Cognitive Rehabilitation for Attention Deficit/Hyperactivity Disorder (ADHD): Promises and Problems

Diana Tajik-Parvinchi PhD¹; Leah Wright BAHons¹; Russell Schachar MD¹

Abstract

Objective: Cognitive training entails the repeated exercise of a specific cognitive process over a period of time to improve performance on the trained task as well as on tasks that were not specifically trained (transfer effect). Cognitive training shows promise in remediating deficits in children with attention deficit/hyperactivity disorder (ADHD) – a disorder believed to stem from deficient cognitive processes – where the focus has been primarily on training working memory and attention. We discuss evidence from studies that have produced broad, limited, or no transfer effects with the goal of identifying factors that may be responsible for this heterogeneity. Results: There are several implicit assumptions that appear to drive researchers’ decisions regarding both the selection of cognitive abilities to train as well as the training tasks chosen to target those abilities. We identify these implicit assumptions and their weaknesses. We also draw attention to design limitations that may be contributing to lack of transfer. Conclusion: Although the overall pattern of findings from these studies is promising, the methodological and theoretical limitations associated with the literature limit conclusions about the efficacy of cognitive training as a rehabilitation method for ADHD. We hypothesize several suggestions that may improve training effects and summarize the evidence which led to our hypotheses.

Key Words: cognitive training, ADHD, children, rehabilitation

Résumé

Objectif: L’entraînement cognitif comporte l’exercice répété d’un processus cognitif spécifique sur une période de temps afin d’améliorer le rendement à la tâche exercée ainsi qu’à des tâches qui ne faisaient pas spécifiquement partie de l’entraînement (effet de transfert). L’entraînement cognitif est prometteur pour remédier aux déficits chez les enfants souffrant du trouble de déficit de l’attention avec hyperactivité (TDAH) – un trouble estimé provenir de processus cognitifs déficients – alors que l’accent avait d’abord été mis sur l’entraînement de la mémoire de travail et de l’attention. Nous discutons des données probantes d’études qui ont produit des effets de transfert vastes, limités ou nuls dans le but d’identifier des facteurs qui peuvent être responsables de cette hétérogénéité. Résultats: Il y a plusieurs hypothèses implicites qui semblent mener les décisions des chercheurs à l’égard de la sélection des capacités cognitives à entraîner et des tâches d’entraînement choisies pour cibler ces capacités. Nous identifions ces hypothèses implicites et leurs faiblesses. Nous attirons aussi l’attention sur les limitations de la méthodologie qui peuvent contribuer à l’absence de transfert. Conclusion: Bien que le modèle global des résultats de ces études soit prometteur, les limitations méthodologiques et théoriques associées à la littérature restreignent les conclusions sur l’efficacité de l’entraînement cognitif comme méthode de réhabilitation du TDAH. Nous proposons plusieurs suggestions qui peuvent améliorer les effets de l’entraînement et résumons les données probantes qui ont mené à nos propositions.

Mots clés: trouble de déficit de l’attention avec hyperactivité, TDAH, entraînement cognitif, remédiation

¹Hospital for Sick Children, Department of Psychiatry, Toronto, Ontario

Corresponding E-Mail: diana.parvinchi@sickkids.ca

Submitted: January 10, 2014; Accepted: July 25, 2014
In the past decade cognitive training has received considerable attention as an intervention method. This attention has been partly stimulated by the demand for non-pharmacological interventions for children with childhood onset disorders. Cognitive training entails the repeated exercise of a specific cognitive process (or multiple processes) over an extended period of time typically several weeks after which performance gains are expected on the trained task, but more importantly on untrained tasks and/or behavioural measures (transfer). Performance gains on tasks similar but not identical to the training task are defined as ‘near transfer’ (e.g. training memory and improving on an untrained task measuring memory) whereas performance gains on dissimilar tasks and/or behavioral measures are defined as ‘far transfer’ (e.g. training on a memory task and improving on a mathematical task). This transfer of benefits to other cognitive skills/behaviour is the distinguishing element of cognitive training. However, evidence supporting cognitive training as an intervention that can produce transfer is mixed. The aim of the present review is to summarize and evaluate this heterogeneous evidence with a focus on its application in children with ADHD.

ADHD is a common, persistent and impairing disorder distinguished by developmentally inappropriate restlessness, inattention and impulsiveness (DSM-5). It is characterized by academic, behavioral and emotional problems in childhood and by increased risk for motor vehicle accidents, antisocial behavior and school dropout in adolescence (Barkley et al., 2006; Raggi & Chronis, 2006; Reinhart & Reinhardt, 2013; Küpper et al., 2012; Wilens, 2004). According to current theories, ADHD is a complex disorder with many genetic contributions, but also with contributory environmental risks (Vaidya & Stollstorff, 2008; Volkow et al., 2002). These underlying risks are not manifest in ADHD symptoms, but rather seem to disturb brain structure and function which in turn affects the higher order functions of the brain called executive functions (Crosbie, Pérusse, Barr, & Schachar, 2008). Executive functions manage other cognitive processes and are a collection of abilities that are related to one another but which are, to a great extent separable (Collette, Hogge, Salmon, & Van der Linden, 2006; Friedman et al., 2006). There are several reasons for the interest in cognitive training as an intervention for ADHD. First, cognitive training claims to directly address the cognitive deficits presumed to underlie ADHD (Castellanos, Sonuga-Barke, Milham, & Tannock, 2006). Second, preliminary evidence suggests that cognitive remediation might be at least partially effective in the treatment of ADHD (Johnstone et al., 2012; Shaley, Tsai, & Mevorach, 2007). And third, if effective, cognitive remediation would offer a non-drug alternative for a disorder that typically, and in very large numbers, involves the use of stimulant medication (Froehlich et al., 2007; MTA Cooperative Group, 1999).

Some studies that have used cognitive training have reported broad transfer effects (Johnstone et al., 2012; Shaley et al., 2007; Van der Molen, Van Luit, Van der Molen, & Jongmans, 2010), others have reported limited transfer effects (Gibson et al., 2011), and others still have failed to generate any transfer at all (Ackerman, Beier, & Boyle, 2005; Chacko et al., 2014; Healy, Wohldmann, Sutton, & Bourne, 2006; Lee et al., 2012; Owen et al., 2010). At this time, we do not know why some studies succeed in producing transfer and some do not. In the present review, we employ two approaches to unravel and understand this heterogeneity. First, there are several implicit assumptions that appear to guide researchers’ decisions regarding which cognitive abilities to target and which training tasks to use. We identify these assumptions as guiding principles that form the foundation for the current state of the literature on cognitive training and explain their weaknesses. Second, we summarize the studies both within the pediatric ADHD population and those outside of this population that have not produced any transfer effects in an effort to decipher the factors that may be responsible for this lack of transfer effects.

Assumptions in Cognitive Training

Several implicit assumptions drive researchers’ decisions regarding which cognitive abilities to train and which training tasks should be used to exercise those abilities. For example, a particular cognitive process might be targeted because it is presumed to be a higher-order function that predicts or influences a range of other cognitive processes. The rationale is that improving that particular process would lead to improvements of all skills under its influence (i.e. broad transfers). For the purposes of the present paper, we will refer to this as the “higher-order assumption”. Although performance on one measure of executive function might be significantly correlated with performance on a task measuring another executive function, there is little known about the specific nature of this relationship. For example, we know little about whether change in one process would induce change in another. The second implicit assumption is that the targeted ability is a central deficit in a particular disorder such as ADHD. We will refer to this as the “central-deficit assumption”. Even though specific cognitive deficits are found regularly in ADHD, the correlation between performance on cognitive tasks and symptom severity is typically moderate at best (McAuley, Chen, Goos, Schachar, & Crosbie, 2010) and the nature of the relationship is unknown. The third assumption relates to the training task and presumes that the training task will target the selected ability of interest. This assumption will be denoted as the “task-purity assumption”. These assumptions will be discussed in the context of the two most commonly targeted cognitive processes in this literature, working memory (WM) and attention.
**Implicit Assumptions and WM**

There are various definitions of WM. It has been described as the ability to maintain task relevant information for easy access during a task (storage capacity only) (Goldman-Rakic, 1995); as storage capacity + the processing of that information (Daneman & Carpenter, 1980; Engle, 2001); and as storage capacity + information retrieval from secondary memory if information-maintenance fails (Gibson, Gondoli, Flies, Dobrzenski, & Unsworth, 2009). Working memory is often the target of cognitive training because of its assumed capacity to influence a range of cognitive processes (Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Olesen, Westerberg, & Klingberg, 2004) sometimes referred to as predictive capacity. Thus, how strongly performance on a particular task (e.g. WM) can predict performance on a different task (e.g. reading comprehension) captures predictive capacity of that task. Empirical evidence supporting this view has been reported within the context of storage + processing and storage + retrieval (Baddeley, 2003; Unsworth & Engle, 2007a), but not when WM is defined as storage capacity-only. This distinction is important since many cognitive training studies justify targeting WM using the higher-order assumption, but use training tasks that exercised short-term memory (STM). STM is defined as the ability to maintain information for a short period of time and does not appear to have a strong predictive capacity (Engle, Tuholski, Laughlin, & Conway, 1999; Conway, Cowan, Bunting, Therriault, & Minkoff, 2002).

Complex span tasks are often used to measure WM (Daneman & Carpenter, 1980; Turner & Engle, 1989; Unsworth & Engle, 2007a). For example, in operation span task, participants are instructed to answer a mathematical equation and to remember the word at the end of the equation, such as “Does 2+1=3? (yes or no) HOUSE”. After a set of 2-7 trials, they are asked to reproduce the words presented at the end of the equations in the correct serial order (memory component). Solving the equation (processing) prevents rehearsal strategies from maintaining the list of words in primary memory and increases the probability that the to-be-remembered words will dissipate from primary memory, in which case retrieval from secondary memory will be required (retrieval) to access the list of words from the secondary memory (Unsworth & Engle, 2007a). Using complex span tasks, studies have shown an association between WM and reading comprehension (Daneman & Carpenter, 1980), language acquisition (Baddeley, 2003), fluid intelligence (the ability to reason and problem solve in novel situations) (Conway et al., 2002; Kane et al., 2004; Unsworth & Engle, 2007b), vocabulary learning, note taking, and reasoning (Engle, 2001). This association between performance on complex spans and performance on other cognitive tasks is viewed as the predictive capacity of complex spans. As stated earlier, this evidence is often used as the rationale for training WM (the higher-order assumption). However, many studies which have aimed to train WM used simple span tasks to target it (Beck, Hanson, Puffenberger, Benninger, & Benninger, 2010b; Klingberg, Forssberg, & Westerberg, 2002).

Simple span tasks are considered to be a measure of STM (Baddeley & Hitch, 1974; Engle, 2001; Unsworth & Engle, 2007b). For example, in forward digit span, participants are presented with a list of digits one at a time and required to repeat the list of digits in the correct order. Simple spans do not have a strong capacity to predict performance on other cognitive tasks (Conway et al., 2002; Daneman & Merikle, 1996; Engle et al., 1999; Kail & Hall, 2001) and are easily influenced by scoring procedures, presentation modality, and trial lengths (Unsworth & Engle, 2007b). This is not always the case with backward span tasks. In backward span tasks the participant must reproduce the list in backward serial order. The backward span is sometimes considered to measure STM (Engle et al., 1999; Swanson, Mink, & Bocian, 1999) and sometimes WM (Ackerman et al., 2005). Overall, empirical evidence depicts simple span tasks as having limited predictive capacity.

The central-deficit assumption pertains to the view that a cognitive process, such as WM or attention, is the key deficit in ADHD (Beck, Hanson, Puffenberger, Benninger, & Benninger, 2010a; Klingberg et al., 2002) and training it would remediate a range of behavioural symptoms associated with ADHD. However, current theories suggest that ADHD is a complex disorder (Brown, 2006). Many executive functions seem to be perturbed in ADHD and each child can present with a distinct profile of deficits, some without working memory impairments (Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005; Qian, Shuai, Chan, Qian, & Wang, 2013). Presently, we do not have a clear knowledge of the hierarchical organization of cognitive processes implicated in ADHD. Furthermore, as stated earlier, the association between cognitive performance and symptom severity is not strong (Coghill, Hayward, Rhodes, Grimmer, & Matthews, 2013; McAuley et al., 2010). Thus, the view that remediating WM deficit would alleviate a range of symptoms in ADHD has not yet received careful examination.

Lastly, there is an assumption that a selected training task will target the ability of interest. It is difficult to measure a specific cognitive process with laboratory tasks because of the task impurity problem (Burgess, 1997). Any given task of executive function will involve a range of abilities since executive functions by definition regulate other cognitive processes. Although it is not necessary to have a pure task for training purposes, it is important to know what a training task measures in order to correctly target and train the ability of interest. For example, if we intend to train WM, but use a task that predominately targets STM, then we will most likely not train WM. With these issues highlighted, the next section will summarize the findings in this field.
**Training Working Memory**

Many of the studies that have trained working memory have used the Cogmed Working Memory Training (2006) software (Table 1). Cogmed consists of verbal and visuo-spatial forward and backward simple span tasks. Studies which have used Cogmed (Table 1) have been criticized for design limitations that did not control for placebo effects. These limitations make it difficult to attribute the reported transfers to the training program unambiguously rather than to simple practice effects or the passage of time (Shipstead, Redick, & Engle, 2012b). Given that methodological weaknesses have been addressed in previous review papers (Shipstead et al., 2012b), we will only focus on theoretical weaknesses in our review.

Studies on Cogmed have reported transfer to tasks that measured the same ability (storage capacity) and some transfer to attention and organizational behaviour based on parent and teacher ratings. There are three well-controlled studies which have used Cogmed as a training program (Chacko et al., 2014; Gray et al., 2012; Green & King, 1998), two of which did not produce transfer effects and will be discussed in a later section. Green et al. (2012) was a well-controlled and double-blind study. They reported significant reductions in “off-task” behaviour (far transfer), albeit with a small effect size (Rapport, Orban, Kofler, & Friedman, 2013) and significant improvements on storage capacity (near transfer) in their training group. However, they did not find any significant differences in parent reports of problem behavior. In other words, they were not able to improve problem behaviour associated with ADHD. A meta-analytic review (Melby-Lervåg & Hulme, 2013) as well as a systematic review (Shipstead, Hicks, & Engle, 2012a) of studies which have used Cogmed training program and other WM training programs (Melby-Lervåg & Hulme, 2013) have revealed short-term near transfer effects only. These findings show limited transfer to other cognitive skills and limited effects on behavior, highlighting the weakness of higher-order and central-deficit assumptions. The limited transfer effects generated by Cogmed working memory training may stem from the failure of simple span tasks to target WM (task-purity assumption) and/or the limited capacity to influence other cognitive processes (predictive capacity). Empirical evidence appears to point to the former notion demonstrating that Cogmed targets short-term memory and not WM (Rapport et al., 2013).

Several studies have used complex span tasks to train WM in children and adolescents and have been able to produce broader transfer effects (Alloway & Alloway, 2008; Loosli, Buschkuehl, Perrig, & Jaeggi, 2012; Van der Molen et al., 2010). One commercially available program which includes complex span tasks is Jungle Memory designed by Alloway and Alloway (2008). Alloway (2012) used Jungle Memory to train WM in adolescents with learning difficulties and reported improvements in working memory capacity (near transfer), vocabulary (far transfer) and mathematical abilities (far transfer). Other researchers who have used complex span tasks to train WM have reported transfer effects to reading ability in healthy children (Loosli et al., 2012) and scholastic abilities in adolescents with mild to borderline intellectual disability (Van der Molen et al., 2010). These findings suggest broader transfers than those produced with simple span tasks corroborating the view that complex spans have greater predictive capacity (Daneman & Carpenter, 1980; Engle et al., 1999; Unsworth & Engle, 2007b) and are perhaps better suited to exercise WM and/or produce transfer effects. However, a meta-analysis of WM training programs which included Jungle memory showed that these training programs were only successful in producing short-term near transfer effects (Melby-Lervåg & Hulme, 2013). Thus further research is required to investigate the effectiveness of complex span tasks as training tasks before conclusive statements can be construed.

**Training Attention**

Attention is broadly defined as the ability to concentrate on a selected aspect of the environment. The type of attention targeted by cognitive training researchers is varied. Four attentional functions have been described, each associated with a separate neural correlates (Posner & Petersen, 1990; Tsai, Shalev, & Mevorach, 2005). These include orienting of attention (directing attention to a specific stimulus in the sensory environment), selective attention (selection of the relevant information from sensory input), sustained attention (the ability to sustain attention over time) and executive attention (the ability to divide or to alternate/shift attention between two tasks).

Studies which have trained attention appear to have based their selection on the central-deficit assumption with the approach that inattentiveness is the key impairment in ADHD and remediating it would alleviate cognitive and behavioral difficulties associated with ADHD. There are two commercially available programs (“Captain’s Log” and “Pay Attention”) that train various attentional networks in addition to other cognitive processes. Several studies have used Captain’s Log as a training program for children with ADHD and have reported improvements on the Continuous Performance Test (CPT) and transfer to parent/teacher rating scales (Table 1). Slate et al. (Slate, Meyer, Burns, & Montgomery, 1998) reported additional transfer effects to math and vocabulary but they did not run formal statistical tests, it is therefore difficult to consider their findings to be significant training-induced improvements.

“Pay Attention” is a paper and pencil program which targets sustained attention, selective attention and alternating/divided attention. Kerns et al. (Kerns, Eso, & Thomson, 1999) used “Pay Attention” to train children with ADHD (7-14 years) and reported broad transfer effects (e.g. improved problem solving and mathematical ability). However, parent reports of inattention-impulsivity and hyperactivity were not significantly different from those described in...
<table>
<thead>
<tr>
<th>Author et al. (year)</th>
<th>Subject Group (age)</th>
<th>Training Method</th>
<th>Sessions</th>
<th>Session Duration (min)</th>
<th>Control Group (n=?)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beek et al. (2010)</td>
<td>ADHD (7-17)</td>
<td>Cogmed (n=27)</td>
<td>25</td>
<td>30-40</td>
<td>Waitlist control (n=24)</td>
<td>P-BRIEF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T-BRIEF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CPRS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DSM-IV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P-BRIEF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T-BRIEF</td>
</tr>
<tr>
<td>Chacko et al. (2014)</td>
<td>ADHD (7-11)</td>
<td>Cogmed (n=44)</td>
<td>25</td>
<td>30-45</td>
<td>Non-adaptive (n=11)</td>
<td>AWMA: Dot Matrix; AWMA: Digit Recall; AWMA: Spatial Recall; AWMA: Listening Recall</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DBD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A-X OPT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WRAAT 4-PMV</td>
</tr>
<tr>
<td>Gibson et al. (2011)</td>
<td>ADHD (11-16)</td>
<td>Cogmed (n=44)</td>
<td>20-25</td>
<td>No controls</td>
<td></td>
<td>IFR Tasks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DuPaul ADHD-RS (parent and teacher)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DuPaul ADHD-RS (parent and teacher)</td>
</tr>
<tr>
<td>Gray et al. (2012)</td>
<td>LD+ADHD (12-17)</td>
<td>Cogmed (n=32)</td>
<td>20-25</td>
<td>45</td>
<td>Math Training (n=20)</td>
<td>WISC-DSF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SWAN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IOWA-CRS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D2 TA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WRA-T-4PM</td>
</tr>
<tr>
<td>Green et al. (2012)</td>
<td>ADHD (7-14)</td>
<td>Cogmed (n=12)</td>
<td>25</td>
<td>40</td>
<td>Non-adaptive (n=14)</td>
<td>WISC-DSF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WRAML-FW</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WAIS-SPF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WAIS-SBB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RAVEN</td>
</tr>
<tr>
<td>Dahlin et al. (2011)</td>
<td>Special needs (9-12)</td>
<td>Cogmed (n=42)</td>
<td>25</td>
<td>30-40</td>
<td>Controls from Klingberg et al. (2005)</td>
<td>WISC-DSF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WISC-DSB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Stoop</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WAIS-SPF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WAIS-SBB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Stoop</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RAVEN</td>
</tr>
<tr>
<td>Olesen et al. (exp1)</td>
<td>Healthy (20-23)</td>
<td>WM (n=3)</td>
<td>20-30</td>
<td>35-45</td>
<td>No training control (n=11)</td>
<td>WAIS-SB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Stoop</td>
</tr>
</tbody>
</table>

Table 1 continued
Table 1. continued A Review of cognitive training studies for children, adolescent, and young adults

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Subject Group (age)</th>
<th>Training Method (n?)</th>
<th>Training Schedule (sessions)</th>
<th>Sess Dur (min)</th>
<th>Control Group (n?)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olesen et al. (exp2)</td>
<td>2004</td>
<td>Healthy (27-31)</td>
<td>Cogmed (n=8)</td>
<td>15-21</td>
<td></td>
<td>No training control (n=11)</td>
<td>WAIS-DS/WAIS-III</td>
</tr>
<tr>
<td>Holmes et al.</td>
<td>2009</td>
<td>Healthy (9-11)</td>
<td>Cogmed (n=22)</td>
<td>20</td>
<td>35</td>
<td>Non-adaptive (n=20)</td>
<td>AWMA</td>
</tr>
<tr>
<td>Looss et al.</td>
<td>2012</td>
<td>Healthy (9-11)</td>
<td>WM training (n=20)</td>
<td>10</td>
<td>12</td>
<td>No training (n=40)</td>
<td></td>
</tr>
<tr>
<td>Alloway</td>
<td>2012</td>
<td>Learning difficulties (11-15)</td>
<td>Jungle Memory (n=8)</td>
<td>24</td>
<td>30</td>
<td>Targeted education support (n=7)</td>
<td>AWMA</td>
</tr>
<tr>
<td>Van der Molen et al.</td>
<td>2010</td>
<td>Learning disabilities</td>
<td>Odd Yellow (n=41)</td>
<td>10</td>
<td>45</td>
<td>Social training (n=10)</td>
<td></td>
</tr>
<tr>
<td>Hoekzema et al.</td>
<td>2010</td>
<td>ADHD (8-14)</td>
<td>WM, Attention, cognitive flexibility, planning, problem solving (n=9)</td>
<td>10</td>
<td>45</td>
<td>Social training (n=10)</td>
<td>GNG</td>
</tr>
<tr>
<td>Johnstone et al.</td>
<td>2010</td>
<td>ADHD (7-13)</td>
<td>WM, RI (n=18)</td>
<td>25</td>
<td>20</td>
<td>Low intensity (n=18)</td>
<td>GNG</td>
</tr>
<tr>
<td>Johnstone et al.</td>
<td>2012</td>
<td>ADHD (7-13)</td>
<td>IC, WM (n=60)</td>
<td>25</td>
<td>15-20</td>
<td>Healthy (n=68)</td>
<td>GNG</td>
</tr>
<tr>
<td>Rabiner et al.</td>
<td>2010</td>
<td>ADHD (6-7)</td>
<td>Attention, reading, math (n=52)</td>
<td>28</td>
<td>75</td>
<td>Waitlist (n=25)</td>
<td>CTRT, BRIEF-P, BRIEF-III</td>
</tr>
<tr>
<td>Steiner et al.</td>
<td>2011</td>
<td>ADHD (11-14)</td>
<td>Attention, WM (n=13)</td>
<td>24</td>
<td>45</td>
<td>Waitlist (n=15)</td>
<td>CPRS, BASC-P, BRIEF-P, CPT</td>
</tr>
</tbody>
</table>

continued
### Table 1. A Review of cognitive training studies for children, adolescent, and young adults

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Subject Group (age)</th>
<th>Training Method (n?)</th>
<th>Training Schedule (sessions)</th>
<th>Sess Dur (min)</th>
<th>Control Group (n?)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kerns et al.</td>
<td>1999</td>
<td>ADHD (7-14)</td>
<td>Attention (n=14)</td>
<td>16</td>
<td>30</td>
<td>Other edu. games</td>
<td>Digit Span Coding, MFFT ADDES Mental (WISC-III) ACT UT Math Work-sheets</td>
</tr>
<tr>
<td>Shaive et al.</td>
<td>2007</td>
<td>ADHD (6-13)</td>
<td>Attention (n=20)</td>
<td>Other video games (n=16)</td>
<td></td>
<td></td>
<td>PRS-Inattention PRS-Hyperactivity VOT (WISC-III) COPT Passage copying, reading</td>
</tr>
<tr>
<td>Kotwal et al.</td>
<td>1996</td>
<td>ADHD (13)</td>
<td>Captain’s Log (n=1)</td>
<td>Single-case study</td>
<td></td>
<td></td>
<td>No statistical Tests: CTRS-on-task behaviour CTRS</td>
</tr>
<tr>
<td>Rabiner et al.</td>
<td>2010</td>
<td>ADHD (6-7)</td>
<td>Captain’s Log (n=25)</td>
<td>Washlist (n=3)</td>
<td>28</td>
<td>75</td>
<td>CTRS–Inattention CTRS–Hyperactive-Impulsive, Oppositional, Social Problems, Anxious/Shy</td>
</tr>
<tr>
<td>Slate et al.</td>
<td>1998</td>
<td>ADHD+severely emotionally disturbed (7-11)</td>
<td>Captain’s Log (n=4)</td>
<td>No Control</td>
<td>64</td>
<td>30</td>
<td>No statistical Tests: CTRSP 4/4 children – impulse control, hyperactivity TRF (3/4 children)– attention CTRS CBCL</td>
</tr>
</tbody>
</table>

Note: 
- “=” indicates that the cognitive process was examined but training related improvements were not found.
- “#” indicates that the cognitive process was examined and training related improvements were reported either at post-training session or at the follow-up session subsequent to training.
- “not examined” indicates that the cognitive process was not examined.

The table includes the following measures:
- **Working Memory**
- **Response Inhibition**
- **Behaviour**
- **Planning/Organization/Reasoning**
- **Attention**
- **Scholastic Abilities**
- **Intelligence/Fluid Intelligence**

The measures used are:
- **WJ-III APRS**
- **WRAT=3–math (3/4 children)**
- **PPVT-R (3/4 children)–receptive vocabulary**
- **CBCL**
- **WRAT-4PMV (Wide-Range Achievement Test-4-Progress Monitoring Version)**
- **VOT = WISC-III Hooper Visual Organization Test**
- **VMR = Working Memory Rating Scale**
- **WOND = Wechsler Objective Number Dimensions**
- **WORD = Wechsler Objective**
- **WRAT-4PMV = Wide-Range Achievement Test-4-Progress Monitoring Version**

Additional measures include:
- **TMT–B (3/4 children)**
- **WRAT=3–math (3/4 children)**
- **PPVT-R (3/4 children)–receptive vocabulary**
- **CBCL**
- **WRAT-4PMV (Wide-Range Achievement Test-4-Progress Monitoring Version)**
- **VOT = WISC-III Hooper Visual Organization Test**
- **VMR = Working Memory Rating Scale**
- **WOND = Wechsler Objective Number Dimensions**
- **WORD = Wechsler Objective**
- **WRAT-4PMV = Wide-Range Achievement Test-4-Progress Monitoring Version**

The assessments used in the studies include:
- **Kerns et al. 1999**
- **Shalib et al. 2007**
- **Kotwal et al. 1996**
- **Rabiner et al. 2010**
- **Slate et al. 1998**
the pre-training phase. In contrast, Shalev et al. (2007) used neuropsychological tasks to train sustained attention, selective attention, orienting attention and executive attention in children 6-13 years with ADHD and reported improvements in reading comprehension, passage copying and a reduction of parent reported inattentiveness. Overall, these findings demonstrate improvements on the targeted ability and limited transfer to other cognitive skills. Although these results are promising, transfer still remains narrow in most studies. This highlights the limitations of the central-deficit assumption.

**Lack of Transfer**

Studies which have not produced any transfer effects are more prevalent in adults than in children. This may be due to publication bias or it may be that cognitive training is more successful in children. The only published studies that have not shown any transfer effects in children are the ones that examined the efficacy of Cogmed. Chacko et al. (2014) examined effectiveness of Cogmed in children 7-11 years of age with ADHD. They found that the active training group improved on WM storage tasks but not on WM tasks measuring storage + processing/manipulation. Similarly, they did not observe transfer to any other measures assessing academic ability, attention, or parent/teacher rating scales (Table 1). Gray et al. (2012) used Cogmed to train adolescents 12-17 years of age with learning disability and ADHD. They reported gains on two tasks measuring WM but not on other WM tasks or any of their transfer tasks (for details see Table 1). These findings emphasize the limitations of the central and higher-order assumptions and the view that simple span tasks are not qualified to produce broad transfers. A series of work by Gibson et al. (Gibson et al., 2011; Gibson et al., 2013; Gibson et al., 2009) indicate that a modification of the adaptive algorithm of simple span tasks may increase their effectiveness as training tasks. Gibson and colleagues describe a dual-component model of WM where task relevant information is stored in the primary memory (PM); but upon loss of information from PM, attempts are made for the relevant information to be recalled from secondary memory (SM). Gibson et al. (2009) demonstrate that actively retaining information in PM is not impaired in ADHD; whereas retrieval of task relevant information from SM is impaired. In their later work Gibson et al. (2011) reveal that Cogmed enhances the maintenance of information in PM, a component of WM which is not affected in ADHD. Gibson and colleagues suggest that the adaptive algorithm of Cogmed for advancing to the next difficulty level is designed in such a way that the tasks currently tax the resources of the PM. They propose a modification to this algorithm to shift the demand to SM component of WM and confirm (Gibson et al., 2013) that by changing this adaptive algorithm, task demand does get shifted to the SM. Although it is plausible that such a modification would improve effectiveness of Cogmed training program or simple span tasks in general, this hypothesis must be examined and verified. In other words, it remains to be determined whether such an alteration would produce broader transfer effects in general or whether it would mitigate symptoms of ADHD more effectively.

The other studies that have not been successful in producing transfer are in adult samples. Lee et al. (2012) trained healthy young adults on a complex video-game (Space Fortress; Donchin, 1989). Space Fortress engages cognitive processes such as working memory, resource management and manual control. It requires players to navigate their ship and destroy a Space Fortress (for a full description of the game see Donchin, 1989). These investigators did not find transfer to untrained tasks measuring memory, attention, visual processing, motor control, reasoning ability and dual-tasking ability. Healy et al. (2006) had similar findings using a different training apparatus. They trained healthy young adults on a perceptual-motor task and aimed to improve inhibitory control. Their rational was that since in several of their conditions, participants had to inhibit prepotent responses (having a pre-existing probability of occurring) transfer should take place to conditions that included the same inhibitory demand. The investigators did not find transfer to other conditions.

One plausible explanation for lack of transfer in the above two studies may be that manual control and visuo-spatial skills can improve over time making it easier to perform the tasks and potentially decrease the cognitive load. When a given training task allows for task specific strategies to develop, it may lose its novelty and performance may become automatic (for further details see Morrison & Chein, 2011). Imaging findings support this view as well (Olesen et al., 2004). It is generally believed that an increase in neural activation is observed at first when the participant is required to carry out novel tasks. However as training continues, performance on the task becomes more automated and does not require as many cognitive resources leading to a decrease in neural activation (see Buschkuehl et al., 2012 for a review). In view of these findings, it may be wise to minimize development of task-specific strategies and strive to maintain high cognitive loads when designing training tasks. Owen et al. (2010) trained healthy adults 18-60 years of age for six weeks on an online brain training program. Participants were required to train for at least ten minutes a day, three times a week for 23-28 training sessions. They randomly assigned participants to two experimental groups and a control group. The training tasks consisted of different combinations of tasks targeting reasoning, planning, problem-solving, STM, attention, visuospatial processing and mathematics. The control group answered abstract questions from six different categories using online resources. Owen et al. found no evidence of transfer and no significant differences between their groups. Ackerman et al. (Ackerman, Kanfer, & Calderwood, 2010) trained healthy adults...
Discussion and Conclusion

The theoretical weaknesses associated with this field cluster around the implicit assumptions summarized earlier. These assumptions require consideration and discussion. Empirical evidence thus far does not appear to support the central-deficit and the higher-order assumptions. In fact, research in the neuroscience and mental health fields suggest a range of cognitive deficits associated with ADHD (Nigg, Blaskey, Stawicki, & Sache, 2004; Nigg et al., 2005) and these cognitive deficits may be similar to those of other childhood onset disorders (Lipszyc & Schachar, 2010). Research is detaching itself from a central-deficit approach and shifting towards multiple-deficit models to capture the heterogeneity of ADHD (Castellanos & Tannock, 2002; Castellanos et al., 2006; Nigg et al., 2005). However, for researchers/clinicians interested in training WM, they may consider complex span tasks or use Gibson et al. (2013) algorithm for adaptation purposes in order to target WM and produce broader transfers.

The task-purity assumption will likely be an issue to tackle when selecting the training tasks and it becomes of greater concern when the training task is shown to measure different processes (e.g. STM vs WM). One approach to bypass this problem may be the same one preferred by many studies examining latent variables associated with a set of executive functions. A latent variable is an ability that is correlated with performance on a set of cognitive tasks. Perhaps in order to successfully target an ability of interest, a collection of tasks known to measure that ability should be selected to train it.

Taken as a whole, research in this field has demonstrated that cognitive processes which were once believed to be hard-wired are in fact trainable and modifiable. Future research aimed at identifying moderator factors that influence transfer effects as well the mechanism underlying transfer will be vital to this field. This line of research together with the objective of unraveling the plastic capacity of the human brain will have far-reaching implications.

Acknowledgments/Conflicts of Interest

Dr. Russell Schachar is a consultant to and has equity in a company that develops cognitive rehabilitation software. Research reported in this paper was supported by the Ontario Brain Institute (OBI-FedDev) and Behavioural Neurological and Applied Solutions (BNAS). The funders had no role in review, decision to publish, or preparation of the manuscript.

References


