, larger rewards in comparison with more immediate, smaller rewards after the training than they had prior to the training. Interventions that have taken an approach that targets both reward and cognitive control pathways also have been shown to work in children at-risk for later problems. For instance, the Tools of the Mind Curriculum is used to teach children "patience" and self-control. It is integrated into preschool settings and has been shown to improve executive functioning for children at-risk [94]. Halperin et al. [95, 96] and Sonuga-Barke and Halperin [97] have proposed that it is possible to prevent or alter the trajectory of ADHD in preschool children through a program of activities focusing on cognitive and motor skills (for more detail see Halperin and Healey [98]), which targets the underlying pathology associated with the dual pathway model. Halperin et al. [95] demonstrated improvement on parent- and teacher-rated ADHD symptom scales for preschoolers with ADHD who underwent a program training a variety of executive functioning skills, including attention, inhibition, working memory, planning, visuospatial, and motor skills. The improvement was maintained at a follow-up period 3 months later. Theoretically, these programs will have long-term beneficial effects in preparing children to cope with temptation, as they directly teach the use of cognitive control strategies to modulate reward response.

We propose that early adolescence (in the 11-13 age range) affords another window of opportunity for reducing ADHD symptoms and specifically targeting impulsivity through cognitive training. This is a developmental period with the largest

discrepancy between prefrontal and striatal performance. The active improvement in cognitive control-prefrontal cortical-parietal functioning during adolescence suggests that it may be sensitive to intervention to address the heightened sensitivity to reward information during the adolescent years that has yet to be directed by cognitive-control functioning.

We know of only one study that has actually tested the role of development in cognitive training in brain functioning. Jolles et al. [99] assessed changes in resting state functional connectivity in adults ($n=15$) (22.04 years of age) and a small group of children ($n=9$) (12.24 years of age) after 6 weeks of working memory training; the study did not include a nontreatment control group for comparison. The adult and pediatric samples improved their performance with training. The adult sample demonstrated increased connectivity between the right middle frontal gyrus and regions in the frontoparietal network (i.e., anterior cingulate cortex, bilateral superior frontal gyrus, and paracingulate gyrus) and decreased functional connectivity between the medial prefrontal cortex and the right posterior middle temporal gyrus with working memory training. In contrast, the pediatric sample did not demonstrate changes in the same network after training, although they used the same network to perform the tasks. The absence of detectable changes in resting state functional connectivity could be due to insufficient power, considering the small sample size of the child group. However, the authors tried to address this by conducting an additional analysis, matching a smaller sample size of adults to the pediatric sample. The matched smaller sample analysis continued to yield differences in the between-age analysis with the adults only demonstrating an increase in functional connectivity between the right middle frontal gyrus and other frontoparietal regions. The between-age analysis with the matched sample size for adults no longer demonstrated age differences for the network associated with reductions in functional connectivity (medial prefrontal cortex to posterior middle temporal gyrus) after training in adults. One of the hypotheses put forth by the authors is that a lack of planning and response preparation in children is responsible for changes in functional connectivity. The authors also postulate that movement artifacts, interindividual variability or variability in the use of cognitive strategies by pediatric subjects also may have caused it to be more difficult to detect changes associated with training in the child group analysis. Future research should attempt to delineate how age relates to effectiveness of cognitive training across multiple domains.

Interventions Targeting Reward Systems

Traditional behavioral approaches [9] to treating ADHD within parent or school-administered behavioral programs

target specific behaviors and do not target “meta” processing in general response to reward. Previous research, based on the operant animal research [100], has also demonstrated that targeting impulsivity and response to reward magnitude and immediacy with a specifically tailored intervention can yield improvements in impulsivity in ADHD. As noted earlier, Schweitzer and Sulzer-Azaroff [93] utilized a teaching procedure that increased self-control in preschool-aged children exhibiting impulsivity by gradually introducing larger delay intervals over numerous training sessions. To date, we are aware of no other studies that have revisited this approach. The effectiveness of the “shaping” procedure in the Schweitzer and Sulzer-Azaroff [93] study may be related to enhanced connections between the striatum and prefrontal cortex or merely a reduction in striatal activity in response to a more immediate reward. Other attempts to reduce the attractiveness of the immediate reward have been tested in the form of introducing extra stimulation during delay periods to the more delayed, larger reward. One study found that participants with ADHD responded more like controls [101], whereas another found that the beneficial effects of the extra stimulation on self-control were not maintained across several sessions [102].

“Nudges” and Reframing Approaches from Behavioral Economics

The assessment of the potential benefits of cognitive training would be improved by supplementing current measurement techniques (e.g., rating scales, cognitive paradigms assessing near transfer effects) with additional objective measures that test for generalizability across settings. Outcome measures assessing far transfer effects that are clinically relevant are particularly important in assessing the effects of cognitive training programs.

Delay discounting paradigms add value to this endeavor by providing an intermediate metric that sensitively assesses a cognitive skill that has demonstrated links to important behavioral outcomes. Delay discounting methods have a rich history in the operant literature, but we focus here on recent adaptations of it in the behavioral economics literature. The behavioral economic literature has arrived at a model of delay discounting that is alluringly similar in form to accounts of ADHD [29]. This body of work moves past what we typically find in cognitive and behavioral research studies and examines behaviors on wider scales across much larger samples. We may therefore have the opportunity to use the information gathered therein, to review what has been successful, and perhaps to apply the findings at the level of the individual in a clinical interventional setting. This brief overview of behavioral economics only begins to postulate the potential for

bridging ideas between the two fields; the countless, highly developed models and theories may continue to provide fruitful avenues to evaluate and improve our clinical interventions.

The behavioral economics literature is concerned with how economic decision-making is influenced by cognitive, emotional, and environmental factors. A segment of the field has been specifically interested in how choice is made in regard to rewards that vary in value depending on the size of the reward and the length of the delay to the reward. Subtle experimental manipulations have been shown to influence behavior by altering the function of specific cognitive processes, and this can enable a better understanding of these systems in ADHD. Beyond the informative capacity of other behavioral, developmental, and cognitive theories, this behavioral economics literature provides uniform ways to assess the impact of different interventions, a strong limitation found in the ADHD intervention literature. The clinical value of this literature is abundantly evident in the substance abuse field, which has repeatedly demonstrated the generalizability between measures used in behavioral economics and clinical functioning (e.g., [103]). We will look to the behavioral economics literature to inform us as to how its theory, models, and research on factors influencing choice may be considered in developing treatments to address impulsivity (choice for more immediate, smaller rewards of delayed, larger rewards) in ADHD.

In economics, attention to future-oriented goals is assumed to depend on the present subjective value, or utility, of the prospective outcomes. The assumption is that attention is directed to a task only if the present subjective value of the goal outcome is greater than other (potentially more proximate) distracters. Present utility derives from some combination of visceral value and cognitively construed benefit relative to an overarching behavioral goal [41]. Measurement of time-dependent subjective values relies on revealed preferences, commonly assessed in the laboratory using delay discounting procedures. In delay discounting, choices are made between immediate or delayed rewards of arbitrary values [104, 105]. For instance, participants are asked to choose between smaller, sooner rewards (e.g., \$20 today) and a series of larger, later rewards (e.g., \$25 in 1 month from now, or \$40 in 6 months from now?). The outcomes of interest are the points in which the participant is indifferent between a relatively large reward after a long delay and a smaller reward after a short delay. Functions fit to these indifference points express how value declines as behavioral objectives become more distant (see Fig. 1D). Usually a summary discount rate is computed, which expresses an individual’s relative preference for immediate rewards, with greater discount rates indicative of stronger bias for immediate outcomes, and hence greater potential

for distractibility [106]. Thus, these procedures provide objective, quantifiable measures of decision-making.

As expected, developmental studies have shown that discount rates decline as a function of age [107, 108]. On the other hand, discount rates are shown to be stable within healthy adult individuals [109]. More importantly, discount rates differ systematically across a number of disorders that involve impulsivity problems, including addiction, aggressive disorders, and ADHD [36, 102, 110, 111]. In all of these cases, the clinical groups have been found to have higher discount rates (i.e., more impulsivity) than appropriately matched controls. Despite the clear relevance of the measure, the recent emphasis in behavioral economics has been on how variable discounting truly is. For example, in a recent article, Frederic and Loewenstein [112] concluded that “given [the] variation in apparent discounting across circumstances, and across various descriptions of the same circumstances, we question the usefulness of attempting to produce any single parametric specification of discounting.” This conclusion followed from the fact that discount rates can be affected by a wide variety of contextual or individual state factors.

Some of these factors are intuitive. For example, delay discounting is elevated when we are hungry, tired, or emotionally aroused so that emotional responses that underlie impulsivity are exaggerated [113–116]. Discounting is further elevated when distracted by a secondary task that competes for attention [117, 118]. The complexity of contextual effects on discounting becomes beguiling when considering other known effects. For example, people value future outcomes more when expressed in definite terms (e.g., to be received on November 1st) as opposed to equivalent delays expressed in relative terms (e.g., to be received in 4 weeks [119]). Discounting also depends on whether preferences are elicited by specific choices (i.e., Would you rather have \$100 today or \$150 in 1 year?) as opposed to more open-ended elicitation methods (i.e., What is the minimum you would accept in a year instead of \$100 today?) [112]. There are also interesting framing effects that influence preferences. People are more patient when presented with the opportunity to “speed up” a delayed outcome compared with an equivalent choice requiring the postponement of an immediate outcome [120]. Discounting is also reduced when foregone outcomes are emphasized. Therefore, preference is greater for future outcomes when they are expressed as “nothing now but more later” and immediate outcomes are expressed as “something now but nothing later” [121]. Preference for greater future outcomes also increases when the choice set consists of options with large rather than small magnitudes (e.g., \$1000 today or \$5000 in a year vs \$10 today or \$50 in a year [122]). Finally, it has been shown that when the most recent option is immediately presented, people are significantly more impulsive than when the most

recent option is also presented at a delayed period [123]. Thus, this so-called “immediacy effect” may suggest how certain context may increase impulsive behavior.

In light of the myriad effects that influence a person’s judgment of future reward, it is understandable that Frederic and Loewenstein [112] urged “an approach that focuses on the reasons, considerations, motives, and perspectives that influence evaluations of temporal prospects”. We anticipate that training interventions would benefit from precisely such considerations. For example, the research discussed previously that identified the relationship between framing manipulations and behavioral effects has been used to promote improvements in far-sighted behaviors [124]. This application suggests that cognitive training, too, is likely to be most effective when combined with environmental changes that facilitate the desired cognitive and behavioral changes. Although cognitive training interventions focus on general processes rather than specific behaviors, and therefore have the capacity to yield broadly reaching effects, targeting cognition alone is not as likely to have a major effect on behavioral change as a combined approach. According to the behavioral economics research body, large-scale behavioral change is best elicited through simultaneous education and environmental change [125]. Therefore, although generalizability of treatment effects can potentially be attained through cognitive training, it stands to reason that the intervention would benefit from added, interactive, environmental manipulations that may accelerate generalization and facilitate the maintenance of sustained outcomes. This may help to explain why positive effects can be seen in some cognitive training programs but not others: Cogmed’s [126] working memory treatment requires a coach to facilitate treatment compliance primarily through the use of a reward system. It is possible that incorporating this environmental change of a reward system that includes real-life choices based on the personal valuation of rewards, along with an automatic forum for practicing its implementation (to encourage completion of the training itself), catalyzes the effects of the cognitive training. Beyond this observation, we offer a number of applications of behavioral economics research in the consideration of ADHD interventions. While we do note that proper investigations of the multiple factors related to delay discounting, as it applies to treatment effects, would be a large and complex undertaking, requiring large sample sizes and multifactorial designs, the suggestions listed here are meant to illustrate more simple relationships between ADHD interventions and factors highlighted by behavioral economics research as starting points. Indeed, there is a large body of research that has been accumulating for decades in which factors related to delay discounting have been systematically studied in animals and humans, which could serve to inform intervention research on ADHD [103].

First, people with ADHD who are predisposed to distraction may be aided by structural changes to their environment

that protect against deleterious behaviors. Subtle changes that emphasize, for example, the foregone outcomes associated with any choice (something now but nothing later) may facilitate behavioral change promoted by the intervention. Children with ADHD are not universally low on attentional capacity – a fact that creates great consternation among parents who wish the same focus was applied to homework as is applied to video games. It is reasonable to suspect that behavioral economics, through the variations in the delay discounting paradigm, has uncovered important constructs whose manipulation underlies such situational variations in attention.

Second, Sonuga-Barke's [29] dual-process theory posits systems that may be differentially influenced by these framing manipulations. Drawing attention to specific future dates elevates dorsal lateral prefrontal cortex activity [127], an effect that is related to the date-delay framing manipulation previously mentioned [119]. By contrast, the presence of an immediate reward in the choice set is thought to elevate activity in the ventral striatum [42, 45], which may be related to the immediacy effect. Accordingly, there may be a spectrum of ADHD phenotypes that may be revealed by studying the relative size of effects identified in behavioral economics. For instance, a recent study by Scheres and colleagues [128] showed that children and adolescents with combined type ADHD discounted steeper than those diagnosed with inattentive type ADHD, and that association between steep discounting and hyperactivity–impulsivity was particularly strong when the magnitude of the options was small. Discovering the different mechanisms underlying the ADHD phenotypes would be beneficial in understanding, diagnosing, and treating the disorder and could reveal specific (neural) targets for training.

Finally, it should be possible to design cognitive training programs that specifically train persons with ADHD to make better choices under the various contexts that we know influence choice. These training programs could come in a variety of formats. For young children in a preschool setting, they may be integrated into the curriculum. For older and younger children they could also be in the form of a computer game with the rewards in the form of points or access to another game. One could integrate findings from behavioral economics into the training programs. For example, this could include gradually manipulating the degree of distraction present during the times of choice for rewards, the emotional valence of the stimuli, the saliency and visceral nature of the stimuli, or social factors (e.g., the presence of peers or observers, real or fictitious). For more mature individuals with ADHD, it might be sufficient to educate them not to make important decisions during a time when they are distracted, experiencing stress, or during periods of high cognitive demands. Marketing research shows that people with greater impulsivity are more likely

to choose with their “heart” than their “mind” and discount more when cognitive processing resources are in demand [117]. Thus, it is important to teach people with ADHD that they should not make decisions during or after a demanding task. Another variation borrowed from the behavioral economics literature would be to actively incorporate “nudging” into a saving program. Interventions could be designed to “nudge” behavior toward the default of “saving” and “banking” earnings won in a game so that there is greater value and reward for saving rather than spending the points as one earns them in a game.

To enhance generalization, computer programs could use avatars with virtual reality software to transfer the generalization of the training programs from a computer to a real setting. Virtual reality software for classroom settings already exists and could be modified to teach better self-control and attention for children with ADHD. One major advantage of the computerized training programs is the opportunity to repeatedly practice choosing in self-control situations.

The Effects of Cognitive Training on Reward Processes

Bickel et al. [129] produced the only study to date that has examined how cognitive training affects performance on a delay discounting task. The authors utilized a working memory training program in a sample of individuals addicted to stimulants to determine what domains improved following training. This population is of particular interest because the rate of substance abuse in ADHD is significantly higher than those who do not have ADHD [130, 131]. Both ADHD and substance abuse are associated with elevated delay discounting. Participants in the study (who were currently being treated for substance abuse) were assigned to either an active ($n=14$) or control ($n=13$) treatment condition. In the active treatment condition (based on the commercially-available cognitive training program PSSCogReHab), participants completed sessions by practicing tasks of serial recall of numbers and words in addition to a free word recall task. These participants completed sessions until they ceased to show improvement across three consecutive sessions. The control treatment condition included the same tasks, but the answers were provided so that participants did not need to invoke working memory to perform the tasks. After treatment, the participants in the active working memory treatment condition showed significantly decreased delay discounting behaviors on a choice task containing hypothetical and real monetary rewards. The causal relationship between working memory training and delay discounting may be explained by the improvement in executive functioning that arises, leading to a greater ability of the system to override the activity of the impulsive frontal-striatal system. A reasonable question would be

whether cognitive training would also decrease impulsivity in delay discounting situations for individuals with ADHD.

Summary and Future Directions

This review discussed the current status of research on interventions, with implications to target cognitive impairments associated with ADHD and a call to incorporate concepts from other fields of study, such as the behavioral economics literature. The proliferation of research articles and the availability of commercially available cognitive training programs for ADHD [48, 53, 55, 65, 72, 132–136] are becoming increasingly more popular; however, the evidence base for these programs is still forming with few randomized, placebo-controlled studies involving well-characterized, large samples (for more detail see Epstein and Tsai [14] and Klingberg [137] for reviews and Klingberg et al., [48], Rabiner et al. [55], and Holmes et al. [136] for research articles). The field is at a critical juncture where it must produce well-designed clinical trials using objective and rating scale data to demonstrate whether or not the training techniques are effective. There is some evidence for the generalization of training to nontrained skills and behaviors, and for the maintenance of training significantly after the training period [48, 136], but importantly, improvements in ADHD behaviors do not always generalize across home and school settings.

Considering the neural processes involved in self-control, which is balanced between reward-driven and cognitive control systems [42], it stands to reason that directly targeting reward systems or the balance between cognitive control and reward systems, could improve ADHD behaviors. Recent research has provided evidence of this assertion as cognitive training also targeting self-control has yielded improvements in ADHD symptoms both at home and in the classroom environment in young children [95]. How well the changes in performance or neural functioning associated with cognitive training relate to the generalization of training effects is unknown and an important area of study.

If the programs are found to be effective, there will need to be finer grained study of how the components within the cognitive training programs relate to their observed benefits. Issues regarding the modality of training (e.g., computerized *vs* paper and pencil; home *vs* school), dose and coaching are still lacking. Findings from a recent study [138] suggest that working memory training [48] should target secondary memory rather than primary memory with the ADHD population. However, another study, using a different methodology found impaired primary (or short-term) capacity and deficits in articulatory rehearsal in children with ADHD [61, 62]. Based on these findings, Rapport et al. [61] and Bolden

et al. [62] suggest that training programs that incorporate practice on rehearsing information over increasingly longer time periods will ultimately be more successful than simply requiring patients with ADHD to remember longer spans of information. Understanding the role of these training parameters is particularly important considering that these programs are commercially available for addressing a clinical disorder. Age of training likely also affects the utility of various forms of training, as neurological structures do not develop at the same rates from birth to adulthood. Our review of the developmental literature suggests that adolescence provides another opportunity for intervention in ADHD and that cognitive training may be more effective during early adolescence over earlier periods of development due to the development of the prefrontal cortex and the parietal cortex in adolescence. There is a great need for interventions that may be effective and palatable to adolescents with ADHD. Cognitive training interventions may help fill that niche. Future studies should also assess if there is a relationship between brain maturation and sensitivity to the training. Other individual differences, beyond age, are also likely to play a role in who will most benefit from cognitive training [139].

Furthermore, the platform of training will most likely evolve as tablets and smart phones are increasingly more available and may be preferred by families over computer or paper and pencil options. Yet, the programs we reviewed are not currently available on these formats. The cost involved in the interventions is also an important factor that should be addressed. At this point, these training programs are not reimbursable by insurance policies and are unlikely to be covered until there is more support for their effectiveness. Currently, there are programs available on the internet for free *versus* others that charge a fee and are only accessed via a certified “coach” training in the product. Cost effectiveness studies will need to be done to determine if the cost is reasonable and the service results in enhanced performance on relevant indices.

Finally, the possibility of medications enhancing cognitive training or cognitive training enhancing the effectiveness of medications for ADHD has yet to be systematically explored and reported on. Theoretically, stimulant medications and other medications may improve active task performance on cognitive training programs; however, the ultimate question is whether or not these medications would improve the generalization and maintenance of the training effects. Both stimulant (e.g., dextroamphetamine and methylphenidate) and nonstimulant medications (e.g., atomoxetine) affect dopamine and norepinephrine and are likely to improve performance on the cognitive training task by improving working memory, but also by enhancing sensitivity to rewards, and thus it is possible that they could increase the saliency of the training stimuli, intrinsic or extrinsic

rewards for participating in the training. Ultimately, there is much work to be done in the evaluation of cognitive training for ADHD, whether it be in combination with other treatments for the disorder or as a stand alone intervention.

Acknowledgments This work is supported by the DHHS Administration for Children and Families (90DD0596/01 to J.B.S.), the National Institutes of Health-National Center for Research and Resources (UL1 RR024146-06 to J.B.S.), and the Rubicon, Netherlands Organisation for Scientific Research (to WvdB). The authors have no real or perceived conflict of interest. Full conflict of interest disclosures is available in the electronic supplementary material for this article.

Required Author Forms Disclosure forms provided by the authors are available with the online version of this article.

References

- American Psychiatric Association Press. Diagnostic and Statistical Manual of Mental Disorder, fourth ed. Arlington, VA: American Psychiatric Publishing, 2000.
- Polanczyk G, de Lima MS, Horta BL, Biederman J, Rohde LA. The worldwide prevalence of ADHD: a systematic review and meta-regression analysis. *The Am J Psychiatry* 2007;164:942-948.
- Breslau J, Miller E, Breslau N, Bohnert K, Lucia V, Schweitzer J. The impact of early behavior disturbances on academic achievement in high school. *Pediatrics* 2009;123:1472-1476.
- Raggi VL, Chronis AM. Interventions to address the academic impairment of children and adolescents with ADHD. *Clin Child Fam Psychol Rev* 2006;9:85-111.
- Hinshaw SP. Academic underachievement, attention deficits, and aggression: comorbidity and implications for intervention. *J Consult Clin Psychol* 1992;60:893-903.
- Barkley RA, Fischer M, Smallish L, Fletcher K. Young adult outcome of hyperactive children: adaptive functioning in major life activities. *J Am Acad Child Adolesc Psychiatry* 2006;45:192-202.
- Polderman TJ, Boomsma DI, Bartels M, Verhulst FC, Huizink AC. A systematic review of prospective studies on attention problems and academic achievement. *Acta Psychiatr Scand* 2010;122:271-284.
- Breslau J, Miller E, Joanie Chung WJ, Schweitzer JB. Childhood and adolescent onset psychiatric disorders, substance use, and failure to graduate high school on time. *J Psychiatr Res* 2011;45:295-301.
- Barkley RA. Attention deficit hyperactivity disorder: a handbook for diagnosis and treatment, 3rd ed. New York: Guilford Press, 2006.
- Wilens T. Attention-deficit/hyperactivity disorder and the substance use disorders: the nature of the relationship, subtypes at risk, and treatment issues. *Psychiatr Clin North Am* 2004;27:361-372.
- Levin FR, Evans SM, Kleber HD. Prevalence of adult attention-deficit hyperactivity disorder among cocaine abusers seeking treatment. *Drug Alcohol Depend* 1998;52:15-25.
- The MTA Cooperative Group. A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry* 1999;56:1073-1086.
- Molina BS, Hinshaw SP, Swanson JM, et al. The MTA at 8 years: prospective follow-up of children treated for combined-type ADHD in a multisite study. *J Am Acad Child Adolesc Psychiatry* 2009;48:484-500.
- Epstein JN, Tsal Y. Evidence for cognitive training as a treatment strategy for children with attention deficit/hyperactivity disorder. *J ADHD Relat Disord* 2010;1:49-64.
- Johnston C, Hommersen P, Seip C. Acceptability of behavioral and pharmacological treatment for attention — deficit hyperactivity disorder: relations to child and parent characteristics. *Behav Ther* 2008;39(1):22-32.
- Hoza B, Johnston C, Pillow D, Ascough JC. Predicting treatment response for childhood attention-deficit hyperactivity disorders: introduction of a heuristic model to guide research. *Applied and Preventive Psychology: Current Scientific Perspectives* 2006;11:215-229.
- Jensen PS, Kettle L, Roper MT, et al. Are stimulants overprescribed? Treatment of ADHD in four U.S. communities. *J Am Acad Child Adolesc Psychiatry* 1999;38:797-804.
- Tremmery S, Buitelaar JK, Steyaert J, et al. The use of health care services and psychotropic medication in a community sample of 9-year-old schoolchildren with ADHD. *Eur Child Adolesc Psychiatry* 2007;16:327-336.
- Chronis AM, Jones HA, Raggi VL. Evidence-based psychosocial treatments for children and adolescents with attention-deficit/hyperactivity disorder. *Clin Psychol Rev* 2006;26:486-502.
- Pelham WE Jr., Wheeler T, Chronis A. Empirically supported psychosocial treatments for attention deficit hyperactivity disorder. *J Clin Child Psychol* 1998;27:190-205.
- Anastopoulos AD, Guevremont D, Shelton T, DuPaul GJ. Parenting stress among families of children with attention deficit hyperactivity disorder. *J Abnorm Child Psychol* 1992;20:503-520.
- Abikoff H, Ganeles D, Reiter G, Blum C, Foley C, Klein RG. Cognitive training in academically deficient ADHD boys receiving stimulant medication. *J Abnorm Child Psychol* 1988;16:411-432.
- Loo S, Makeig S. Clinical utility of EEG in attention-deficit/hyperactivity disorder: a research update. *Neurotherapeutics* 2012; doi:10.1007/s13311-012-0123-z.
- Moriyama T, Polanczyk GV, Caye A, Banaschewski T, Brandeis D, Rohde LA. Evidence based information on the clinical use of neurofeedback for ADHD. *Neurotherapeutics*. 2012; doi:10.1007/s13311-012-0136-7.
- Hoekzema E, Carmona S, Tremols V, et al. Enhanced neural activity in frontal and cerebellar circuits after cognitive training in children with attention-deficit/hyperactivity disorder. *Hum Brain Map* 2010;31(12):1942-1950.
- Klingberg T. Development of a superior frontal-intraparietal network for visuospatial working memory. *Neuropsychologia* 2006;44:2171-2177.
- McNab F, Varrone A, Farde L, et al. Changes in cortical dopamine D1 receptor binding associated with cognitive training. *Science* 2009;323:800-802.
- Abikoff H. ADHD psychosocial treatments: generalization reconsidered. *J Atten Disord* 2009;13:207-210.
- Sonuga-Barke. Psychological heterogeneity in AD/HD — a dual pathway model of behaviour and cognition. *Behav Brain Res* 2002;130:29-36.
- Sonuga-Barke EJ. The dual pathway model of AD/HD: an elaboration of neuro-developmental characteristics. *Neurosci Biobehav Rev* 2003;27:593-604.
- Barkley RA. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol Bull* 1997;121:65-94.
- Sergeant J. The cognitive-energetic model: an empirical approach to attention-deficit hyperactivity disorder. *Neurosci Biobehav Rev* 2000;24:7-12.
- Castellanos FX, Sonuga-Barke EJ, Milham MP, Tannock R. Characterizing cognition in ADHD: beyond executive dysfunction. *Trends Cogn Sci* 2006;10:117-123.
- Nigg JT. Attention-deficit/hyperactivity disorder: endophenotypes, structure, and etiological pathways. *Curr Dir Psychol Sci* 2010;19:24-29.
- Sagvolden T, Johansen EB, Aase H, Russell VA. A dynamic developmental theory of attention-deficit/hyperactivity disorder

- (ADHD) predominantly hyperactive/impulsive and combined subtypes. *Behav Brain Sci* 2005;28:397-419.
36. Sonuga-Barke EJ, Taylor E, Sembi S, Smith J. Hyperactivity and delay aversion — I. The effect of delay on choice. *J Child Psychol Psychiatry* 1992;33:387-398.
 37. Volkow ND, Wang GJ, Newcorn JH, et al. Motivation deficit in ADHD is associated with dysfunction of the dopamine reward pathway. *Mol Psychiatry* 2011;16:1147-1154.
 38. Volkow ND, Wang GJ, Kollins SH, et al. Evaluating dopamine reward pathway in ADHD: clinical implications. *JAMA* 2009;302:1084-1091.
 39. Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 2001;24:167-202.
 40. Sonuga-Barke EJ, Coghill D, DeBacker M, Swanson J. Measuring methylphenidate response in attention-deficit/hyperactivity disorder: how are laboratory classroom-based measures related to parent ratings? *J Child Adolesc Psychopharmacol* 2009;19:691-698.
 41. Loewenstein G. Out of control: visceral influences on behavior. *Organ Behav Hum Decis Process* 1996;65:272-292.
 42. McClure SM, Laibson DI, Loewenstein G, Cohen JD. Separate neural systems value immediate and delayed monetary rewards. *Science* 2004;306:503-507.
 43. McClure SM, Berns GS, Montague PR. Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 2003;38:339-346.
 44. O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal difference models and reward-related learning in the human brain. *Neuron* 2003;38:329-337.
 45. McClure SM, Ericson KM, Laibson DI, Loewenstein G, Cohen JD. Time discounting for primary rewards. *J Neurosci* 2007;27:5796-5804.
 46. Hare TA, Camerer CF, Rangel A. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 2009;324:646-648.
 47. Kerns KA, Eso K, Thompson J. Investigation of a direct intervention for improving attention in young children with ADHD. *Dev Neuropsychol* 1999;16:273-295.
 48. Klingberg T, Fernell E, Olesen PJ, et al. Computerized training of working memory in children with ADHD—a randomized, controlled trial. *J Am Acad Child Adolesc Psychiatry* 2005;44:177-186.
 49. Shipstead Z, Redick TS, Engle RW. Is working memory training effective? *Psychol Bull* 2012;138(4):628-654.
 50. Sohlberg MM, Mateer CA. Effectiveness of an attention-training program. *J Clin Exp Neuropsychol* 1987;9:117-130.
 51. Steiner NJ, Sheldrick RC, Gotthelf D, Perrin EC. Computer-based attention training in the schools for children with attention deficit/hyperactivity disorder: a preliminary trial. *Clin Pediatr (Phila)* 2011;50:615-622.
 52. Tamm L, Hughes C, Ames L, et al. Attention training for school-aged children with ADHD: results of an open trial. *J Atten Disord* 2010;14:86-94.
 53. Shalev L, Tsal Y, Mevorach C. Computerized progressive attentional training (CPAT) program: effective direct intervention for children with ADHD. *Child Neuropsychol* 2007;13:382-388.
 54. Posner MI, Rothbart MK. Influencing brain networks: implications for education. *Trends Cogn Sci* 2005;9:99-103.
 55. Rabiner DL, Murray DW, Skinner AT, Malone PS. A randomized trial of two promising computer-based interventions for students with attention difficulties. *J Abnorm Child Psychol* 2010;38:131-142.
 56. Fassbender C, Schweitzer JB, Cortes CR, et al. Working memory in attention deficit/hyperactivity disorder is characterized by a lack of specialization of brain function. *PLoS One* 2011;6:e27240.
 57. Schoechlin C, Engel RR. Neuropsychological performance in adult attention-deficit hyperactivity disorder: meta-analysis of empirical data. *Arch Clin Neuropsychol* 2005;20:727-744.
 58. Martinussen R, Hayden J, Hogg-Johnson S, Tannock R. A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 2005;44:377-384.
 59. Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF. Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biol Psychiatry* 2005;57:1336-1346.
 60. Mills KL, Bathula D, Dias TG, et al. Altered cortico-striatal-thalamic connectivity in relation to spatial working memory capacity in children with ADHD. *Front Psychiatry* 2012;3:2.
 61. Rapport MD, Alderson RM, Kofler MJ, Sarver DE, Bolden J, Sims V. Working memory deficits in boys with attention-deficit/hyperactivity disorder (ADHD): the contribution of central executive and subsystem processes. *J Abnorm Child Psychol* 2008;36:825-837.
 62. Bolden J, Rapport MD, Raiker JS, Sarver DE, Kofler MJ. Understanding phonological memory deficits in boys with attention-deficit/hyperactivity disorder (ADHD): dissociation of short-term storage and articulatory rehearsal processes. *J Abnorm Child Psychol* 2012;40(6):999-1011.
 63. McVay JC, Kane MJ. Conducting the train of thought: working memory capacity, goal neglect, and mind wandering in an executive-control task. *J Exp Psychol Learn Mem Cogn* 2009;35:196-204.
 64. Unsworth N, Redick TS, Spillers GJ, Brewer GA. Variation in working memory capacity and cognitive control: goal maintenance and microadjustments of control. *Q J Exp Psychol (Hove)* 2012;65:326-355.
 65. Klingberg T, Forssberg H, Westerberg H. Training of working memory in children with ADHD. *J Clin Exp Neuropsychol* 2002;24:781-791.
 66. Tallal P, Miller SL, Bedi G, et al. Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*. 1996;271:81-84.
 67. Holmes J, Gathercole SE, Dunning DL. Adaptive training leads to sustained enhancement of poor working memory in children. *Dev Sci* 2009;12:F9-F15.
 68. Holmes J, Gathercole SE, Place M, Dunning DL, Hilton KL, Elliott JG. Working memory deficits can be overcome: impacts of training and medication on working memory in children with ADHD. *Appl Cogn Psychol* 2010;24:827-836.
 69. Alloway TP. Automated working memory assessment. London: Harcourt Assessment, 2007.
 70. Alloway TP, Gathercole SE, Pickering SJ. Verbal and visuospatial short-term and working memory in children: are they separable? *Child Dev* 2006;77:1698-1716.
 71. Wechsler D. Wechsler Abbreviated Scale of Intelligence (WASI). San Antonio: Harcourt Assessment, 1999.
 72. Beck SJ, Hanson CA, Puffenberger SS, Benninger KL, Benninger WB. A controlled trial of working memory training for children and adolescents with ADHD. *J Clin Child Adolesc Psychol* 2010;39:825-836.
 73. Green CT, Long DL, Green D, et al. Will working memory training generalize to improve off-task behavior in children with attention-deficit/hyperactivity disorder? *Neurotherapeutics* 2012; doi:10.1007/s13311-012-0124-y.
 74. Olesen PJ, Westerberg H, Klingberg T. Increased prefrontal and parietal activity after training of working memory. *Nat Neurosci* 2004;7:75-79.
 75. Rainer G, Miller EK. Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* 2000;27:179-189.
 76. Schweitzer JB, Lee DO, Hanford RB, et al. Effect of methylphenidate on executive functioning in adults with attention-deficit/hyperactivity disorder: normalization of behavior but not related brain activity. *Biol Psychiatry* 2004;56:597-606.

77. Nigg JT, Casey BJ. An integrative theory of attention-deficit/hyperactivity disorder based on the cognitive and affective neurosciences. *Dev Psychopathol* 2005;17:785-806.
78. Luna B. Developmental changes in cognitive control through adolescence. *Adv Child Dev Behav* 2009;37:233-278.
79. Hoekzema E, Carmona S, Ramos-Quiroga JA, et al. Training-induced neuroanatomical plasticity in ADHD: a tensor-based morphometric study. *Human Brain Mapp* 2011;32:1741-1749.
80. Takeuchi H, Sekiguchi A, Taki Y, et al. Training of working memory impacts structural connectivity. *J Neurosci* 2010;30:3297-3303.
81. Baddeley A. Working memory: looking back and looking forward. *Nat Rev Neurosci* 2003;4:829-839.
82. Davidson MC, Amso D, Anderson LC, Diamond A. Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 2006;44:2037-2078.
83. Shaw P, Eckstrand K, Sharp W, et al. Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proc Natl Acad Sci U S A* 2007;104:19649-19654.
84. Somerville LH, Casey BJ. Developmental neurobiology of cognitive control and motivational systems. *Curr Opin Neurobiol* 2010;20:236-241.
85. Morrison JH, Baxter MG. The ageing cortical synapse: hallmarks and implications for cognitive decline. *Nat Rev Neurosci* 2012;13:240-250.
86. Casey BJ, Epstein JN, Buhle J, et al. Frontostriatal connectivity and its role in cognitive control in parent-child dyads with ADHD. *Am J Psychiatry* 2007;164:1729-1736.
87. Casey BJ, Jones RM, Hare TA. The adolescent brain. *Ann N Y Acad Sci* 2008;1124:111-126.
88. Steinberg L. A behavioral scientist looks at the science of adolescent brain development. *Brain Cogn* 2010;72:160-164.
89. Abikoff H. Cognitive training in ADHD children: less to it than meets the eye. *J Learn Disabil* 1991;24:205-209.
90. Safren SA, Otto MW, Sprich S, Winett CL, Wilens TE, Biederman J. Cognitive-behavioral therapy for ADHD in medication-treated adults with continued symptoms. *Behav Res Ther* 2005;43:831-842.
91. Solanto MV, Marks DJ, Wasserstein J, et al. Efficacy of meta-cognitive therapy for adult ADHD. *Am J Psychiatry* 2010;958-968.
92. Galvan A, Hare TA, Parra CE, et al. Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *J Neurosci* 2006;26:6885-6892.
93. Schweitzer JB, Sulzer-Azaroff B. Self-control: teaching tolerance for delay in impulsive children. *Journal of the experimental analysis of behavior*. 1988;50:173-186.
94. Diamond A, Barnett WS, Thomas J, Munro S. Preschool program improves cognitive control. *Science* 2007;318:1387-1388.
95. Halperin JM, Marks DJ, Bedard AC, et al. Training executive, attention, and motor skills: a proof-of-concept study in preschool children with ADHD. *Journal of attention disorders*. 2012; doi:10.1007/s13311-012-0123-z.
96. Halperin JM, Bedard AC, Curchack JT. Preventive interventions for ADHD: a neurodevelopmental perspective. *Neurotherapeutics* 2012; doi:10.1007/s13311-012-0123-z.
97. Sonuga-Barke EJ, Halperin JM. Developmental phenotypes and causal pathways in attention deficit/hyperactivity disorder: potential targets for early intervention? *J Child Psychol Psychiatry* 2010;51:368-389.
98. Halperin JM, Healey DM. The influences of environmental enrichment, cognitive enhancement, and physical exercise on brain development: can we alter the developmental trajectory of ADHD? *Neurosci Biobehav Rev* 2011;35:621-634.
99. Jolles DD, van Buchem MA, Crone EA, Rombouts SA. Functional brain connectivity at rest changes after working memory training. *Hum Brain Mapp* 2011; doi:10.1002/hbm.21444.
100. Mazur JE, Logue AW. Choice in a "self-control" paradigm: effects of a fading procedure. *J Exp Anal Behav* 1978;30:11-17.
101. Antrop I, Stock P, Verte S, Wiersma JR, Baeyens D, Roeyers H. ADHD and delay aversion: the influence of non-temporal stimulation on choice for delayed rewards. *J Child Psychol Psychiatry* 2006;47:1152-1158.
102. Schweitzer JB, Sulzer-Azaroff B. Self-control in boys with attention deficit hyperactivity disorder: effects of added stimulation and time. *J Child Psychol Psychiatry* 1995;36:671-686.
103. Madden GJ, Bickel WK. Impulsivity: the behavioral and neurological science of discounting. Washington, DC: APA Books, 2009.
104. Mazur JE. An adjusting procedure for studying delayed reinforcement. In: Commons ML, Mazur JE, Nevin JA, Rachlin H, eds. *Quantitative Analysis of Behavior, Vol. 5. The Effect of Delay and Intervening Events on Reinforcement Value*. Hillsdale, NJ: Erlbaum, 1987.
105. Frederick S, Loewenstein G, O'Donoghue T. Time discounting and time preference: a critical review. *J Econ Lit* 2002;40:351-401.
106. Ainslie G. Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychol Bull* 1975;82:463-496.
107. Scheres A, Dijkstra M, Ainslie E, et al. Temporal and probabilistic discounting of rewards in children and adolescents: effects of age and ADHD symptoms. *Neuropsychologia* 2006;44:2092-2103.
108. Steinberg L, Graham S, O'Brien L, Woolard J, Cauffman E, Banich M. Age differences in future orientation and delay discounting. *Child Dev* 2009;80:28-44.
109. Kirby KN. One-year temporal stability of delay-discount rates. *Psychon Bull Rev* 2009;16:457-462.
110. Solanto MV, Abikoff H, Sonuga-Barke E, et al. The ecological validity of delay aversion and response inhibition as measures of impulsivity in AD/HD: a supplement to the NIMH multimodal treatment study of AD/HD. *J Abnorm Child Psychol* 2001;29:215-228.
111. Sonuga-Barke EJ, Williams E, Hall M, Saxton T. Hyperactivity and delay aversion. III: The effect on cognitive style of imposing delay after errors. *J Child Psychol Psychiatry* 1996;37:189-194.
112. Frederick S, Loewenstein G. Conflicting motives in evaluations of sequences. *J Risk Uncertain* 2008;37:221-235.
113. Van den Bergh B, Dewitte S, Warlop L. Bikinis instigate generalized impatience in intertemporal choice. *J Consum Res* 2008;35:85-97.
114. Wilson M, Daly M. Do pretty women inspire men to discount the future? *Proc Biol Sci* 2004;271(Suppl 4):S177-S179.
115. Wang XT, Dvorak RD. Sweet future: fluctuating blood glucose levels affect future discounting. *Psychol Sci* 2010;21:183-188.
116. Giordano LA, Bickel WK, Loewenstein G, Jacobs EA, Marsch L, Badger GJ. Mild opioid deprivation increases the degree that opioid-dependent outpatients discount delayed heroin and money. *Psychopharmacology* 2002;163:174-182.
117. Shiv B, Fedorikhin A. Heart and mind in conflict: the interplay of affect and cognition in consumer decision making. *J Consum Res* 1999;26:278-292.
118. Hinson JM, Jameson TL, Whitney P. Impulsive decision making and working memory. *J Exp Psychol Learn Mem Cogn* 2003;29:298-306.
119. Read JP, Frederick S, Orsel B, Rahman J. Four score and seven years from now: the date/delay effect in temporal discounting. *Manage Sci* 2005;41:1326-1335.
120. Weber EU, Johnson EJ, Milch KF, Chang H, Brodscholl JC, Goldstein DG. Asymmetric discounting in intertemporal choice: a query-theory account. *Psychol Sci* 2007;18:516-523.

121. Magen E, Dweck CS, Gross JJ. The hidden-zero effect: representing a single choice as an extended sequence reduces impulsive choice. *Psychol Sci* 2008;19:648-649.
122. Thaler RH. Some empirical evidence on dynamic inconsistency. *Econ Lett* 1981;8:201-207.
123. Prelec D, Loewenstein G. Decision making over time and under uncertainty: a common approach. *Manage Sci* 1991;37:770-786.
124. Thaler RH, Benartzi S. Save More Tomorrow™: using behavioral economics to increase employee saving. *J Polit Econ* 2004;112: S164-S187.
125. Thaler RH, Sunstein CR. *Nudge: improving decisions about health, wealth, and happiness*. New York: Penguin Books, 2009.
126. *Coaching Manual: Cogmed Working Memory Training*. Retrieved from: <http://www.cogmed.com/>. London, UK: Pearson, 2012.
127. Peters J, Buchel C. Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediotemporal interactions. *Neuron* 2010;66:138-148.
128. Scheres A, Tontsch C, Thoeny AL, Kaczurkin A. Temporal reward discounting in attention-deficit/hyperactivity disorder: the contribution of symptom domains, reward magnitude, and session length. *Biol Psychiatry* 2010;67:641-648.
129. Bickel WK, Yi R, Landes RD, Hill PF, Baxter C. Remember the future: working memory training decreases delay discounting among stimulant addicts. *Biol Psychiatry* 2011;69:260-265.
130. Schubiner H, Tzelepis A, Milberger S, et al. Prevalence of attention-deficit/hyperactivity disorder and conduct disorder among substance abusers. *J Clin Psychiatry* 2000;61:244-251.
131. McAweeney M, Rogers NL, Huddleston C, Moore D, Gentile JP. Symptom prevalence of ADHD in a community residential substance abuse treatment program. *J Atten Disord* 2010;13:601-608.
132. Clarfield J, Stoner G. Research brief: the effects of computerized reading instruction on the academic performance of students identified with ADHD. *School Psych Rev* 2005;34:246-254.
133. Kotwal DB, Burns WJ, Montgomery DD. Computer-assisted cognitive training for ADHD — a case study. *Behav Modif* 1996;20:85-96.
134. Mautone JA, DuPaul GJ, Jitendra AK. The effects of computer-assisted instruction on the mathematics performance and classroom behavior of children with ADHD. *J Atten Disord* 2005;9:301-312.
135. Mezzacappa E, Buckner JC. Working memory training for children with attention problems or hyperactivity: a school-based pilot study. *School Men Health* 2010;2(4):202-208.
136. Holmes J, Gathercole SE, Place M, Dunning DL, Hilton KA, Elliott JG. Working memory deficits can be overcome: impacts of training and medication on working memory in children with ADHD. *Appl Cogn Psychol* 2009;12:9-15.
137. Klingberg T. Training and plasticity of working memory. *Trends Cogn Sci* 2010;14:317-324.
138. Gibson BS, Gondoli DM, Johnson AC, Steeger CM, Dobrzanski BA, Morrissey RA. Component analysis of verbal versus spatial working memory training in adolescents with ADHD: a randomized, controlled trial. *Child Neuropsychol* 2011;17:546-563.
139. Jaeggi SM, Buschkuhl M, Jonides J, Shah P. Short- and long-term benefits of cognitive training. *Proc Natl Acad Sci U S A* 2011;108:10081-10086.