Routledge
Taylor & Francis Group

Applying cognitive training to target executive functions during early development

Sam V. Wass

Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK

Developmental psychopathology is increasingly recognizing the importance of distinguishing *causal processes* (i.e., the mechanisms that cause a disease) from *developmental outcomes* (i.e., the symptoms of the disorder as it is eventually diagnosed). Targeting causal processes early in disordered development may be more effective than waiting until outcomes are established and then trying to reverse the pathogenic process. In this review, I evaluate evidence suggesting that neural and behavioral plasticity may be greatest at very early stages of development. I also describe correlational evidence suggesting that, across a number of conditions, early emerging individual differences in attentional control and working memory may play a role in mediating later-developing differences in academic and other forms of learning. I review the currently small number of studies that applied direct and indirect cognitive training targeted at young individuals and discuss methodological challenges associated with targeting this age group. I also discuss a number of ways in which early, targeted cognitive training may be used to help us understand the developmental mechanisms subserving typical and atypical cognitive development.

Keywords: Cognitive training; Attentional control; Working memory; Infant; Toddler; Early intervention; At-risk; Preventative intervention.

A number of authors in recent years have advocated the desirability of early interventions (Bryck & Fisher, 2012; Heckman, 2006; Shonkoff & Levitt, 2010; Sonuga-Barke & Halperin, 2011). Several studies have suggested, for example, that teacher- and parent-mediated interventions providing increased social and educational provision for young children from low socioeconomic status backgrounds are more effective the earlier the training is applied (Campbell et al., 2008; Olds, Sadler, & Kitzman, 2007). Similarly, clinician-, parent-, and teacher-mediated programs are currently being set up to assess the impact of intervening early in disrupted development for individuals at high risk of developing conditions such as attention deficit/hyperactivity disorder (ADHD; Sonuga-Barke & Halperin, 2011) and Autism Spectrum Disorders (ASD; Wallace & Rogers, 2010).

Parallel to clinician-, parent-, and teacher-mediated interventions (that tend to be cognitively heterogeneous in nature), a separate research field exists that examines the effect of applying cognitive training targeted at particular, prespecified cognitive domains. Within this field, there appears to be surprisingly little appreciation of the importance of the developmental perspective. Cognitive training is administered to individuals throughout the lifespan with the majority of new studies in this area targeting older individuals (e.g., Brehmer et al., 2012; Richmond, Morrison, Chein, & Olson, 2011; Wang, Chang, & Su, 2011). Of studies that have applied cognitive training to children, the majority have targeted children in the 8–12 years age range (Holmes, Gathercole, & Dunning, 2009; Klingberg et al., 2005), with only a small number of studies targeting children of a younger age (e.g., 4–6 years: Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005; 4 years: Thorell, Lindqvist, Nutley, Bohlin, & Klingberg, 2009) or during infancy (Wass, Porayska-Pomsta, & Johnson, 2011).

In Section 1 of this review, I outline a priori arguments suggesting the importance of understanding the very early stages of development in both typical and atypical populations. In Section 2, I evaluate evidence from longitudinal studies of high-risk and atypical populations, including individuals born preterm, with genetic disorders, from low socioeconomic-status backgrounds, and at familial risk of ASD and ADHD. I conclude that, across a number of populations, there is correlational evidence suggesting that early emerging deficits in attentional control and working memory may be important in mediating later-emerging deficits in other areas. In Section 3, I review the currently small number of studies that have applied cognitive training to young individuals and discuss methodological issues involved in applying targeted training at these populations. I discuss a number of avenues for future work in this area.

WHY IS EARLY DEVELOPMENT IMPORTANT?

Evidence from neuroimaging, computational modelling, and animal studies is increasingly revealing neural development as a dynamic and interactive process (Johnson, 2010; Quartz & Sejnowski, 1997). For example, research investigating the effect of brain lesions early in development suggests that early disruption can either be compensated for (Stiles, Reilly, Paul, & Moses, 2005) or can lead to cascade-like patterns of systemic disruption across other, nonlesioned parts of the system due to the disruption of normal interactive maturational processes (Johnson, Halit, Grice, & Karmiloff-Smith, 2002; Spencer-Smith et al., 2011).

A similar picture has emerged from work using functional imaging. Early in development, functional cortical activation patterns are relatively unlocalized and undifferentiated: Specific tasks evoke larger functional activation patterns, and cortical areas are relatively less specialized (Bell & Wolfe, 2007; Cohen Kadosh & Johnson, 2007; Redcay, Haist, & Courchesne, 2008). Neural maturation involves the increasing localization and specialization of neural circuitry (Durston et al., 2006; Fair et al., 2008, 2010). Research using neuroimaging and computational modelling has suggested that these processes arise, at least in part, as the emergent property of competition and cooperation between brain areas (Johnson, 2010; Kelly et al., 2009); atypical development shows activation patterns becoming progressively more abnormal over developmental time due to the disruption of normal maturational processes (Johnson et al., 2002; Oliver, Johnson, M. H., Karmiloff-Smith, A., & Pennington, 2000; however, see Shaw et al., 2008).

Over recent years similar arguments have also been advanced in favor of studying how behavior develops over time at the systemic level (see Smith & Sheya, 2011). These approaches emphasize the importance of studying not just the end-state of cognition but

also the developmental pathways by which the end-state has been arrived at (Cornish, Sudhalter, & Turk, 2004; Cornish, Scerif, & Karmiloff-Smith, 2007; Karmiloff-Smith, 1998, 2007). Thus, rather than viewing disorders in terms of static neuropsychological deficits ("intact" vs. "impaired" cognitive modules), we should instead seek to develop longitudinal, developmentally plausible models of disease causation (Karmiloff-Smith, 1998, 2009). For example, research with individuals with Williams syndrome has suggested that early developing atypicalities in eye movement control at the microtemporal (subsecond) scale may lead to subsequently impaired learning across other domains including social communication and number perception (Brown et al., 2003; Karmiloff-Smith et al., 2012). Research in typical development is similarly suggesting that many tools for early learning, such as gaze following and other forms of joint attention, may emerge as learnt behaviors, with later attainments building on foundations that are laid down early in development (Corkum & Moore, 1998; Mareschal et al., 2007; Triesch, Teuscher, Deak, & Carlson, 2006). This suggests the vital importance of researching the very early stages of cognitive development.

THE SPECIAL ROLE OF ATTENTIONAL CONTROL/WORKING MEMORY IN MEDIATING EARLY LEARNING

Amongst these dynamic approaches to studying development, two particular cognitive faculties have received particular attention: These are attentional control, defined as "an individual's ability to choose what they pay attention to and what they ignore" and working memory, the "maintenance of task-relevant information in mind for brief periods of time to guide behaviour" (Gazzaley & Nobre, 2012, p. 129). These two faculties are thought to have substantially overlapping neural correlates (Duncan & Owen, 2000; Munakata et al., 2011), particularly early in development (Astle & Scerif, 2009; Scherf, Longhi, Cole, Karmiloff-Smith, & Cornish, 2006; Shing, Lindenberger, Diamond, Li, & Davidson, 2010; Velanova, Wheeler, & Luna, 2008). Attentional control in particular has been discussed as a "hub" cognitive domain, gating subsequent skill acquisition in other areas (Cornish, Cole, et al., 2012; Cornish, Scerif, & Karmiloff-Smith, 2012; Scerif, 2010). The ability to regulate and direct attention releases a child from the constraints of only responding to environmental events and means they are able actively to guide their attention toward the information-rich areas key for learning (Ruff & Rothbart, 1996; Scerif, 2010).

Longitudinal neuroimaging studies suggest that cortical maturation follows a non-uniform trajectory, with certain areas (occipital, parietal) becoming relatively mature at an age when other areas (frontal) are relatively immature (e.g., Gogtay et al., 2004). Similarly, behavioral research has suggested that attentional control and working memory are relatively late-maturing relative to other cognitive faculties (e.g., Davidson, Amso, Anderson, & Diamond, 2006; Johnson, 2010). Some researchers have even suggested that these faculties may be absent during the first year of life and only begin to emerge at around the 12-month age range (Colombo & Cheatham, 2006; however, see Gilmore & Johnson, 1995; Johnson, 1995; Johnson, Posner, & Rothbart, 1991).

Rose, Feldman, and Jankowski, (2012) administered a battery assessing memory, processing speed, and attention in a cohort of individuals at 7, 12, 24, and 36 months and, in the same individuals, measured working memory, inhibition, and shifting at 11 years. They found that memory when assessed during infancy and toddlerhood predicted working memory performance at 11 years; they also found that processing speed (psychomotor reaction time) predicted performance on assessments of shifting and working memory at 11 years (Rose et al., 2012; see also Rose, Feldman, Jankowski, & Van Rossem, 2008).

Rose, Feldman, and Jankowski, (2009) administered a battery of nonverbal "information processing" assessments (including memory, attention, processing speed, and representational competence) in typically developing infants at 12 months and assessed language in the same individuals at 12 and 36 months. They found that some (but not all) of their information-processing measures (memory and representational competence, but not attention and processing speed) correlated with language performance at 12 months and predicted subsequent language performance at 36 months, independent of birth status (see also Dixon & Smith, 2008; Kannass & Oakes, 2008; Snyder & Munakata, 2011). Comparable findings have been reported in infants and toddlers with Autism Spectrum Disorders (Bopp, Mirenda, & Zumbo, 2009), as well as using similar longitudinal tracking studies with older children (Dice & Schwanenflugel, 2012; Gathercole, Alloway, Willis, & Adams, 2006; Kegel & Bus, 2012).

A number of groups have also used techniques such as Structural Equation Modelling (SEM) to explore possible mediators between early development and longer term cognitive outcomes in clinical populations, such as infants born preterm (Rose, Feldman, & Jankowski, 2005; Voigt, Pietz, Pauen, Kliegel, & Reuner, 2012; Weindrich, Jennen-Steinmetz, Laucht, & Schmidt, 2003). Voigt and colleagues found that deficits in effortful control (assessed using a behavioral battery and the Early Child Behavior Questionnaire at 24 months) partially mediated deficits in other cognitive outcomes in early preterm but not in late preterm infants (Voigt et al., 2012). Rose and colleagues administered a battery of assessments to preterm and full-term infants at 7 months, 12 months, and 2–3 years and identified two "elementary" factors that appeared to mediate the more complex factors (Rose et al., 2008). The first of these elementary factors was labeled "attention" (defined from peak look duration and shift rate during a habituation task) and the second was "speed" (defined as the number of trials required to reach criterion in a face familiarization test). Subsequent work from this group has tracked individuals through to 11 years; SEM conducted on these data suggested a cascade of effects, in which prematurity influences processed speed, which then influences executive function (EF), which in turn influences academic achievement (Rose, Feldman, & Jankowski, 2011; Feldman, & Jankowski, et al., 2011).

Research has also suggested that early developing deficits in attentional control/working memory (WM) may play a role in disrupting learning in individuals with genetic disorders such as Williams Syndrome (WS), Down Syndrome (DS), and Fragile X Syndrome (FXS). Cornish and colleagues administered assessments of attentional control on three occasions over 24 months to a group of 4- to 10-year-old individuals with FXS syndrome, as well as tracking the development of autistic symptomatology, hyperactivity/inattention, and other nonverbal cognitive indices. They found that attentional markers in the visual and auditory modality correlated longitudinally with later assessments of intellectual abilities and classroom behavior, whereas auditory markers correlated longitudinally with later autistic symptomatology (Cornish, Cole, et al., 2012; Cornish, Scerif, et al., 2012; Scerif, Longhi, Cole, Karmiloff-Smith, & Cornish, 2012). In earlier work, the same group has also identified early-developing abnormalities in attentional control in 3- to 55-month-old individuals with FXS and WS and documented differences in the developmental trajectories of the deficits observed across different conditions (Cornish et al., 2007; see also Breckenridge, Atkinson, & Braddick, 2012; Brown et al., 2003; Scerif, Cornish, Wilding, Driver, & Karmiloff-Smith, 2004).

Due to the problems inherent in identifying prediagnosis individuals, similar investigations into pathogenic mechanisms in the early development of ADHD and ASD are comparatively more limited. Lawson and Ruff (2004) found that ratings of focused attention in 7-month-old infants during free play with toys correlated with maternal ratings on ADHD rating scales at 4–5 years, as well as with cognitive abilities at 2, 3, and 4–5 years (see also Auerbach,

Atzaba-Poria, Berger, & Landau, 2004; Friedman, Watamura, & Robertson, 2005; Nigg, 2006). Holmboe and colleagues administered a task in which 9- to 10-month-old infants at high familial risk of ASD were required selectively to inhibit their looks to a peripherally occurring distractor and found that a subset showed difficulty disengaging attention, as well as less selective inhibition (Holmboe et al., 2010). Webb and colleagues reported group differences in 18- to 30-month-olds with more severe ASD symptoms during a habituation protocol (longer peak look and more time required to habituate) that were present for social and nonsocial stimuli but markedly stronger for social stimuli (Webb et al., 2010). Several groups have also reported problems with disengaging visual attention under competition but not noncompetition conditions in individuals with or at risk of ASD (Elsabbagh et al., 2009; Landry & Bryson, 2004), although these findings are not reported universally and appear to be contingent on the exact nature of the visual stimulus that is used (Chawarska, Volkmar, & Klin, 2010; Kikuchi et al., 2011). Systematic longitudinal mediation studies in this area are, however, lacking.

Summary—The Shortcomings of Correlational Findings

Across a range of disorders within both typical and atypical development, research has suggested that individual differences in early AC/WM, along with the related domain of processing speed, correlate with subsequent learning abilities in a range of different domains. These findings suggest that these domains may play a role in mediating subsequent learning. However, it is vital to recognize that all of the findings reported above are correlational and, therefore, are insufficient demonstrations of causal relationships. Even techniques such as Structural Equation Modelling are vulnerable to the possibility that confounding variables have not been included in the model.

Willoughby and colleagues, for example, examined the well-replicated finding that the performance of older children on EF tasks relates to later academic learning (Willoughby, Kupersmidt, Voegler-Lee, & Bryant, 2011). Their analyses replicated the commonly found relationship, even after including an earlier measure of academic achievement as a covariate; however, when they used a different technique, fixed effects analysis, which capitalizes on repeated measures data to control for time stable measured and unmeasured covariates, the observed relationships disappeared. The authors interpreted this as suggesting that the well-replicated association between EF abilities and academic achievement may be spurious (Willoughby et al., 2011). A conclusive investigation of how two domains are causally linked requires an experimental study to establish a counterfactual dependence between two: If we can demonstrate how training "x" improves "y," then we have taken a significant step toward demonstrating how "x" is causally implicated in "y." Thus, in addition to applied goals, training techniques have considerable potential to address questions motivated by "basic science" of how interactions subsist between cognitive domains over developmental time.

APPLYING COGNITIVE TRAINING DURING EARLY DEVELOPMENT

Despite the evidence reviewed above, only a small number of studies have provided targeted cognitive training aimed at individuals early in development (see Table 1).

Researchers working with infants face a unique problem of identifying a means by which the individual can interact with a computerized training paradigm, since fine motor skills are poor in this age range (Aslin, 2007). One solution is to use eye-gaze control as the means by which the infant interacts with the training by using eyetrackers to design

Table 1 Summary of training studies included in the review.

Authors	Year	Description of participants	Age of participants	Nature of training	Amount of training	Control	N trained	Pre- and Posttests
Wass et al.	2011	Typically developing (TD)	11-month-olds	Mixed Attention/ WM (eye-gaze contingent)	4 training sessions (variable length)—average of 77 ($SD = 19.1$) mins training administered in total	Watched infant-friendly animations and videos for a matched program of sessions	21	Cognitive flexibility (y); processing speed (y); sustained attention (y); working memory (n); spontaneous orienting during free play (s)
Kloo and Perner	2003	TD	3- to 5-year-olds	Cognitive Flexibility—Card sorting task (similar to Wisconsin Task).	2 sessions (15 mins per session) over 2 weeks (30 mins total)	Group trained at number conservation tasks or relative clauses	14	False belief (y); switching (card-sorting) (y)
Bergman Nutley et al. (WM group)	2011	TD	4- to 4.5-year-olds	WM—visuospatial (Cogmed)	25 sessions (15 mins per session) over 5–7 weeks (375 mins total)	Received nonadaptive training (combined NVR and WM)	24	Working memory/ short-term memory (y); reasoning (fluid intelligence (Gf latent variable) (n)
Bergman Nutley et al. (NVR group)	2011	TD	4- to 4.5-year-olds	Computerized nonverbal reasoning (NVR) training based on three tests from the Leiter Battery	25 sessions (15 mins per session) over 5–7 weeks (375 mins total)	Received nonadaptive training (combined NVR and WM)	25	Working memory/ short-term memory (n); reasoning (Gf latent variable) (y)
Thorell et al. (inhibition group)	2009	TD	4- to 5-year-olds	Inhibition (variant of Go/No-Go)	25 sessions - 5 weeks of 15 mins per school day (375 mins total)	Active group played commercially available computer games; passive group only took part in pre- and posttesting	17	Selective attention (Stroop) (n); visual WM (Wechsler/span board) (n); sustained attention (Continuous performance task (CPT) (n); reasoning (Wechsler) (n); inhibition (Go/No-Go) (n)

(Continued)

Table 1 (Continued).

Authors	Year	Description of participants	Age of participants	Nature of training	Amount of training	Control	N trained	Pre- and Posttests
Thorell et al. (WM group)	2009	TD	4- to 5-year-olds	WM—visuospatial (Cogmed)	25 sessions: 5 weeks of 15 mins per school day (375 mins total)	Active group played commercially available computer games; passive group only took part in pre- and posttesting	17	Selective attention (Stroop) (n); visual WM (Wechsler/ span board) (y); sustained attention (CPT) (y); reasoning (Wechsler) (n); inhibition (Go/No-Go) (s)
Rueda et al.	2005	TD	4-year-olds and 6-year-olds (separate groups)	Mixed Attention—tracking an object; anticipation; stimulus discrimination; inhibitory control	5 sessions (45 mins per session, spread out over 2 to 3 weeks) (225 mins total)	Brought into the lab for the same no. of sessions, watched children's videos	24 four-year-olds (18 for ANT); twelve 6-year-olds	Executive attention (Attention Network Tests (ANT) conflict) (y); alerting attention (ANT) (n); orienting attention (ANT) (n); reasoning (Kaufman-Brief Intelligence Test (K-BIT)) (s); general behavior (Childhood Behavior Questionnaire) (n)
Rueda, Checa, and Cómbita	2012	TD	5.5-year-olds	Mixed Attention—tracking an object; anticipation; stimulus discrimination; inhibitory control	10 sessions (45 mins per session) over 5 weeks (450 mins total)	Brought into the lab for the same no. of sessions, watched children's videos	18	Reasoning (K-BIT) (s); Attention (ANT, all subcomponents) (n); gambling task (n); Delay of gratification (s)

Note. In the final column, "Pre- and Posttests," "y," indicates that a significant training improvement was observed relative to controls, "s" indicates that some training improvement was observed (either $p < .1$ on the core measure or significant improvement at some but not all subcomponents), and "n" indicates no training improvement was observed.

training stimuli that change contingent on where on the screen the infant was looking. Using this interface, Wass et al. (2011) administered a battery of tasks targeting interference resolution, inhibition, task switching, and working memory for objects embedded in scenes of varying complexity to typically developing 11-month-old infants. Seventy-seven minutes of training were administered over four visits spread over 2 weeks; the effect of training was assessed relative to a control group who attended a matched number of ersatz training visits. Immediately posttraining, increased cognitive control and sustained attention were observed (Wass et al., 2011); attentional disengagement and saccadic reaction time latencies were reduced following training, and marginally nonsignificant changes in looking behavior during free play were also observed. No changes were found in working memory. Current ongoing work is investigating whether these findings can be replicated in "high-risk" populations, such as infants from low socio-economic status backgrounds.

Researchers working with toddlers have used computerized point-and-click interfaces to administer training targeting different executive domains: working memory, nonverbal reasoning, inhibition, and attentional control (mixed). Thorell and colleagues (2009) trained visuospatial working memory in typically developing 4- to 5-year-old children. A total of 6 hours of training was administered in one 5-week phase. They observed improvement posttraining at nontrained working memory tasks, on an auditory Continuous Performance Task (CPT), and on Go/No-Go omissions but no improvement on problem solving, Go/No-Go response speed, and on a Stroop-like task. Bergman Nutley and colleagues (2011) trained WM in typically developing 4-year-old children and identified posttraining transfer to nontrained working memory tasks but not to problem solving tasks. Subsequent analyses of this study suggested that the degree of training improvement observed was related to variation in the dopamine transporter gene DAT1 (Söderqvist et al., 2012; see also Klingberg, 2010; McNab et al., 2009).

Thorell and colleagues (2009) also trained typically developing 4- to 5-year-old children at inhibition using variants of the Go/No-Go paradigm and flanker task; training was spread over 6 hours across one 5-week phase. They identified no significant transfer to tasks such as Stroop and CPT, with less transfer observed than in the group that had received WM training. In their discussion, Thorell and colleagues suggest that the larger training effects observed following WM than inhibition training may be attributable to methodological issues, such as problems defining how the difficulty of the inhibition tasks changes adaptively during training (see Klingberg, 2010).

Rueda and colleagues administered a battery of training tasks targeting object tracking, anticipation, stimulus discrimination, conflict resolution, and inhibitory control to groups of 4- and 6-year-old children. A total of 3.5 hours of training was administered over 2–3 weeks. They found substantial within-task training effects; pre- and posttests identified some transfer to reasoning tasks but no significant changes to performance on the Attention Network Test or Childhood Behaviour Questionnaire (Rueda et al., 2005). Subsequent work replicated some of these effects and showed that some (weaker) effects of training were also discernable at 2-month follow-up. Event-related potentials (ERPs) were also recorded, which suggested a more efficient and faster activation of the executive attention network after training (Rueda et al., 2005, 2012). Kloo and Perner (2003) administered 30 minutes of noncomputerized training targeting either Dimensional Card Change Sorting or false belief to typically developing 3- to 5-year-old children and observed bidirectional transfer at posttesting relative to an active control group.

The majority of the developmental work in this field has involved older children (aged 7+ years). Holmes et al. (2009) administered WM training sessions to 8- to

11-year-old children and examined transfer to other academic measures. Loosli, Buschkuehl, Perrig, and Jaeggi, (2012) administered ten 12-minute WM training sessions to typically developing 9- to 11-year-old children and identified evidence of improved reading performance after training but no improvement on a reasoning task. St Clair Thompson (2007) administered training targeting explicit mnemonic strategies to typically developing 7-year-olds and found transfer to some language and WM tasks but not to standardized reading arithmetic or math tests, either immediately or 5 months later. Klingberg et al. (2005) administered WM training for at least 20 days to 7- to 12-year-old children with ADHD and identified improved performance at Stroop, non-verbal reasoning and nontrained working memory tasks, along with some evidence of reduction on parental (but not teacher) ratings of ADHD symptom severity. Green et al. (2012) applied similar training to children with ADHD and found reductions posttraining in experimentally assessed off-task behaviors but not in parent ratings of ADHD severity. Kray, Karbach, Haenig, and Freitag, (2011) trained 8- to 12-year-old children with ADHD at a variant of the Wisconsin Card Sorting task and identified improvements posttraining on the Stroop task but not on assessments of nonverbal reasoning and processing speed. Kerns, Eso, and Thomson (1999) administered similar training to 7- to 11-year-old children with ADHD and found improvement posttraining on some (but not other) experimental assessments of nonverbal reasoning, sustained attention, as well as the Stroop task. Improvements were also noted on some but not on other ratings of inattention-impulsivity.

Do studies targeting younger individuals report more widespread transfer of training effects? Melby-Lervag and Hulme (2013) looked at transfer reported to nontrained working memory tasks following working memory training and found that younger children showed significantly larger benefits from training than do older children; however, no evidence was found of increased transfer to nonverbal abilities. Wass, Scerif, & Johnson, (2012) analyzed 34 studies that applied cognitive training targeting working memory or attentional control to individuals aged between 1–80 years, and analyzed the posttraining transfer observed. They identified a significant relationship between the age of participants and the degree of training transfer reported ($r = -.31$), suggesting that training targeted at younger participants tended to lead to more widespread transfer of training effects. This effect became stronger when the amount of training administered was included as a covariate ($r = -.37$), and when those studies targeting typically developing individuals were considered independently ($r = -.53$). However, comparing the studies targeting 4- to 6-year-olds with those targeting 7- to 10-year-olds suggests a contrary effect, namely that most of the largest observed training effects are found in the 7- to 10-year-old age range. Possible reasons for this are discussed below.

Summary and Recommendations for Future Work

The number of studies that have successfully applied targeted cognitive training to individuals in the 0- to 5-year age range is low. However, the fact that several studies have successfully reported training effects, together with the number of studies that have reported similar findings in older children, suggests the future potential of these methods.

However, a number of limitations should be recognized to the studies reviewed here. All studies included in the 0–5 age range were conducted with typically developing rather than high-risk populations; future work should also explore whether similar training effects can be identified in clinical populations or those identified as "high-risk" via

epidemiological, familial, and genetic risk factors. It should also be noted that all the studies reviewed here have administered a single, discrete "dose" of training (e.g., 3.5 hours over 2–3 weeks), which from a developmental perspective may be suboptimal; future work should explore the effect of administering much larger doses of training spaced over longer time periods (cf., e.g., Slagter et al., 2007), as well as assessing the degree to which training improvements are maintained over longer time periods. The total amount of training administered in these studies also tends to be small (e.g., 77 minutes in Wass et al., 2011). One reason for this is a practical one: Cognitive training regimes are often intrinsically repetitive and, with infants and young toddlers, meta-cognitive factors (an awareness that what they are doing should be good for them) cannot be used to encourage participation.

Future work should explore practical ways of addressing these challenges to make longer training phases viable with very young individuals: first, using a number of different training tasks in rotation; these can either be heterogeneous (e.g., Rueda et al., 2005; Wass et al., 2011) or different tasks targeting similar cognitive mechanisms (Klingberg et al., 2005); second, using adaptive change criteria such that both the difficulty of the training task and the audiovisual content of the training task change contingent on task performance; third, actively monitoring participants' engagement levels during training, including the use of exogenously salient stimuli to re-attract participants' attention when they become distracted; fourth, careful design of responses for correct and incorrect rewards to reward participation over longer time scales; fifth, the use of different methods for interacting with the training paradigms (Wass & Porayska-Pomsta, 2013). For infants, who lack the fine motor skills to interface via a point-and-click or touchscreen interface, eye gaze appears to be an effective interface – particularly because the control of visual attention is thought to be important in mediating learning (e.g., Frischen, Bayliss, & Tipper, 2007). Future work with toddlers can incorporate touchscreen technology and motion-contingent interfaces to provide a more immersive training environment.

One further important point is the heterogeneity or homogeneity of the training regime. I have reviewed cognitive training studies that administered a relatively heterogeneous battery of training tasks targeting different subcomponents of attentional control (e.g., Rueda et al., 2005; Wass et al., 2011) and others that administered more homogenous training targeting a single component of cognition (e.g., visuospatial working memory; Thorell et al., 2009). Some authors have suggested that heterogeneous training batteries may be more effective in influencing global behavioral outcomes such as academic learning or clinical diagnoses (Wallace & Rogers, 2010), although this question has not to our knowledge been assessed systematically. The disadvantage of heterogeneous training, however, is that the results are often inconclusive as to which of the elements of the battery has been responsible for the observed changes in behavior; this can make causal mechanistic pathways hard to untangle. A homogenous training battery, in contrast, may be more informative in helping us to understand underlying developmental mechanisms but less effective in influencing global behavioral outcomes.

The most crucial avenue for future work, though, will involve applying early targeted training to clinical or high-risk populations. Possible future targets include infants born prematurely (Voigt et al., 2012), infants from "high-risk" backgrounds, such as low-socioeconomic status (SES) families (Welsh, Nix, Blair, Bierman, & Nelson, 2010), as well as infants with family histories of clinical conditions, such as ASD (Elsabbagh & Johnson, 2012) and ADHD (Auerbach et al., 2004).

Research with these atypical populations will allow us to address a number of key questions about the role that domain-general faculties such as executive control play in mediating other aspects of development. For example, Johnson argued that it may be that deficits in EFs are observed across a range of developmental disorders because individuals with strong EF skills are better able to compensate for atypicalities in other brain systems early in life (Johnson, 2012). Dynamic, multidomain disease models of this type are hard to assess using correlational (even longitudinal correlational) techniques for reasons documented above. However, they make specific and falsifiable predictions for the differential transfer of effects that would be observed across individuals following targeted training to EFs early in development.

Another question that can be addressed in research with atypical populations is that of whether some individuals may benefit more from training than others (cf. Söderqvist et al., 2012). Multiple cognitive domains are involved in the achievement of learning goals such as early language acquisition (e.g., Rose et al., 2008). Is it the case that training executive control improves language acquisition only in cases where executive control was deficient and thereby exerting a limiting influence on language learning? Or does training executive control to supranormal levels also improve language learning, even in those individuals who show no initial executive control deficit? Addressing these hypotheses will enrich our understanding of the mechanisms underlying cognitive development.

A third question that can be assessed using targeted training is that of whether critical periods subsist during cognitive development — for example, for the involvement of executive control in language acquisition. Although a number of authors have speculated that this may be the case (e.g., Richardson & Thomas, 2008; Tomalski & Johnson, 2010), these questions are virtually impossible to assess using correlational methods. Examining how the effect of applying targeted training differs at different stages of cognitive development would potentially be informative here.

CONCLUSION: THE IMPORTANCE OF TARGETING THE EARLY, FORMATIVE STAGES OF COGNITIVE DEVELOPMENT

Researchers are increasingly recognizing the importance of developmentally informed models that understand how pathogenic disease mechanisms operate early in disrupted development. In this article, I have described studies from both typical and atypical development that suggested that early developing individual differences in attentional control and working memory may play a role in mediating later-emerging differences in learning in academic and other settings. These findings have been reported within typical development (Snyder & Munakata, 2011) as well as within a number of disorder or at-risk groups including individuals born preterm (Rose et al., 2008), from low-SES backgrounds (Welsh et al., 2010), at risk of ADHD (Lawson & Ruff, 2004), and with genetic disorders such as Fragile X syndrome and Down's syndrome (Cornish et al., 2007; Cornish, Cole, et al., 2012; Cornish, Scerif, et al., 2012). These findings point to the potential utility of investigating early and intensive interventions designed to remediate early emerging deficits in attentional control.

I have also described evidence suggesting that the effects of training attentional control and working memory can be detected, even following only very small doses of training (0.5–6 hours), in individuals in the 0- to 6-year age range. I have concluded, however, that the number of studies in this area is currently low. I have discussed possible

directions for future work, including assessing medium-term training effects and working with young, "high-risk" populations.

REFERENCES

Aslin, R. N. (2007). What's in a look? *Developmental Science, 10*(1), 48–53.

Astle, D. E., & Scerif, G. (2009). Using developmental cognitive neuroscience to study behavioral and attentional control. *Developmental Psychobiology, 51*(2), 107–118.

Auerbach, J. G., Atzaba-Poria, N., Berger, A., & Landau, R. (2004). Emerging developmental pathways to ADHD: Possible path markers in early infancy. *Neural Plasticity, 11*(1–2), 29–43.

Bell, M. A., & Wolfe, C. D. (2007). Changes in brain functioning from infancy to early childhood: Evidence from EEG power and coherence during working memory tasks. *Developmental Neuropsychology, 31*(1), 21–38.

Bergman Nutley, S., Soderqvist, S., Bryde, S., Thorell, L. B., Humphreys, K., & Klingberg, T. (2011). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: A controlled, randomized study. *Developmental Science, 14*(3), 591–601.

Bopp, K. D., Mirenda, P., & Zumbo, B. D. (2009). Behavior predictors of language development over 2 years in children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research, 55*(5), 1106–1120.

Breckenridge, K., Atkinson, J., & Braddick, O. (2012). Attention. In E. Farran & A. Karmiloff-Smith (Eds.), *Neuroconstructivism: The multidisciplinary study of genetic syndromes from infancy to adulthood* (pp. 119–134). Oxford: Oxford University Press.

Brehmer, Y., Rieckmann, A., Bellander, M., Westerberg, H., Fischer, H., & Backman, L. (2012). Neural correlates of training-related working-memory gains in old age. *NeuroImage, 58*(4), 1110–1120.

Brown, J. H., Johnson, M. H., Paterson, S. J., Gilmore, R., Longhi, E., & Karmiloff-Smith, A. (2003). Spatial representation and attention in toddlers with Williams syndrome and down syndrome. *Neuropsychologia, 41*(8), 1037–1046.

Bryck, R. L., & Fisher, P. A. (2012). Training the brain: Practical applications of neural plasticity from the intersection of cognitive neuroscience, developmental psychology, and prevention science. *American Psychologist, 67*(2), 87–100.

Campbell, F. A., Wasik, B. H., Pungello, E., Burchinal, M., Barbarin, O., Kainz, K., … Ramey, C. T. (2008). Young adult outcomes of the Abecedarian and CARE early childhood educational interventions. *Early Childhood Research Quarterly, 23*(4), 452–466.

Chawarska, K., Volkmar, F., & Klin, A. (2010). Limited attentional bias for faces in toddlers with autism spectrum disorders. *Archives of general psychiatry, 67*(2), 178–185.

Clair-Thompson, H. L. S. (2007). The influence of strategies on relationships between working memory and cognitive skills. *Memory, 15*(4), 353–365.

Cohen Kadosh, K., & Johnson, M. H. (2007). Developing a cortex specialized for face perception. *Trends in Cognitive Sciences, 11*(9), 367–369.

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Advances in Child Development and Behavior, 34*, 283–322.

Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental Psychology, 34*(1), 28–38.

Cornish, K., Cole, V., Longhi, E., Karmiloff-Smith, A., & Scerif, G. (2012). Does attention constrain developmental trajectories in Fragile X syndrome? A 3-year prospective longitudinal study. *American Journal on Intellectual and Developmental Disabilities, 117*(2), 103–+.

Cornish, K., Scerif, G., & Karmiloff-Smith, A. (2007). Tracing syndrome-specific trajectories of attention across the lifespan. *Cortex, 43*(6), 672–685.

Cornish, K., Scerif, G., & Karmiloff-Smith, A. (2012). Developmental trajectories of visual attention in young children with Fragile X syndrome: Developmental delay or developmental freeze? *Journal of Intellectual Disability Research, 55*, 950–950.

Cornish, K., Sudhalter, V., & Turk, J. (2004). Attention and language in fragile X. *Mental Retardation and Developmental Disabilities Research Reviews, 10*(1), 11–16.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037–2078.

Dice, J. L., & Schwanenflugel, P. (2012). A structural model of the effects of preschool attention on kindergarten literacy. *Reading and Writing, 25*(9), 2205–2222.

Dixon Jr, W. E., & Smith, P. H. (2008). Attentional focus moderates habituation—Language relationships: Slow habituation may be a good thing. *Infant and Child Development, 17*(2), 95–108.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23*(10), 475–483.

Durston, S., Davidson, M. C., Tottenham, N., Galvan, A., Spicer, J., Fossella, J. A., & Casey, B. J. (2006). A shift from diffuse to focal cortical activity with development. *Developmental Science, 9*(1), 1–8.

Elsabbagh, M., & Johnson, M. H. (2012). Getting answers from babies about autism. *Trends in Cognitive Sciences, 14*(2), 81–87.

Elsabbagh, M., Volein, A., Holmboe, K., Tucker, L., Csibra, G., Baron-Cohen, S., … Johnson, M. H. (2009). Visual orienting in the early broader autism phenotype: Disengagement and facilitation. *Journal of Child Psychology and Psychiatry, 50*(5), 637–642.

Fair, D. A., Cohen, A. L., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., Barch, … Schlaggar, B. L. (2008). The maturing architecture of the brain's default network. *Proceedings of the National Academy of Sciences of the United States of America, 105*(10), 4028–4032.

Fair, D. A., Posner, J., Nagel, B. J., Bathula, D., Dias, T. G., Mills, K. L., & Nigg, J. T. (2010). Atypical default network connectivity in youth with attention-deficit/hyperactivity disorder. *Biological psychiatry, 68*(12), 1084–1091.

Friedman, A. H., Watamura, S. E., & Robertson, S. S. (2005). Movement-attention coupling in infancy and attention problems in childhood. *Developmental Medicine and Child Neurology, 47*(10), 660–665.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin, 133*(4), 694–724.

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A. M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93*(3), 265–281.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences, 16*(2), 129–135.

Gilmore, R. O., & Johnson, M. H. (1995). Working memory in infancy—6-month-olds performance on 2 versions of the oculomotor delayed-response task. *Journal of Experimental Child Psychology, 59*(3), 397–418.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., … Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America, 101*(21), 8174–8179.

Green, C. T., Long, D. L., Green, D., Iosif, A.-M., Faye Dixon, J., Miller, M. R., … Schweitzer, J. B. (2012). Will working memory training generalize to improve off-task behavior in children with attention-deficit/hyperactivity disorder? *Neurotherapeutics, 9*(3), 639–648.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900–1902.

Holmboe, K., Elsabbagh, M., Volein, A., Tucker, L. A., Baron-Cohen, S., Bolton, P., … Johnson, M. H. (2010). Frontal cortex functioning in the infant broader autism phenotype. *Infant Behavior & Development, 33*(4), 482–491.

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12*(4), F9–F15.

Johnson, M. H. (1995). The inhibition of automatic saccades in early infancy. *Developmental Psychobiology, 28*(5), 281–291.

Johnson, M. H. (2010). *Developmental cognitive neuroscience* (3rd ed.). Oxford: Wiley-Blackwell.

Johnson, M. H. (2012). Executive function and developmental disorders: The flip side of the coin. *Trends in Cognitive Sciences, 16*(9), 454–457.

Johnson, M. H., Halit, H., Grice, S. J., & Karmiloff-Smith, A. (2002). Neuroimaging of typical and atypical development: A perspective from multiple levels of analysis. *Development and Psychopathology, 14*(3), 521–536.

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy—contingency learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience, 3*(4), 335–344.

Kannass, K. N., & Oakes, L. M. (2008). The development of attention and its relations to language in infancy and toddlerhood. *Journal of Cognition and Development, 9*(2), 222–246.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2*(10), 389–398.

Karmiloff-Smith, A. (2007). Atypical epigenesis. *Developmental Science, 10*(1), 84–88.

Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: Rethinking the study of developmental disorders. *Developmental Psychology, 45*(1), 56–63.

Karmiloff-Smith, A., D'Souza, D., Dekker, T. M., Van Herwegen, J., Xu, F., Rodic, M., & Ansari, D. (2012). Genetic and environmental vulnerabilities in children with neurodevelopmental disorders. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 17261–17265.

Kegel, C. A. T., & Bus, A. G. (2012). Links between DRD4, executive attention, and alphabetic skills in a nonclinical sample. *Journal of Child Psychology and Psychiatry, 54*(3), 305–312.

Kelly, A. M. C., Di Martino, A., Uddin, L. Q., Shehzad, Z., Gee, D. G., Reiss, P. T., … Milham, M. P. (2009). Development of anterior cingulate functional connectivity from late childhood to early adulthood. *Cerebral Cortex, 19*(3), 640–657.

Kerns, K. A., Eso, K., & Thomson, J. (1999). Investigation of a direct intervention for improving attention in young children with ADHD. *Developmental Neuropsychology, 16*(2), 273–295.

Kikuchi, Y., Senju, A., Akechi, H., Tojo, Y., Osanai, H., & Hasegawa, T. (2011). Atypical disengagement from faces and its modulation by the control of eye fixation in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 41*(5), 629–645.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences, 14* (7), 317–324.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., … Westerberg, H. (2005). Computerized training of working memory in children with ADHD—A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*(2), 177–186.

Kloo, D., & Perner, J. (2003). Training transfer between card sorting and false belief understanding: Helping children apply conflicting descriptions. *Child Development, 74*(6), 1823–1839.

Kray, J., Karbach, J., Haenig, S., & Freitag, C. (2011). Can task-switching training enhance executive control functioning in children with attention deficit/-hyperactivity disorder? *Frontiers in Human Neuroscience, 5*, 180.

Landry, R., & Bryson, S. E. (2004). Impaired disengagement of attention in young children with autism. *Journal of Child Psychology and Psychiatry, 45*(6), 1115–1122.

Lawson, K. R., & Ruff, H. A. (2004). Early focused attention predicts outcome for children born prematurely. *Journal of Developmental and Behavioral Pediatrics, 25*(6), 399–406.

Loosli, S. V., Buschkuehl, M., Perrig, W. J., & Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, 18*(1), 62–78.

Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M., Thomas, M., & Westermann, G. (2007). *Neuroconstructivism, Vol. I: How the brain constructs cognition*. Oxford: Oxford University Press.

McNab, F., Varrone, A., Farde, L., Jucaite, A., Bystritsky, P., Forssberg, H., & Klingberg, T. (2009). Changes in cortical dopamine D1 receptor binding associated with cognitive training. *Science, 323*(5915), 800–802.

Melby-Lervag, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*(2), 270–291.

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences, 15*(10), 453–459.

Nigg, J. T. (2006). *What causes ADHD? Toward a multi-path model for understanding what goes wrong and why* (422p). New York, NY: The Guilford Press.

Olds, D. L., Sadler, L., & Kitzman, H. (2007). Programs for parents of infants and toddlers: Recent evidence from randomized trials. *Journal of Child Psychology and Psychiatry, 48*(3–4), 355–391.

Oliver, A., Johnson, M. H., Karmiloff-Smith, A., & Pennington, B. (2000). Deviations in the emergence of representations: A neuroconstructivist framework for analysing developmental disorders. *Developmental Science, 3*(1), 1–40.

Quartz, S. R., & Sejnowski, T. J. (1997). The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences, 20*(4), 537–596.

Redcay, E., Haist, F., & Courchesne, E. (2008). Paper: Functional neuroimaging of speech perception during a pivotal period in language acquisition. *Developmental Science, 11*(2), 237–252.

Richardson, F. M., & Thomas, M. S. C. (2008). Critical periods and catastrophic interference effects in the development of self-organizing feature maps. *Developmental Science, 11*(3), 371–389.

Richmond, L. L., Morrison, A. B., Chein, J. M., & Olson, I. R. (2011). Working memory training and transfer in older adults. *Psychology and Aging, 26*(4), 813–822.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2005). The structure of infant cognition at 1 year. *Intelligence, 33*(3), 231–250.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2009). A cognitive approach to the development of early language. *Child Development, 80*(1), 134–150.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2011). Modeling a cascade of effects: The role of speed and executive functioning in preterm/full-term differences in academic achievement. *Developmental Science, 14*(5), 1161–1175.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2012). Implications of infant cognition for executive functions at age 11. *Psychological Science, 23*(11), 1345–1355.

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Van Rossem, R. (2008). A cognitive cascade in infancy: Pathways from prematurity to later mental development. *Intelligence, 36*(4), 367–378.

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Van Rossem, R. (2011). Basic information processing abilities at 11 years account for deficits in IQ associated with preterm birth. *Intelligence, 39*(4), 198–209.

Rueda, M. R., Checa, P., & Cómbita, L. M. (2012). Enhanced efficiency of the executive attention network after training in preschool children: Immediate changes and effects after two months. *Developmental Cognitive Neuroscience, 2S*, S192–S204.

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences of the United States of America, 102*(41), 14931–14936.

Ruff, H. A., & Rothbart, M. K. (1996). *Attention in early development: Themes and variations*. New York, NY: Oxford University Press.

Scerif, G. (2010). Attention trajectories, mechanisms and outcomes: At the interface between developing cognition and environment. *Developmental Science, 13*(6), 805–812.

Scerif, G., Cornish, K., Wilding, J., Driver, J., & Karmiloff-Smith, A. (2004). Visual search in typically developing toddlers and toddlers with Fragile X or Williams syndrome. *Developmental Science, 7*(1), 116–130.

Scerif, G., Longhi, E., Cole, V., Karmiloff-Smith, A., & Cornish, K. (2012). Attention across modalities as a longitudinal predictor of early outcomes: The case of fragile X syndrome. *Journal of Child Psychology and Psychiatry, 53*(6), 641–650.

Scherf, K. S., Sweeney, J. A., & Luna, B. (2006). Brain basis of developmental change in visuospatial working memory. *Journal of Cognitive Neuroscience, 18*(7), 1045–1058.

Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., … Wise, S. P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *Journal of Neuroscience, 28* (14), 3586–3594.

Shing, Y. L., Lindenberger, U., Diamond, A., Li, S.-C., & Davidson, M. C. (2010). Memory maintenance and inhibitory control differentiate from early childhood to adolescence. *Developmental Neuropsychology, 35*(6), 679–697.

Shonkoff, J. P., & Levitt, P. (2010). Neuroscience and the future of early childhood policy: Moving from why to what and how. *Neuron, 67*(5), 689–691.

Slagter, H. A., Lutz, A., Greischar, L. L., Francis, A. D., Nieuwenhuis, S., Davis, J. M., & Davidson, R. J. (2007). Mental training affects distribution of limited brain resources. *PLoS Biology, 5*(6), 1228–1235.

Smith, L. B., & Sheya, A. (2011). Is cognition enough to explain cognitive development? *Topics in Cognitive Science, 2*(4), 725–735.

Snyder, H. R., & Munakata, Y. (2011). Becoming self-directed: Abstract representations support endogenous flexibility in children. *Cognition, 116*(2), 155–167.

Sonuga-Barke, E. J. S., & Halperin, J. M. (2011). Developmental phenotypes and causal pathways in attention deficit/hyperactivity disorder: Potential targets for early intervention? *Journal of Child Psychology and Psychiatry, 51*(4), 368–389.

Spencer-Smith, M., Anderson, P., Jacobs, R., Coleman, L., Long, B., & Anderson, V. (2011). Does timing of brain lesion have an impact on children's attention? *Developmental Neuropsychology, 36*(3), 353–366.

Stiles, J., Reilly, J., Paul, B., & Moses, P. (2005). Cognitive development following early brain injury: Evidence for neural adaptation. *Trends in Cognitive Sciences, 9*(3), 136–143.

Söderqvist, S., Bergman Nutley, S., Peyrard-Janvid, M., Matsson, H., Humphreys, K., Kere, J., & Klingberg, T. (2012). Dopamine, working memory, and training induced plasticity: Implications for developmental research. *Developmental Psychology, 48*(3), 836–843.

Thorell, L. B., Lindqvist, S., Nutley, S. B., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science, 12*(1), 106–113.

Tomalski, P., & Johnson, M. H. (2010). The effects of early adversity on the adult and developing brain. *Current Opinion in Psychiatry, 23*(3) 233–238.

Triesch, J., Teuscher, C., Deak, G. O., & Carlson, E. (2006). Gaze following: Why (not) learn it? *Developmental Science, 9*(2), 125–147.

Velanova, K., Wheeler, M. E., & Luna, B. (2008). Maturational changes in anterior cingulate and frontoparietal recruitment support the development of error processing and inhibitory control. *Cerebral Cortex, 18*(11), 2505–2522.

Voigt, B., Pietz, J., Pauen, S., Kliegel, M., & Reuner, G. (2012). Cognitive development in very vs. moderately to late preterm and full-term children: Can effortful control account for group differences in toddlerhood? *Early Human Development, 88*(5), 307–313.

Wallace, K. S., & Rogers, S. J. (2010). Intervening in infancy: Implications for autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 51*(12), 1300–1320.

Wang, M.-Y., Chang, C.-Y., & Su, S.-Y. (2011). What's cooking?—cognitive training of executive function in the elderly. *Frontiers in Psychology, 2*, 228.

Wass, S. V., & Porayska-Pomsta, K. (2013). The uses of cognitive training technologies in the treatment of autism spectrum disorders. *Autism, 17*(6). doi:10.1177/1362361313499827

Wass, S. V., Porayska-Pomsta, K., & Johnson, M. H. (2011). Training attentional control in infancy. *Current Biology, 21*(18), 1543–1547.

Wass, S. V., Scerif, G., & Johnson, M. H. (2012). Training attentional control and working memory: Is younger, better? *Developmental Review, 32*(4), 360–387.

Webb, S. J., Jones, E. J. H., Merkle, K., Namkung, J., Toth, K., Greenson, J., … Dawson, G. (2010). Toddlers with elevated Autism symptoms show slowed habituation to faces. *Child Neuropsychology, 16*(3), 255–278.

Weindrich, D., Jennen-Steinmetz, C., Laucht, M., & Schmidt, M. H. (2003). Late sequelae of low birthweight: Mediators of poor school performance at 11 years. *Developmental Medicine and Child Neurology, 45*(7), 463–469.

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology, 102*(1), 43–53.

Willoughby, M., Kupersmidt, J., Voegler-Lee, M., & Bryant, D. (2011). Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. *Developmental Neuropsychology, 36*(2), 162–180.