Review

Training attentional control and working memory – Is younger, better?

S.V. Wass [a,*], G. Scerif [b], M.H. Johnson [a]

[a] Centre for Brain and Cognitive Development, Psychological Sciences, Birkbeck, University of London, Malet Street, London WC1E 7HX, UK
[b] Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK

ARTICLE INFO

ABSTRACT

Authors have argued that various forms of interventions may be more effective in younger children. Is cognitive training also more effective, the earlier the training is applied? We review evidence suggesting that functional neural networks, including those subserving attentional control, may be more unspecialised and undifferentiated earlier in development. We also discuss evidence suggesting that certain skills such as attentional control may be important as 'hub' cognitive domains, gating the subsequent acquisition of skills in other areas. Both of these factors suggest that attentional training administered to younger individuals ought to be relatively more effective in improving cognitive functioning across domains. We evaluate studies that have administered forms of cognitive training targeting various subcomponents of attention and the closely related domain of working memory, and we contrast their reported transfer to distal cognitive domains as a function of the age of the participants. Although negative findings continue to be common in this literature we find that cognitive training applied to younger individuals tends to lead to significantly more widespread transfer of training effects. We conclude that future work in this area should concentrate on understanding early intensive training, and discuss a number of practical steps that might help to achieve this aim.

* Corresponding author.
E-mail address: sam.wass@bbk.ac.uk (S.V. Wass).

Introduction

In recent years a number of authors within applied cognitive neuroscience have advocated the desirability of early interventions (Burger, 2010; Heckman, 2006; Shonkoff & Levitt, 2011; Sonuga-Barke, Koerting, Smith, McCann, & Thompson, 2011; Wallace & Rogers, 2010). Several studies have suggested, for example, that interventions providing increased social and educational provision for young children from low socio-economic status backgrounds are more effective the earlier the training is applied (Campbell et al., 2008; Olds, Sadler, & Kitzman, 2007). Similarly, programs are currently being set up to assess the impact of early interventions for individuals at high risk of developing conditions such as Attention Deficit/Hyperactivity Disorder (ADHD) (Sonuga-Barke & Halperin, 2010) and Autism Spectrum Disorders (ASD) (Wallace & Rogers, 2010).

Why are early interventions considered desirable? Heckman (2006) argued this point from an inter-disciplinary perspective at four levels: first, that cognitive and neural development is influenced by gene–environment interactions; second, that the mastery of skills needed for economic success follows hierarchical rules, with later attainments building on earlier; third, that cognitive, linguistic, social and emotional competencies are interdependent; fourth, that human abilities are formed in a predictable sequence of sensitive periods, during which development appears plastic (see also Karmiloff-Smith, 1998; Sonuga-Barke & Halperin, 2010).

Within the field of cognitive training, however, there appears to be relatively little appreciation of the importance of the developmental perspective. A number of studies have looked at the effect of administering training in older adults, young adults and school-age children, with only a small number of studies looking at the effect of applying training in very young populations. Furthermore, no review has yet considered whether cognitive training applied to younger participants tends to report more widespread transfer of training effects than similar training applied to older participants. This is despite considerable *a priori* evidence at both the neural and behavioural levels suggesting the plausibility of this hypothesis. In the next three introductory sections we outline some of this evidence.

Overlapping functional connectivity patterns in early brain development

Evidence from developmental cognitive neuroscience is gradually eroding the idea that the postnatal functional development of the cortex unfolds in a mosaic-like manner, with static structure–function correspondences during development (Elman et al., 1996; Fair et al., 2008, 2009; Johnson, 2000, 2010; Karmiloff-Smith, 1992, 2009). Instead, evidence of how functional activation patterns change during typical development (Bell & Wolfe, 2007; Durston et al., 2006; Kelly et al., 2009), in particular in the context of the acquisition of skills such as face perception (e.g. Cohen Kadosh & Johnson, 2007; Grossmann & Johnson, 2007), language (Redcay, Haist, & Courchesne, 2008), and also in terms of reorganisation following acquired damage early in life (e.g. Spencer-Smith et al., 2011) is leading to a better appreciation that postnatal cortical development is a non-stationary, dynamic process (see also Edelman, 1993; Quartz & Sejnowski, 1997).

Amongst these developmentally dynamic perspectives, the Interactive Specialization hypothesis describes how increasing functional localization and specialization of cortical circuits arises (in part) as the emergent property of competition and cooperation between different circuits and networks (Johnson, 2000, 2010; Johnson, Halit, Grice, & Karmiloff-Smith, 2002; Mareschal et al., 2007). Early in postnatal development some regions of the human cortex are relatively unspecialised and undifferentiated: specific tasks evoke larger functional activation patterns, and cortical areas are relatively less functionally specialised. As particular areas become better tuned to particular tasks, Hebbian competitive and cooperative learning algorithms mean that functional activation patterns become gradually more localised and neural regions become more specialised over developmental time (Oliver, Johnson, Karmiloff-Smith, & Pennington, 2000).

However compelling, evidence on these dynamic developmental changes does not identity causal factors: they could just as well be determined by maturational changes as by interactions with the environment. Recently, cognitive training programs have offered the opportunity to test, instead, the extent to which such developmental changes are affected by controlled environmental input,

and differentially so over development. Firstly, plastic changes induced by training are not compatible with a static maturational framework. Secondly, and especially of interest to developmentalists, if we assume (following Klingberg, 2010) that the effects of training on a particular cortical region using a specific task would only be expected to transfer to other tasks and functions to the extent that the tasks rely on at least partially overlapping neural networks, then it follows that training targeted earlier rather than later in development ought to lead to more widespread transfer of training effects (see Bellander et al. (2011), Brocki, Clerkin, Guise, Fan, and Fossella (2009), Diamond, Briand, Fossella, and Gehlbach (2004), Kolata et al. (2010), McNab et al. (2009) and Posner and Rothbart (2009) for discussions of possible mechanisms underlying these neural changes).

Interactions between cognitive domains during development

In addition to the evidence from neuroimaging, behavioural faculties may interact to a greater extent during early development compared to later in life, as particular skills are required for the subsequent acquisition of other abilities (e.g. Cornish, Sudhalter, & Turk, 2004; Karmiloff-Smith, 1992, 1998). Indications that interactions across cognitive domains are critical early in development come from developmental disorders. Particularly in disorders of known genetic origins, interactions identified across all levels of description, from behavioural manifestations, to cognitive processes, to neural systems and molecular pathways, question modular deficits and highlight cascading effects of changes across domains (Karmiloff-Smith, 2009; Scerif & Karmiloff-Smith, 2005). These interactions point to the importance of studying cognitive deficits not just as the adult end-state, but rather as the developmental pathways by which the end-state has been arrived at (Cornish, Scerif, & Karmiloff-Smith, 2007; Cornish, Turk, et al., 2004; Karmiloff-Smith, 1998, 2007; Thomas & Karmiloff-Smith, 2002).

For example, in the genetic disorder Williams Syndrome (WS) infants and toddlers are impaired early on in planning saccadic eye movements (Brown et al., 2003; Scerif, Cornish, Wilding, Driver, & Karmiloff-Smith, 2004). It has been suggested that these difficulties may affect their ability to follow pointing (Laing et al., 2002), and which in turn may be detrimental to their ability to use parental referential pointing to learn vocabulary (Karmiloff-Smith, 2007). Although their language becomes relatively proficient much later in development, initially language in toddlers with WS is extremely delayed and follows a deviant developmental trajectory (Annaz, Karmiloff-Smith, Johnson, & Thomas, 2009; Paterson, Brown, Gsodl, Johnson, & Karmiloff-Smith, 1999). Thus, an early problem within the visuomotor system may potentially disrupt the later acquisition of language. The idea that an initial atypicality may have compounded, and more widespread, effects later in development suggests that early interventions may be critically important.

As illustrated by the case of developmental disorders, developmental dynamics need to be investigated over typical development. Theoretically relevant cognitive functions to explore in this context are domain-general processes, like attentional control, given their potential to gate and influence the development of multiple other processes (Astle & Scerif, 2009; Scerif, 2010). It is to attentional control that we now turn.

Attentional control as a tool for learning

According to Ruff and Rothbart's influential formulation, infants during the first year of life are thought to be able to exercise little volitional control over their allocation of attention (Ruff & Rothbart, 1996; see also Colombo & Cheatham, 2006; Courage, Reynolds, & Richards, 2006; Johnson, 2010). Towards the end of the first year of life, the capacity to exercise attentional control – i.e. the capacity of an individual to choose what is paid attention to and what is ignored – is thought to emerge as the neural circuitry substantiating these cognitive functions matures (Deoni et al., 2011; Gordon et al., 2011; Johnson, 2010). Attentional control capacities continue to progress slowly throughout development (Davidson, Amso, Cruess Anderson, & Diamond, 2006), relative to exogenous (i.e. stimulus-driven) aspects of attention which are thought to be relatively more mature at an earlier age (Iarocci, Enns, Randolph, & Burack, 2009). In recent years considerable attention has also been devoted by investigators in this field to working memory, the "maintenance of task-relevant information in mind for brief periods of time to guide subsequent behaviour" (Gazzaley & Nobre, 2011). Working

memory is thought to have substantially overlapping neural correlates with attentional control (Duncan & Owen, 2000; McNab et al., 2009), particularly early in development (Scherf, Sweeney, & Luna, 2006; Velanova, Wheeler, & Luna, 2008). Behavioral studies have similarly pointed to covariation in attentional control and working memory performance, which again may be more extensive earlier in development (Shing, Lindenberger, Diamond, Li, & Davidson, 2010).

The ability to exercise control over attention is thought to be a 'hub' cognitive faculty – i.e. a faculty required for the acquisition of skills in a range of other areas (Karmiloff-Smith, 1992; Posner & Rothbart, 2007). The ability to regulate and direct attention releases the child from the constraints of only responding to environmental events, and means they are able actively to guide their attention towards the information-rich areas key for learning (Ruff & Rothbart, 1996; Scerif, 2010). Individual differences in early orienting of attention and sustained attention have been shown to correlate longitudinally with later attentional development, as well as with other aspects of cognitive and behavioural functioning (Ruff, 1990). For example, individual differences in aspects of attentional control have been shown to correlate with early language development (Kannass & Oakes, 2008; Rose, Feldman, & Jankowski, 2009), with early learning in academic settings (Razza, Martin, & Brooks-Gunn, 2010; Rose, Feldman, & Jankowski, 2011; Steele, Karmiloff-Smith, Cornish, & Scerif, 2012; Welsh, Nix, Blair, Bierman, & Nelson, 2010) and with later hyperactive/impulsive behaviors (Lawson & Ruff, 2004a, 2004b).

Early disruptions to the development of attentional control may also disrupt the subsequent acquisition of other skills (Cornish, Cole, Longhi, Karmiloff-Smith, & Scerif, 2012b; Cornish, Sudhalter, et al., 2004; Cornish et al., 2007; Mulder, Pitchford, & Marlow, 2010; Scerif, Longhi, Cole, Karmiloff-Smith, & Cornish, 2012; Stevens, Lauinger, & Neville, 2009; van de Weijer-Bergsma, Wijnroks, & Jongmans, 2008). In fragile X syndrome, for example, attentional control of very simple responses is atypical from infancy (Scerif et al., 2005) into toddlerhood (Scerif, Cornish, Wilding, Driver, & Karmiloff-Smith, 2007; Scerif et al., 2004) and into childhood, both when mapped cross-sectionally (Cornish et al., 2007) and longitudinally (Cornish, Cole, Longhi, Karmiloff-Smith, & Scerif, 2012a). Recent longitudinal data suggests that early atypical attention to auditory and visual stimuli correlates longitudinally with later difficulties across cognitive functions and behaviour, including autistic symptomology (Cornish et al., 2012b; Scerif et al., 2012).

Summary

We have described *a priori* evidence suggesting that interventions that train attentional control and the closely overlapping domain of working memory ought to have more widespread effects the earlier the intervention is targeted. Firstly, we have argued that the neuroimaging evidence suggests that functional neural networks are more unspecialised and undifferentiated earlier in development. Secondly, empirical data and theoretical considerations from developmental disorders suggest that different domains may interact during development. Thirdly, attentional control in particular may be important as a 'tool for learning', gating the subsequent acquisition of skills in other domains.

With regard to all three points, however, it is important to note that the evidence we have reviewed is purely correlational in nature. Correlations, including longitudinal correlations (e.g. Rose et al., 2009), are insufficient demonstration of causality, since any observed relationship may be mediated by some unknown third factor. Even investigations using Structural Equation Modelling (e.g. Rose, Feldman, & Jankowski, 2011; Rose, Feldman, Jankowski, & Van Rossem, 2008) are vulnerable to the possibility that confounding variables have not been included in the model. Conclusive proof that domains interact during development requires an experimental study to establish a counterfactual dependence – showing, for example, that training attentional control improves the subsequent acquisition of new skills in other areas (e.g. vocabulary acquisition).

The current review aims, therefore, to evaluate the evidence of observed training transfer following cognitive training and to assess whether cross-domain effects operate differentially over developmental time. Our hypothesis is there will be a relationship between the degree of improvement found at post- vs. pre-testing and the age of the participants involved in study. Specifically, we predict that studies targeting younger participants ought to report relatively more widespread transfer of training effects.

Different methods for training attentional control

In the analyses that follow we concentrate on two strands of research within cognitive training. First, we discuss studies that have administered a mixed training regime targeting a battery of sub-components of attention (mixed attention (MA)); second, we discuss studies that explicitly trained working memory (WM). Both groups are substantially heterogenous on multiple dimensions – on the exact nature of the training that was administered, on the amount of training that was administered, on the age group of the participants, and on whether the participants were typically developing or members of a disorder group.

Mixed attention (MA) training

The majority of studies in this category administered a mixed training battery that includes a number of partially complementary different training tasks targeting one or a combination of the following cognitive domains: sustained attention, selective attention, task switching and inhibition. For example, Wass, Porayska-Pomsta, and Johnson (2011) administered to typically developing 11-month-old infants a battery of training tasks targeting different subcomponents of attentional control including task switching, inhibition, focused attention and working memory for objects embedded in complex scenes. The tasks were administered using an eyetracker and a gaze-contingent interface. They found that a short training period (average 77 min) led to improvements relative to an active control group in cognitive flexibility, sustained attention, and to reduced saccadic reaction time latencies, although they found no improvements post-training on working memory. Rueda, Rothbart, McCand-liss, Saccomanno, and Posner (2005) used more conventional point-and-click software to train 4- and 6-year-old participants using a battery of tasks targeting object tracking, anticipation, stimulus discrimination, conflict resolution and inhibitory control; they found some transfer to reasoning tasks but no significant changes to performance on the Attention Network Test or Childhood Behavior Questionnaire. Studies have also applied a mixed training battery to older adults (Schmiedek, Lovden, & Lindenberger, 2010) and to adults with schizophrenia (Greig, Zito, Wexler, Fiszdon, & Bell, 2007; Haut, Lim, & MacDonald, 2010; Lopez-Luengo & Vazquez, 2003; Medalia, Aluma, Tryon, & Merriam, 1998) and acquired brain injury (Sohlberg, McLaughlin, Pavese, Heidrich, & Posner, 2000). These have reported some limited improvements at tasks such as WM (Greig et al., 2007; Haut et al., 2010; Schmi-edek et al., 2010) and cognitive flexibility (Greig et al., 2007; Lopez-Luengo and Vazquez, 2003).

We have also included in this category studies that looked at transfer following training of one particular task, such as dimensional card-sorting or go/no-go. The strongest results in this area come from the three studies that trained participants using adapted versions of dimensional card sorting tasks (Karbach & Kray, 2009; Kerns, Eso, & Thomson, 1999; Kloo & Perner, 2003). Karbach and Kray (2009; Kray, Karbach, Haenig, & Freitag, 2012) used computerised card sorting to compare the effects of training children (8–10), young adults (18–26) and older adults (62–76), and reported transfer to other executive tasks and fluid intelligence across all age groups. Thorell, Lindqvist, Nutley, Bohlin, and Klingberg (2009) trained inhibition (a go/no-go paradigm) in 4–5-year-old children and found that transfer was weaker than that observed in a matched group that had received WM training. In their discussion they mention that this may be because inhibition of a prepotent response is presumably a short neural process occurring over milliseconds, while tasks such as working memory require sustained activity in both parietal and prefrontal areas over a time-course of seconds (Goldman-Rakic, 1995).

Some of these studies have additionally assessed whether the effects of cognitive training can be detected not just using lab-based cognitive assessment techniques but also using clinician, teacher or parent-administered assessments of real-world behaviors. For example, two adequately controlled studies have reported that computer-administered MA training can lead to reductions in some but not all measures of symptom severity in ADHD. Kerns et al. (1999) trained 7–11-year-olds at dimensional card sorting, and reported reductions in some (but not other) ratings of ADHD symptom severity. Rabiner, Murray, Skinner, and Malone (2010) trained 6–7-year-olds rated by their teacher as having attention difficulties using two commercial MA training programs targeting mixed attention and

reported significant improvements in behaviour. Both of these studies also looked at the effect of training on academic performance: Kerns et al. (1999) reported improved maths performance in an ADHD group following dimensional card sorting training, and Rabiner et al. (2010) reported improvements at some measures of academic performance in a group of children identified as having attention difficulties by their teachers following mixed attention training.

Surprisingly few studies in this area have looked at whether the effect of training persists immediately beyond the cessation of training. Rabiner et al. (2010) found no significant improvements at their 6-month follow-up, which they attributed to a reduction in inattentive symptoms across their whole sample. In a large-scale study looking at the effects of training on cognitive decline with ageing, Willis et al. (2006) reported good maintenance of the effects of cognitive training at 5-year follow-up, although the effects of training remained relatively specific to the domain being trained.

Working memory (WM) training

Studies included in this category are substantially more homogenous than those in the last, primarily because the majority of studies have used the CogMed package. This is a commercially available cognitive training package originally developed by the Klingberg group (Klingberg, 2010; Klingberg, Forssberg, & Westerberg, 2002; Klingberg et al., 2005). This software administers visuo-spatial and verbal working memory training that finds a participant's maximum threshold (e.g. number of locations of objects that can be remembered) and administers tasks aimed to increase that threshold. Other studies have administered different forms of working memory training, such as single and dual *n*-back tasks (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Jaeggi et al., 2010). Willis et al. (2006) explicitly trained participants at using mnemonic strategies for remembering verbal material.

WM training has been repeatedly shown to lead to improvements on measures of reasoning/fluid intelligence (Jaeggi et al., 2008, 2010; Klingberg et al., 2002, 2005), although other groups have reported negative or only trend improvements (Chein & Morrison, 2010; Dahlin, Nyberg, Backman, & Neely, 2008; Holmes, Gathercole, & Dunning, 2009; Thorell et al., 2009; Westerberg et al., 2007). WM training has also been shown to generalise to other measures of attentional control, such as Stroop (Chein & Morrison, 2010; Klingberg et al., 2002, 2005), although again not universally (Thorell et al., 2009; Van der Molen, Van Luit, Van der Molen, Klugkist, & Jongmans, 2010; Westerberg et al., 2007), as well as to measures of sustained attention (Lundqvist, Grundstrom, Samuelsson, & Ronnberg, 2010; Thorell et al., 2009; Westerberg et al., 2007), task switching (Lundqvist et al., 2010) and go/no-go (trend – Thorell et al., 2009). Reports of transfer to processing speed are rarer, with three of four studies (Dahlin et al., 2008 grp a and b; Klingberg et al., 2002 grp b) reporting negative results. Schweizer, Hampshire, and Dalgleish (2011) looked at the differential transfer effects in adults following WM training using either emotional or neutral material and found that only training with emotional material yielded transferable gains to improved control over affective information on an emotional Stroop task.

Klingberg et al. (2005) reported significant improvements on parent (and trend on teacher) ratings of ADHD symptom severity following 475 min of WM training, that were robust (decreasing only slightly) at 3-month follow-up. Given that drug efficacy stops shortly after cessation of treatment in ADHD (Halperin & Healey, 2011) the finding of a follow-up effect here is promising; however, this result has not, to our knowledge, been successfully replicated.

Well-controlled reports of improved academic performance following WM training are still patchy. Van der Molen et al. (2010) reported improved maths ability in a group with mild to borderline intellectual disability at 10-week follow-up. Chein and Morrison (2010) reported that WM training led to improved performance in a reading test in adults and St Clair Thompson (2007) found significant improvements in 7-year-old children at following instructions following training, as did Holmes et al. (2009). A number of other groups (Holmes et al., 2009; van der Molen et al., 2010) have, though, reported non-significant results on school-age children immediately following training. Holmes et al. (2009) reported improvements at 6-month follow-up on standardised math test performance, although they did not re-assess the control group at that visit.

Medium-term maintenance of training improvements has been assessed by a number of WM training studies between 10 weeks and 18 months after the cessation of training (Dahlin et al.,

2008; Klingberg et al., 2005; Lundqvist et al., 2010; St Clair Thompson, 2007; van der Molen et al., 2010; results described in Column M of the Table). The effects of one period of WM training appear to get progressively weaker over time, although significant effects of training have been identified on some measures even as far as 18 months after the cessation of training (Dahlin et al., 2008).

Other cognitive training methods

Although in the quantitative analyses that follow we only include the cognitive studies targeting WM and MA described above, in this section we also provide a brief overview of other forms of cognitive training that behavioural or neuroimaging evidence suggests may be targeting similar attentional control systems. The training methods described below were not included in our quantitative analyses for two principal reasons: either because they administer training that is cognitively heterogenous, in which attentional control may be targeted only indirectly, or because they have not yet produced a body of work that is large enough to allow for quantitative analyses.

Non-verbal reasoning training

Bergman Nutley et al. (2011) administed computerised training targeting non-verbal reasoning in typically developing 4-year-old children and reported transfer to Gf relative to a placebo trained group. Owen et al. (2010) administered a large, home-based training study to adults (mean (*sd*) 41 (*11.8*) years) for which one group received training at reasoning, planning and problem-solving tasks, and reported no transfer following training. Willis et al. (2006) administered non-verbal reasoning training to older adults and found within-domain training improvements that were robust at 5-year follow-up but no transfer of training improvements.

Mindfulness meditation

Lutz, Slagter, Dunne, and Davidson (2008) classify mindfulness meditation (MM) training in two stages: focused attention and open monitoring. Focused attention involves a number of faculties: sustained attention to a target object, disengaging from a distracting object (attention switching) and redirecting focus promptly to the chosen object (selective attention) (Lutz et al., 2008). A number of researchers in cognitive neuroscience have started to take an interest in MM in recent years (Brefczynski-Lewis, Lutz, Schaefer, Levinson, & Davidson, 2007; Cahn & Polich, 2006; Chiesa, Calati, & Serretti, 2010; Jang et al., 2010; Lutz, Greischar, Rawlings, Ricard, & Davidson, 2004; Lutz et al., 2008; Rubia, 2009; Slagter et al., 2007; Tang & Posner, 2009), as studies have suggested links between MM and performance on tasks such as the Attention Network Test (Jha, Stanley, Kiyonaga, Wong, & Gelfand, 2011; van den Hurk, Giommi, Gielen, Speckens, & Barendregt, 2009), the Stroop task (Chan & Woollacott, 2007) and working memory (van Vugt & Jha, 2011). Tang et al. (2007, 2009, 2010) used diffusion tensor imaging, single photon emission computed tomography and electroencephalography (EEG) to point to changes in activity of the anterior cingulate cortex (ACC) and other frontal areas after MM training; these are areas that are also implicated in more traditional executive attention and working memory tasks (Bush, Luu, & Posner, 2000; Duncan & Owen, 2000; Kelly et al., 2009; Pessoa, 2008; Posner & Rothbart, 2009). Relatively few studies have applied mindfulness meditation to younger participants, perhaps because it requires a level of meta-cognitive awareness; however, Napoli, Krech, and Holley (2005) applied 12 sessions of mindfulness meditation training to a group of 114 children aged 6–9 years and found, relative to a wait-list control group, that training led to significant improvements on the Test of Everyday Attention on the selective but not the sustained attention subcomponents, to some changes on the ADD-H Comprehensive Teacher Rating Scale and on the self-reported Test Anxiety Scale (see also Zylowska et al., 2008).

Action computer games

A series of studies (Boot, Kramer, Simons, Fabiani, & Gratton, 2008; Castel, Pratt, & Drummond, 2005; Chisholm, Hickey, Theeuwes, & Kingstone, 2010; Dye, Green, & Bavelier, 2009; Green & Bavelier, 2007, 2008; Mishra, Zinni, Bavelier, & Hillyard, 201; reviewed in Bavelier, Green, and Dye (2010)) have shown that regular action computer game players show superior performance on visual attention tasks such as useful field of view and attentional blink. These correlational findings may, however,

arise because people who are better at these kinds of tasks are drawn to these kinds of computer games. Only a small number of studies have looked at transfer after non-video-game players were trained using action video-games. Some studies have reported that training leads to improvements on aspects of visual attention (Green & Bavelier, 2003, 2007). However, Boot et al. (2008) assessed the effect of c. 20 h of video-game playing in young adult non-video-game players and looked for transfer to a number of other executive tasks such as working memory, task switching and reasoning, with strikingly negative results. Perhaps surprisingly, no study to our knowledge has assessed similar transfer of training improvements using younger participants.

Neurofeedback

During neurofeedback (Arns, de Ridder, Strehl, Breteler, & Coenen, 2009; Cannon, Congedo, Lubar, & Hutchens, 2009) the participant wears an EEG net and views a computer feedback of their electrical brain activity. The subject is encouraged to learn to 'control' this readout by receiving coaching from a clinician on maintaining effort and focus using metacognitive strategies. Neurofeedback has been shown to be associated with changes in neural activity in areas such as the ACC and dorsolateral pre-frontal cortex (Beauregard & Levesque, 2006; Cannon et al., 2009). A number of controlled studies have demonstrated improved performance on standardised cognitive assessment measures following neurofeedback training at WM (Cannon et al., 2009; Vernon et al., 2003), Stroop (Beauregard & Levesque, 2006; Kouijzer, de Moor, Gerrits, Congedo, & van Schie, 2009), go/no-go (Beauregard & Levesque, 2006), sustained attention (Kouijzer et al., 2009) and cognitive flexibility (Kouijzer et al., 2009). However, the widespread use of passive control groups in these studies means that caution should be exercised in interpreting these findings (see Methods section). A number of neurofeedback papers have also reported reductions in attention deficit hyperactivity disorder (ADHD) symptom severity following training. Arns et al. (2009) have recently reviewed this literature, although they include a number of papers that are inadequately controlled.

Clinician-/parent-/teacher-administered interventions

A number of behavioural intervention packages explicitly include an executive or attentional control training component administered within a social context, normally via interaction either with parent or teacher. For example, the Tools of the Mind package is a classroom-based curriculum aimed to boost the development of the executive functions (Barnett et al., 2008; Blair & Diamond, 2008; Diamond, Bernett, Thomas, & Munro, 2007). Within the field of clinician-administered ADHD interventions, a number of programs now also include elements of structured working memory and turn-taking training with the primary caregiver (Sonuga-Barke & Halperin, 2010; Sonuga-Barke et al., 2011; Thompson et al., 2009).

Methods

In the analyses that follow we focus exclusively on studies that applied Mixed Attention (MA) and Working Memory (WM) training.

Search strategy

All searches were carried out on 27th February 2012 by two separate raters, one of whom was blind to the purposes of the review. An *a priori* decision was made only to include studies dating from 1990 to the present, and only to search for studies published in English. These were the only filters that were applied. The search terms used were: 'cognitive training' (167 titles returned from Medline), 'cognitive rehabilitation (234), 'executive control (253)', 'task switching' (310), 'attention training' (40), 'working memory training' (44), and 'inhibition training' (4); the abstracts for studies which contained those search terms in the title were inspected. All studies losely fitting our inclusion criteria were downloaded, and where unavailable, full texts were requested from the corresponding authors. Since the sensitivity of searches using Medline alone has been reported to be low (Adams, Power, Frederick, & Lefebvre, 1994), additional databases were also used for the search, namely ISI Web of

Knowledge and Google Scholar. Identical search terms were entered into these search engines and possibly relevant articles were downloaded. Reference lists and bibliographies were searched from all retrieved articles and relevant published reviews. Papers were read by each of the two raters and a consensus decision was taken as to whether the data contained material suitable for the analysis, based on the inclusion criteria below. To our knowledge no study fitting the inclusion criteria below has been excluded from this review.

Inclusion criteria

To be included in this review, studies must be: (1) published in a peer-reviewed journal; (2) apply training primarily designed to require the use of one or several of the following cognitive domains: Working Memory; Mixed Attention, defined as including the terms: Sustained Attention; Selective Attention; Task Switching; Inhibition. (3) include at least a passive control group (see Controls section below); (4) apply testing to both trained and control groups before and after training that includes at least one standardised assessment that measures either executive attention or working memory (see Defining transfer section below).

Controls

A number of studies (including several of those reporting strong transfer following training (Holmes et al., 2010; Kerns, MacSween, Wekken, & Gruppuso, 2010; Tamm et al., 2008; Zylowska et al., 2008)) do not include any control groups so have been excluded from this review. In those studies we have included, some have used a passive (wait-list) control group (in which control participants attend only two testing sessions, at pre- and post-test) whereas others have used an active control group (in which control participants attend for the same number of sessions and participate in *ersatz* training sessions).

We consider active-controlled studies to be substantially preferable. Although passive-controlled studies control for test–retest effects and for the passage of developmental time, they do not (in most cases) control for the placebo effect (since participants in the trained group knew that they had participated in some form of training, whereas those in the control group knew that they had not). There exists a large body of evidence (see e.g. Beauregard, 2007) suggesting that expectation of improvements after training can substantially alter performance at post-test. It also does not control for the effect of repeated exposure to experimenters/lab settings. Beauregard and Levesque (2006), for example, report a study in which one group of 8–12-year-olds with ADHD attended the lab for 40 sessions in total whereas the other group attended only twice – at pre- and post-testing.

Statistical analyses reported

The studies we review have administered a variety of statistical analyses (described in Column K of the table). Of these, the two most commonly used are:

(i) Repeated measures ANOVAs with group (trained vs. control) as the between-subjects factor, visit (pre vs. post-test) as the within-subjects factor and test result as the dependent variable.
(ii) Analyses of covariance (ANCOVAs) with group (trained vs. control) as the between-subjects factor, difference scores on the test result (post–pre-test) as the dependent variable and pre-test performance as the covariate. Assuming that the trained and control groups are randomly assigned, the advantage of the ANCOVA is that it controls for differences between the scores of the two groups at pre-test (Dimitrov & Rumrill, 2003; Jamieson, 2004; Miller & Chapman, 2001).

Pre–post assessments

The assessments included in our analyses were generally administered immediately after the cessation of training. In some cases (9/37 studies) the initial post-assessment was followed by a subse-

quent reassessment session administered between 10 weeks and 5 years after the cessation of training to assess the longer term maintenance of training improvements. Where applied, these follow-up assessments have been described in Column M of the table. However, the small number of studies that admininstered a follow-up assessment together with the variability of the date of administration post-training mean that it has not been possible to include these assessments in our quantitative analyses.

Defining transfer

Assessing whether training improvements have generalised from the specific task being trained to other tasks is a conceptually complicated area where no clear benchmarks apply (see discussions in Klingberg, 2010; Loevden, Backman, Lindenberger, Schaefer, & Schmiedek, 2010; Shipstead, Redick, & Engle, 2010).

To allow standardised comparison of the efficacy of training across a number of training studies, we first classified each of the pre–post tests as assessing a particular cognitive faculty. The categories we have used are: Working Memory; Sustained Attention; Selective Attention; Task Switching; Inhibition; Reasoning; Response Speed; Academic Performance; Emotional Control (including standardised ratings of symptom severity in ADHD); Language and Social; Other (including other standardised assessments).

Where possible we have tried to follow the descriptions used by the authors of the studies, although in cases where there is disagreement we have used the most common definition. To promote transparency we have included, in Column L of the table, the exact name of each pre–post that was administered, followed by our classification of the cognitive faculty that we considered that test to be assessing. A full list of all the tests classified by cognitive faculty, together with the abbreviations that have been used, is given in the table legend.

For the present review we have adopted a stringent definition of transfer. Rather than assessing the efficacy of training of individual constructs themselves (e.g., by assessing whether WM or attentional control can be modified by training), we wished to assess whether these improvements generalise to other cognitive functions. Therefore if the pre–post test was, according to our classification, in the same category as the training stimulus, we have excluded the result of this test from our quantitative analyses. For example, Jaeggi et al. (2010) administered WM training (using single and dual n-back tasks) and assessed the effect of training by administering pre- and post-training two fluid intelligence tasks (Bochumer Matrices and Raven's Advanced Progressive Matrices) and a different assessment of working memory (the operation span task). In our analysis we have included the results of the fluid intelligence tasks but excluded the results of the working memory task, since these are judged to be in the same category as the training that was administered and therefore not to be assessments of distal transfer. It should be noted that, particularly within the WM training literature, this has led to many tests that are described by their authors as assessments of "near" transfer being excluded from our quantitative calculations. The results of these additional "near transfer" assessments have, however, been included (in square brackets) in Column L of the Table, together with the significance of the results observed.

Calculation of effect sizes and confidence intervals

Where possible we have calculated Cohen's d based on the raw means and standard deviations using the standard formula:

$$d = \frac{M_E - M_C}{s}$$

where M_E is the change in means (post–pre) in the experimental (i.e. trained) group, and M_C is the change in means in the control group, and s is the pooled standard deviation. Where the raw means and standard deviations were not available, we have based our calculation on the reported p value and sample size using the procedures suggested by Hunter, Schmidt, and Jackson (1982) and Hunter and Schmidt (1990) (available online at http://www.lyonsmorris.com/ma1/index.cfm). We have also reported the statistical analysis from which this p value was derived in Column K of the table.

We then calculated the confidence intervals (CIs) using the method given by Hedges and Olkin (1985):

$$\sigma[d] = \sqrt{\frac{N_E + N_C}{N_E \times N_C} + \frac{d^2}{2(N_E + N_C)}}$$

where N_E and N_C are the numbers in the experimental and control groups, respectively. The 95% confidence interval has been calculated as $d \pm 1.96\sigma$.

For some studies neither the raw means nor the p values/sample sizes were available. (These studies typically report just that a test was administered, but that the results were not significant.) Since excluding these measures would have lead to a positive bias, we included these tests as reporting an effect size of 0.

The pooled effect sizes and CIs for individual studies can be seen in Column N of the Table and the effect sizes and CIs for the individual pre–post tests can be seen in Column N of the Table.

Results

The studies fitting our inclusion criteria are shown in Fig. 1 and Table 1.
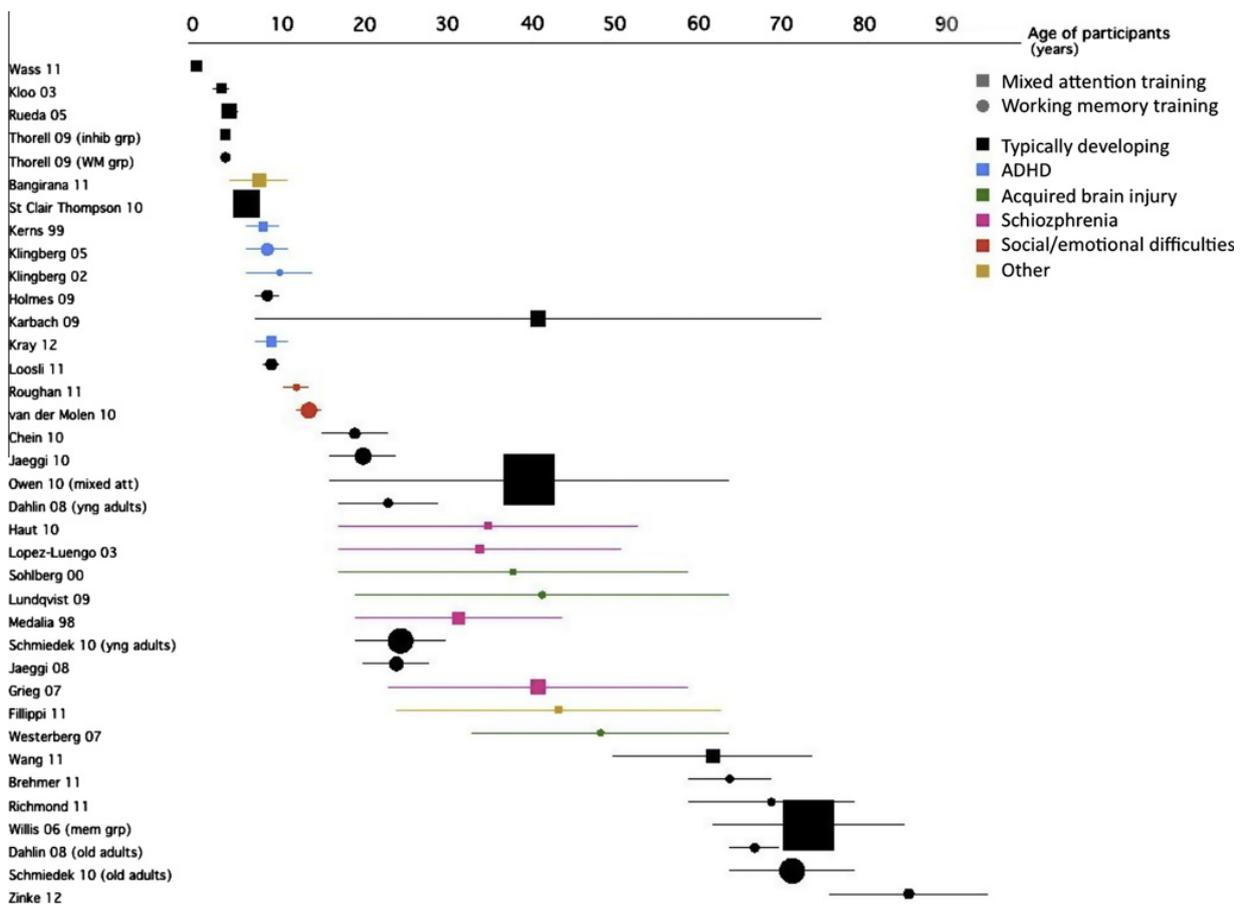
Fig. 1. A graphical representation of the studies included in this review. Studies have been listed by the age of the participants trained. The age of the participants is represented on the *x*-axis. The error bars span from the youngest participant included in each study to the oldest. The size of the central dots represents the number of participants included in each study; a bigger dot indicates a larger participant pool. For Owen et al. (2010) and Willis et al. (2006) these have not been drawn to scale to promote legibility.

Quantitative analyses

Our primary research question was that of whether cognitive training targeted at younger participants tends to report more widespread transfer of training effects. Addressing this question was complicated by the fact that the studies we are reporting are heterogenous on multiple independent variables, namely:

– *Amount of training applied.* This ranges from 30 min (Kloo & Perner, 2003) through to 6000 min (Schmiedek et al., 2010).
– *Exact type of training applied.* On the figures and table we have grouped the studies we are reviewing into two categories, WM and MA, although as we discuss above both categories are substantially heterogeneous.
– *Exact population targeted.* We have classified the studies we have included as targeting six particular populations: typically developing (TD); ADHD; acquired brain injury; schizophrenia; and individuals with social and emotional difficulties; other.

Figs. 2 and 3 show forest plots of the effect sizes and 95% confidence intervals of the individual pre–post tests that were administered.

The effect sizes and variances have then been pooled across all of the different tests administered to give a single estimate of the average degree of transfer observed per study, together with an estimate of the degree of variance observed between pre–post tests. These data have been listed in Column N of the table.

Exploratory analyses conducted to examine the relationship between degree of transfer reported by individual studies and the age of participants in that study revealed that both variables were not parametrically distributed (age–skewness: 1.33, kurtosis: 0.52; transfer–skewness: 1.66, kurtosis: 5.45) and therefore non-parametric correlations (Spearman's rho) have been calculated. When all studies were pooled together a relationship between transfer and age was identified ($r(1,43) = -.31$, p(2-tailed) = .045). When amount of training time administered was included as a covariate this relationship became stronger ($r(1,40) = -.37$, p(2-tailed) = .015). This suggests that the observed relationship between transfer and age was confounded (weakened) by the fact that smaller amounts of training tended to be applied to studies targeting different age groups.

When just those studies targeting typical development are considered a correlation between age and transfer score observed is also identified ($r(1,28) = -.47$, p(2-tailed) = .011), which again becomes stronger when amount of training administered is controlled for ($r(1,25) = -.53$, p(2-tailed) = .005). The results also suggested that studies targeting certain disorder group populations (e.g. children with ADHD and adults with acquired brain injury) may tend to report more transfer than studies targeting typically developing participants. However there are insufficient studies available to us to allow us to calculate the relationship between transfer reported and age independently for the different disorder groups.

The relationship observed between age and training transfer was, although consistent, a relatively weak relationship. This is shown by the fact that when the WM and MA groups are considered individually, the smaller number of studies included in the correlation means that the relationship is non-significant for both groups (WM – $r(1,19) = -.45$, p(2-tailed) = .053; MA – $r(1,24) = -.20$, p(2-tailed) = .35).

Discussion

The quantitative analyses we conducted have identified significant relationships between the degree of transfer reported by individual studies and the age of participants in that study. The initial significant relationship we identified ($r = -.31$) became stronger when the amount of training time was included as a covariate ($r = -.37$), because studies targeting younger individuals tend to administer smaller amounts of training. When just those studies that had applied training to typically developing individuals were considered the relationship was found to be stronger ($r = -.47$), which again increased when the amount of training time was included as a covariate ($r = -.53$). The relationship appeared to be stronger for WM ($r = -.45$) than for MA ($r = -.20$) studies.

Table 1
Summary of studies included in this review

A and B – Authors and year	C – Description of participants	D – Age of participants (range, or mean (sd))	E – Control	F – Amount of training	G – Mins training	H – N trained	I – Nature of training	J – Training group (1 = WM, 2 = MA)	K – Statistical analyses	L – Pre- and Post (see table legend)	M – Follow-up (time since completion of training)	N – Average transfer observed across all pre–post tests administered	O – Std in transfer observed across all pre–post tests administered
Bangirana et al. (2011)	Previously infected with malaria	5–12-y-o	p	16 × 45 min sessions over 8 weeks	720	28	Mixed Att/WM – Captain's Log	2		TOVA (sus); KABC (beh gen); CBC (beh gen); WRAT (beh gen)		0.15	0.21
Brehmer et al. (2011)	Ageing	60–70-y-o	Received non-adaptive training (i.e. no WM load)	25 × 25-min sessions over 5 weeks	625	12	WM – mixed	1	RM ANOVA just on trained group	PASAT (sus) (?); RPMP (reas) (?); Stroop (inh) (?); [RAVLT (WM); SB (WM) DS (WM)]		0.21	0.66
Chein and Morrison (2010)	TD	20-y-o	p	5 sessions per week for 4 weeks (do not say how long)	Do not report	21	WM – mixed	1	Separate paired sample t-tests for T and C groups; independent samples t-test on gain scores; RM ANOVA	ETS (reas) (s); Stroop (selec) (y); RPM (reas) (n); NDRT (lit) (y); [WM (y)]	None	0.29	0.16
Dahlin et al. (2008) (old adults)	Ageing	68 (1)-y-o	p	Three 45-min sessions per week for 5 weeks	675	15	WM – mixed	1	RM ANOVA	RPM (reas) (n); WechPSI (spd) (n); COWAT (lang) (n); [Episodic memory (n); verWM (n); vsWM(n)]	18-m - verWM (y)	0.14	0.19
Dahlin et al. (2008) (yng adults)	Ageing	24 (3)-y-o	p	Three 45-min sessions per week for 5 weeks	675	15	WM – mixed	1	RM ANOVA	RPM (reas) (n); WechPSI (spd) (n); COWAT (lang) (n); [Episodic memory (s); verWM (s); vsWM(s)]	18-m - verWM (y); vsWM (y); Episodic memory (n)	0.44	0.55
Filippi et al. (2011)	Relapsing–remitting multiple sclerosis	25–64-y-o	p	Three 60-min sessions per week for 12 weeks	2160	10	Mixed attetnion, information processing and executive functions	2	Hierarchic liner mixed model	PASAT (sus) (y); WCST (flex) (y); COWAT (lang) (s); SDC (spd) (n); TEA (beh gen) (n); SRT (beh gen) (n)	None	0.55	0.64
Greig et al. (2007)	Schizophrenia	42 (9)-y-o	?p	1 year	?	33	Various EF – Programs from SciLearn and CogRehab and others	2	MANCOVA	WechWM (s); WCST (switch) (y); HT (soc) (n); BLERT (soc) (n)	None	0.17	0.16
Haut et al. (2010)	Schizophrenia	36 (9)-y-o	Received cognitive behavioural social skills training	Up to 25 h of training over 4–6 weeks		10	Mixed Att/Mem – CogPack Marker Software	2	RM ANOVA	vsWM (y); verWM (s) [Lexical decision (n)]	None	0.89	0.12
Holmes et al. (2009)	TD	8–11-y-o	Same but task didn't respond adaptively	20 sessions – approximately 35 min a day in school for at least 20 days in a period of between 5 and 7 weeks	700	22	WM – mixed	1	ANOVA where no sig diff at pre; ANCOVA where sig diff at pre	WechReas (reas) (n); WechRead (lit) (n); WechMath (math) (n); FI (lang) (y); [verWM (y); verSTM(y); vsWM (y); vsSTM(y)]	(NB control group not re-tested at follow-up) 6-m - WechReas(n); WechRead (lit) (n); WechMath (math) (y); FI (lang) (y); (verWM(y); verSTM(n); vsWM(y); vsSTM(y))	0.29	0.53
Jaeggi et al. (2010)	TD	21 (2.2)-y-o	p	20 × 18 min sessions	405	46 (2 different groups)	WM – single and dual n-back WM tasks	1	RM ANOVA	RPM (reas) (y); BOMAT (reas) (y) [OSPAN (WM) (y)]	None	0.40	0.00

Study	Population	Age	Control	Schedule	min	N	Domain	#	Analysis	Outcomes	Follow-up	ES1	ES2
Jaeggi et al. (2008)	TD	25-y-o	p	9–18 sessions	360	34 (4 different groups, received different amounts of training)	WM – vis-spat dual *n*-back task	1	ANOVA; ANCOVA	RPM (reas) (y); BOMAT (reas) (y)	None	0.54	0.01
Karbach and Kray (2009)	TD	8–10, 18–26, 62–76	Performed same sorting task, but without have to switch sort criteria	4 sessions c. 1 per week	200	14 per group	Switch – sorting according to changing criteria	2	4-way ANOVA (age × training session × pre–post × group)	verWM (y); vsWM (y); Stroop (selec) (y); RPM (reas) (y) [NB not possible to derive separate measures for different age groups]	None	0.48	0.07
Kerns et al. (1999)	ADHD	7–11-y-o	Played video games	16 sessions twice a week for 30 min over 8 weeks	480	14	Switch – Pay attention program – card-sorting according to different criteria	2	ANCOVA	WechWM (n); WechReas (reas) (s); ACT (sus) (y); underlining (sus) (s); CPT (sus) (n); Stroop (selec) (y); VOT (reas) (n); arithmetic (math) (y); MFFT (impuls.) (s); ADDES (impuls.) (n)	None	0.51	0.96
Klingberg et al. (2005)	ADHD	7–12-y-o	Identical to the treatment except that the WM trials remained on the initial low level instead of being increased to match the WM span of the child	>20 days	475	27	WM – visuospatial WM	1	Post-test scores compared using a general linear model, controlling for age, training time and pre-test performance. Equivalent to ANCOVA	Stroop (selec) (y); RPM (reas) (y); motor activity (beh gen) (n); DSM_ADHD (inatt) (s); CRS (inatt) (s); [WechWM (vsWM, verWM) (y)]	3-m - Stroop (selec) (n); RPM (reas) (n); motor activity (n); DSM_ADHD (inatt) (s); CRT (inatt) (s); (WechWM (vsWM, verWM) (y);)	0.71	0.89
Klingberg et al. (2002)	ADHD	7–15-y-o	Same training, no WM component	24 days, 25 min per day over 5–6 weeks	600	7	WM – mixed	1	Paired sample *t*-test for T group; independent samples t-test on gain scores	Stroop (selec) (s); RPM (reas) (y); Reaction time (spd) (s); motor activity (beh gen) (y); [vsWM (y)]	None	1.64	1.04
Kloo and Perner (2003)	TD	3–5-y-o	Group trained at number conservation tasks or relative clauses	2 sessions (15 min per session) over 2 weeks	30	14	Flex – Card sorting task (similar to WCST)	2	ANOVA	False belief (soc) (y); [card-sorting (switch) (y)]	None	0.34	0.33
Kray et al. (2012)	ADHD	8–12-y-o	Performed same sorting task, but without have to switch sort criteria; multiple baseline design	4 × 35 min sessions	200	20	Switch – sorting according to changing criteria	2		Stroop (inh) (y); verWM(n); WISC (reas) (n); DSST (spd) (n)	Administered another 4 training sessions and repeated the assessments: Stroop (inh) (n); verWM(n); WISC (reas) (n); DSST (spd) (n)	0.42	0.59
Loosli et al. (2011)	TD	9–11-y-o	p	5 sessions (12 min per session) per week for 2 weeks	120	24	WM span	1	MANOVA - group × diff score	TONI (reas) (n); SLT (lit) (s)		0.31	0.21
Lopez-Luengo and Vazquez (2003)	Schizophrenia	35 (8.5)-y-o	Received no ATT, but participated in the same multidisciplinary day-treatment program	c. 50 sessions of c. 40mins per session, spread over c. 45 weeks	2000	13	Mixed Att – attention process training (APT) (sus, selec, alternating and divided attention)	2	RM ANOVA	CPT (sus) (n); Cancellation (sus) (n); Dichot. listen (selec) (y); PASAT (sus) (n); TMT (selec)(n); VLT (verWM) (n); WCST (switch) (y); MMSE (beh gen) (n); BPRS (beh gen) (n); EAQ (beh gen) (n); GAF (beh gen) (n); SAPS (soc) (s); SANS (soc) (n)	None	0.14	0.34

Table 1 (*continued*)

A and B – Authors and year	C – Description of participants	D – Age of participants (range, or mean (sd))	E – Control	F – Amount of training	G – Mins training	H – N trained	I – Nature of training	J – Training group (1 = WM, 2 = MA)	K – Statistical analyses	L – Pre- and Post (see table legend)	M – Follow-up (time since completion of training)	N – Average transfer observed across all pre–post tests administered	O – Std in transfer observed across all pre–post tests administered
Lundqvist et al. (2010)	Acquired brain injury	20–65-y-o	p	60 min per day, 5 days per week, for five weeks (repeated exactly same program before 6-m follow-up)	1500	10	WM – mixed	1	Paired sample *t*-test	PASAT (sus) (y); CWIT (switch) (y); [WechWM (y); LST (verWM) (y); PS (vsWM) (y); self-report]	6-m (following more training): PASAT (sus) (y); CWIT (switch) (y); COPM (beh gen) (y); EQ-5D (beh gen) (n); (WechWM (y); LST (verWM) (y); PS (vsWM) (y); self-report (y))	0.54	0.00
Medalia et al. (1998)	Schizophrenia	20–45-y-o	Same number of visits, viewed National Geographic documentaries	18 sessions – three 20-min sessions per week for 6 weeks	360	27	Mixed Att	2	Wilcoxon signed rank tests on gain scores	CPT (sus) (y); BPRS (beh gen) (y)	None	0.22	0.11
Owen et al. (2010) (mixed att grp)	TD	41 (11.8)-y-o	Researched general knowledge questions over the internet	Avg 4 (2.7) sessions (>10 min) per week for 6 weeks	240	4014	Mixed Att – tests of memory, attention, visuospatial processing and mathematical calculations	2	Means and 95% CIs	vsWM (n); verWM (n); PAL (WM) (n); grammatical reasoning (Reas) (n)	None	0.12	0.06
Richmond et al. (2011)	Cog decline with ageing	60–80-y-o	Completed a matched number of trivia training sessions	Five 30-min sessions per week for 4 weeks	600	21	vs WM and ver WM	1	ANOVA (group × session)	RPM (reas) (n); TEA (beh gen) (n); CVLT (lang) (s); ([RS (WM) (y) DS (WM) (n)]		−0.18	0.37
Roughan and Hadwin (2011)	Social, emotional and behavioural difficulties	13 (0.65)-y-o	p	Five 35-min sessions per week for 5 weeks	875	9	vs WM and ver WM	1	ANOVA on difference scores	go/no-go (inhib) (s); RPM/MHVS (reas) (y); TAS/SDQ (anx) (y); SDQ/Att (beh gen) (y) [DS (WM) (y) SS (WM) (y)]	3-m: go/no-go (inhib) (n); RPM/MHVS (reas) (y); TAS/SDQ (anx) (n);; SDQ/Att (beh gen) (n) [DS (WM) (y) SS (WM) (y)]	0.90	0.24
Rueda et al. (2005)	TD	4-/6-y-o	Brought into the lab for the same no. of sessions, watched children's videos	Five sessions (45 min per session, spread out over 2–3 weeks)	225	24 four-y-os (18 for ANT); 12 six-y-os	Mixed Att – tracking an object; anticipation; stimulus discrim; Stroop; inhib control	2	RM ANOVA	ANT conflict (switch) y; K-BIT (reas) (s); CBQ (beh gen) (n); [ANT alert (n); ANT orient (n)]	None	0.17	0.29
Schmiedek et al. (2010) (young adults)	TD	20–31-y-o	p	100 one-hour sessions	6000	101	Six tests of perceptual speed (PS), three tests of working memory (WM), and three tests of episodic memory (EM)	2	Mixed models and latent difference score models	AS (WM) n, 3B (WM) y; MU (WM) n; RS (WM) n; CS (WM) n; RS (WM) n; WP (WM) n; BISV (reas) n; BISN (reas) y; BISFS (reas) y; RPM (reas) n	None	0.20	0.24

Schmiedek et al. (2010) (older adults)	Cog decline with ageing	65–80-y-o	p	100 one-hour sessions	6000	103	Six tests of perceptual speed (PS), three tests of working memory (WM), and three tests of episodic memory (EM)	2	Mixed models and latent difference score models	AS (WM) y, 3B (WM) n; MU (WM) n; RS (WM) n; CS (WM) n; RS (WM) y; WP (WM) y; BISV (reas) n; BISN (reas) n; BISFS (reas) n; RPM (reas) y	None	0.23	0.15
Sohlberg et al. (2000)	Acquired brain injury	18–60-y-o	Received training in brain injury education and supportive listening.	24 h over 10 weeks	1440	7	Mixed Att - attention process training (APT) (sus, selec, alternating and divided attention)	2	ANOVA including various combinations of factors	TMT (selec) (s); PASAT (sus) (y); CPT (sus) (n); Gordon(sus) (n); COWAT (lang) (n); Stroop (selec) (s); Sternberg (WM) (s); BAFQ (beh gen) (n); BADS (beh gen) (n); [self-report]	None	0.33	0.40
St Clair Thompson (2007)	TD	7-y-o	p	12–16 × 30-min session (over 6–8 weeks)	420	117	WM – Memory booster – explicitly teaches mnemonic strategies	1	RM ANOVA	FI (lang) (y); various arithmetic (math) (s); GRT (lit) (n); MentMath (math) (n); [WMTB-C (vsWM) (s); verWM (s)]	5-m - various arithmetic (n); GRT (lit) (n); MentMath (math) (n)	0.50	0.46
Thorell et al. (2009) (inhib grp)	TD	4–5-y-o	Active group – played commercially available computer games, passive group only took part in pre- and post-testing	25 sessions – 5 weeks of 15 min per school-day	375	17	Inhibition	2	ANCOVA	Stroop (selec) (n); WechWM (visWM) (n); span board (vsWM) (n); CPT (sus) (n); WchReas (reas) (n) [go/no-go (inhib) (n)]	None	0.15	0.33
Thorell et al. (2009) (WM grp)	TD	4–5-y-o	Active group – played commercially available computer games, passive group only took part in pre- and post-testing	25 sessions – 5 weeks of 15 min per school-day	375	17	WM – visuo-spatial WM	1	ANCOVA	Stroop (selec) (n); go/no-go (inhib) (s); CPT (sus) (y); WchReas (reas) (n); [WechWM (visWM) (y); span board (vsWM) (y)]	None	0.35	0.45
van der Molen et al. (2010)	Mild to borderline intellectual disabilities	13–16-y-o	Identical but non-adaptive	3 × 6-min sessions per week for 5 weeks	90	41	WM – mixed	1	Post-test scores compared using a general linear model, controlling for pre-test performance. Equivalent to ANCOVA	Arithmetic (math) (n); Reading (lit) (n); Stroop (selec) (n); RPM (reas) (n); [verSTM (DR, NR) (s); visSTM (BR, VP) (n); verWM (BDR, LR) (n); SS (vsWM) (n); Story Recall (n)]	10-w - Arithmetic (math) (y); Reading (lit) (n); Stroop (selec) (n); RPM (reas) (n); [verSTM (DR, NR) (s); visSTM (BR, VP) (y); verWM (BDR, LR) (n); SS (vsWM) (y); Story Recall (y)]	−0.09	0.15
Wang et al. (2011)	Cog decline with ageing	63 (6)-y-o	p	5 sessions × 45 min per session	225	31	MA – computerised 'cooking' task	2	ANCOVA	WAIS WM (s); WAIS_Spd(spd) (s); WAIS_arith (arith) (n); MMSE (beh gen) (n)	None	0.23	0.06

Table 1 (continued)

A and B – Authors and year	C – Description of participants	D – Age of participants (range, or mean (sd))	E – Control	F – Amount of training	G – Mins training	H – N trained	I – Nature of training	J – Training group (1 = WM, 2 = MA)	K – Statistical analyses	L – Pre- and Post (see table legend)	M – Follow-up (time since completion of training)	N – Average transfer observed across all pre–post tests administered	O – Std in transfer observed across all pre–post tests administered
Wass et al. (2011)	TD	11-m-o	Active group – watched infant-friendly animations and videos for a matched program of training sessions	4 sessions (variable length)	77 (19.1)	21	Mixed Att/WM (eye-gaze contingent)	2	ANCOVA	Cog flex (y); Gap (spd) (y); Sus att (y); WM (n); spontaneous orienting during free play (beh gen) (s)	None	0.84	0.63
Westerberg et al. (2007)	Stroke	34–65-y-o	p	90 trials (40 min) per day, 5 days a week for 5 weeks	1000	9	WM – vis-spat and auditory	1	Post-test scores compared using a general linear model, controlling for pre-test performance. Equivalent to ANCOVA	RPM (reas) (n); PASAT (sus) (y); RUFF (selec) (y); Stroop (selec) (n); CFQ (beh gen) (y); [WechWM (y); CD (verWM) (s)]	None	0.49	0.42
Willis et al. (2006) (mem grp)	Cog decline with ageing	74 (6)-y-o		Ten 60–75-min sessions; booster training (four 75-min sessions) at 11 and 35 months	675	703	Memory – mnemonic strategies for remembering verbal material	1	Repeated-measures, mixed-effects model	Reas (n); IADL (gen beh) (n); other gen perf measures (gen beh) (s); [Memory (HVLT; RAVLT; RBPRT) (y); UsefulFoV (n)]	5-y - Memory (HVLT; RAVLT; RBPRT) (n); Reas (n); IADL (gen beh); other gen perf measures (gen beh) (n) [UsefulFoV]	0.29	0.51
Zinke et al. (2011)	Cog decline with ageing	77–96-y-o	p	5 sessions (30 min per session) over 2 weeks	300	20	vis-spat and ver WM	1	ANOVA	Strrop (selec) (n); RPM (reas) (n)		−0.05	0.21

Column E: Control group – p stands for passive (i.e. wait-list) control group.

Column F: Amount of training administered – excluding baseline assessment sessions.

Column H: N training – initial number allocated to training group, not including drop-out.

Column L: For each measure, we include: (i) the abbreviated name of the test administered, (ii) the abbreviated name of the cognitive faculty that the test is assessing, and (iii) the result of the test – either y,s or n. 'y' means a significant training improvement was observed, 's' means some training improvement was observed (either $p < 0.1$ on the core measure or significant improvement at some but not all subcomponents), and 'n' means no training improvement was observed. Thus, for example, 'RPM (reas) (n)' means that (i) Raven's Progressive Matrices was administered, (ii) that we have classified this test as an assessment of reasoning, and (iii) that no transfer was found to this assessment as a result of training. Measures in square brackets are excluded measures that have not been included in our graphical representations (Figs. 1–3) or in our calculation of between-domain transfer (see section 1.5).

Working Memory/Short-Term Memory (WM): Wechsler digit span/letter-number sequencing (WechWM); Operation Span Task (OSPAN); Working Memory Test Battery for Children (WMTB-C); Paired Associates Learning (PAL); Working Memory Index (WMI); Conceptual Span Task (CST); Wechsler Block Span (WechsWM); Picture Span (PS); Spatial Span (SS); digit span (DS); Rivermead Behavioral Paragraph Recall Test (RBPRT); Listening Span Task (LST); Verbal Learning Test (VLT); Claeson-Dahl word list recall (CD); Hopkins Verbal Learning Test (HVLT); Rey Auditory Verbal Learning Test (RAVLT); Block Recall (BR); Visual Patterns (VP); Backward Digit Recall (BDR); Listening Recall (LR); Digit Recall (DR); Nonword Recall (NR); Animal span (AS); 3-back numerical (3B), memory updating (MU), reading span (RS), counting span (CS), rotation span (RS), word pairs (ES).

Reasoning/fluid intelligence (reas): Wechsler picture complete, coding, mazes, block design, matrix (WechReas); Hooper Visual Organization Test (VOT); Ravens Progressive Matrices (RPM); Kaufman Brief Intelligence Test (KBIT); Bochumer Matrizen-Test (BOMAT); ETS Inference and nonsense syllogisms and surface development and paper folding (ETS); Tower of London (TOL); Delis-Kaplan Executive Functioning System – Tower Test (D-KEFS_TT); BIS verbal (BISV), BIS numerical (BISN), BIS figural-spatial (BISF-S), Test of Nonverbal Intelligence (TONI).

Response speed (spd): Symbol Digit Coding/Symbol Digit Modalities Test (SDC); Processing Speed Index (PSI); Wechsler processing speed index (WechPSI); Saccadic RTs (Gap); Digit Symbol Substitution Test (DSST); Weschler Digit Symbol Coding/Symbol Search (Wesch_Spd).

Sustained attention (sus): Underlining task (UT); Attentional Capacity Test (ACT); Continuous Performance Task (CPT); Paced Auditory Serial Addition Test (PASAT); cancellation task; Test of Everyday Attention for Children_sustained attention component(TEA-Ch_sus); Vigil Continuous Performance Test (VCPT); D2; Test of Variables of Attention (TOVA).

Selective attention (selec): Stroop; Test of Sustained Selected Attention (TOSSA); dichotic listening; Delis-Kaplan Executive Functioning System – colour-Word Interference (D-KEFS_CWI); Ruff 2&7 (RUFF); Test of Everyday Attention for Children_selective attention component(TEA-Ch_selec).

Switching/cognitive flexibility/conflict resolution (switch): Card sorting/Wisconsin Card Sorting Task (WCST); Trail Making Test (TMT); internal switching task (IST); ANT conflict; Colour Word Interference Test condition 4 – Inhibition/Switching (CWIT).

Inhibition (inhib): go/no-go; Hayling Test (HT); GoStop.

Symptom severity in ADHD (inatt): Matching Familiar Figures Test (MFFT); Attention Deficit Disorder Evaluation Scale (ADDES); parental/teacher report of impulsivity/inattentiveness; DSM-IV ADHD subscales (DSM_ADHD) (inattentive/hyperactive-impulsive) (Inc. Swanson, Nolan and Pelham (SNAP) and Conners' Teacher Rating Scale-Revised (CTRS-R); Conners Rating Scale (CRS) (for ADHD); Clinical Global Impressions (CGI) (for ADHD); ADHD Rating Scale IV (ADHD_RS); ADD-H Comprehensive Teacher Rating Scale (ACTeRS); Fremdbeurteilungsbogen fuer hyperkinetische Stroerungen (German ADHD rating scale) (FBB-HKS).

Stress/anxiety (anx/emot): Profile of Mood States (POMS); biophysical stress markers; Beck Anxiety/Beck Depression Inventories (BDI); Revised Children's Manifest Anxiety Scale (RCMAS); Test Anxiety Scale (TAS); Positive and Negative Affect Schedule (PANAS).

Other behaviour (beh gen): Brief Psychiatric Rating Scale (BPRS); Brock Adaptive Functioning Questionnaire (BAFQ); Behavioral Assessment of Dysexecutive Syndrome (BADS); Instrumental Activities of Daily Living (IADLs); Mini-Mental State Examination (MMSE); Global assessment of functioning (GAF); Everyday Attention Questionnaire (EAQ); Test of Everyday Attention (TEA); Childhood behavioural questionnaire (CBQ); Cognitive Failure Questionnaire (CFQ); Behavior Rating Inventory of Executive Function (BRIEF); DSM-IV oppositional. social problems, anxious/shy subscales; Canadian Occupational Performance Measure (COPM); motor activity; spontaneous attentional reorienting during free play; Kaufman Assessment Battery for Children (KABC); Child Behaviour Checklist CBC); Selective Reminder Test.

General academic (acad gen): Academic performance Rating Scale (APRS); Woodcock Johnson Tests of Achievement (WJTA) Inc.: Understanding Directions (WJTA_UD); Wide Range Achievement Test (WRAT).

Maths (math): Group Reading Test (GRT); Wechsler Objective Number Dimensions (WechMath); Mental Mathematics 7/8 (MentMath); Wechsler Adult Intelligence Scale_arithmetic (WAIS_m).

Literacy (lit): Basic Reading Skills (BRS); Dynamic Indicators of Basic Early Literacy Skills (DIBELS); Wechsler Objective Reading Dimensions (WechRead); Nelson-Denny Reading Test (NDRT); Salzburger Lesetest (SLT).

Language (lang): Wechsler vocabulary subtest/Wechsler Inelligence Scale verbal IQ (WechVerb); Controlled Oral Word Association Test (COWAT); Verbal fluency task (VFT); following instructions (FI); California Verbal Learning Test (CVLT); Mill Hill Vocabulary Scale (MHVS).

Social (soc): false belief; Hinting Task (HT); Bell-Lysaker Emotion Recognition Task (BLERT); Scale for the Assessment of Positive Symptoms (SAPS)/Scale for the Assessment of Negative Symptoms (SANS); Children's Communication Checklist (CCC); AUTI-R; Autism Treatment Evaluation Checklist (ATEC).
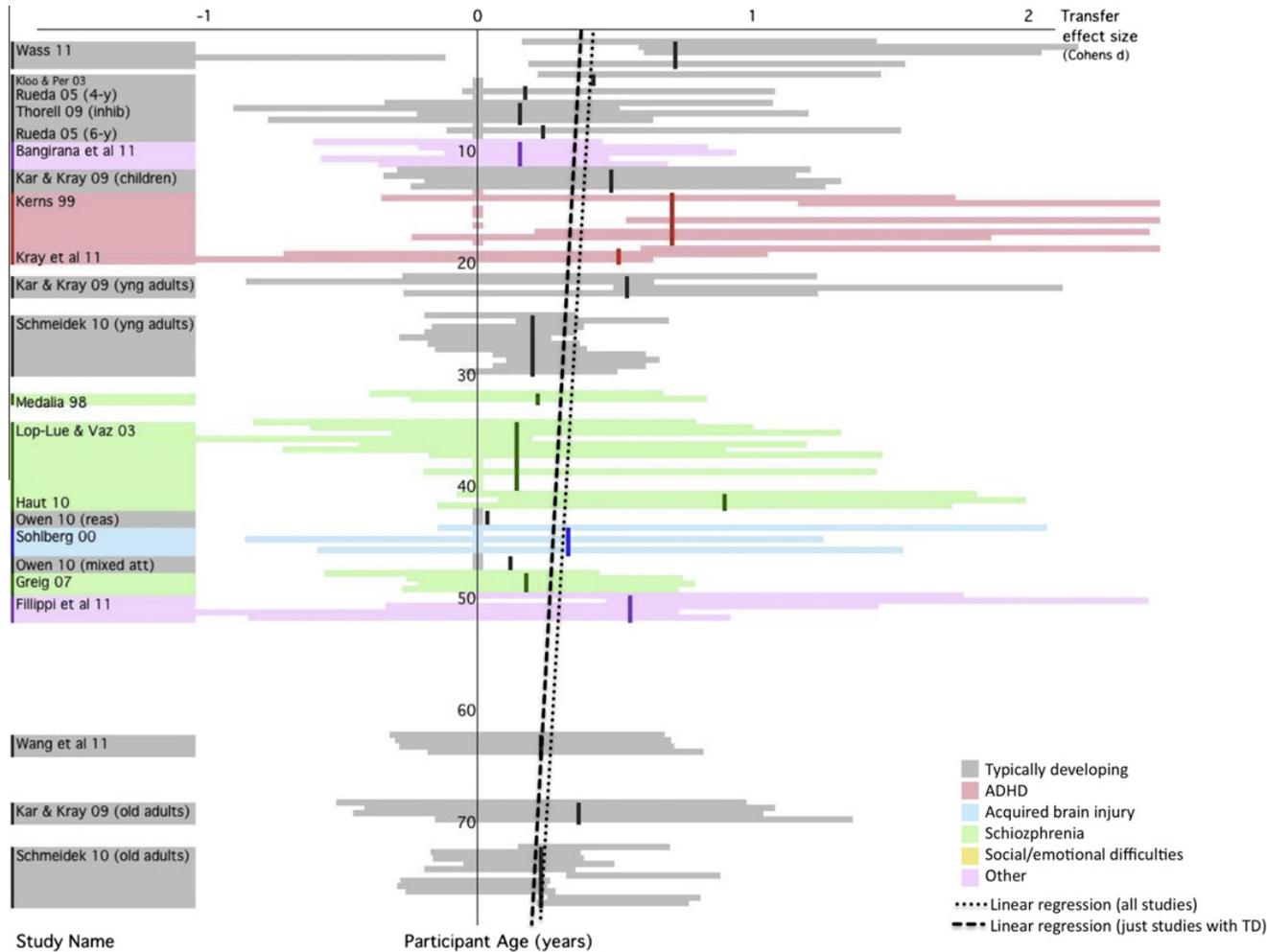
Fig. 2. Plot of MA studies. Column L of Table 1 shows, for each study included, the names of each of the pre-post tests administered to assess transfer following training; the number of pre-post tests administered varies between studies (mean: 4.5, range 2–13). This figure shows the results of these tests drawn, study by study, in ascending order of age; the position of the study on the y axis corresponds to the mean age of the participants (although this has been staggered in some cases to avoid overlap). For each study, the results of each of the pre-post tests has been drawn as a pale, horizontal line (from 95% low to 95% high CI); the results of the average of all the pre-post tests administered by that study has been drawn as a darker, vertical line. For example the first study, 'Wass 11', administered 5 different pre-post assessments (listed in Column L of Table 1) for which the CIs varied from –1.05 to –0.10 (for the WM pre-post assessment) to 0.60 to 2.1 (for the sustained attention pre-post assessment); the results of these tests have been drawn as five horizontal grey lines. The average effect size across the five tests administered was 0.81, which has been drawn as a vertical line. Different colours depict different populations that have been targeted. The smaller dotted line shows the linear regression between transfer observed and age for all studies. The other line shows the linear regression just for those studies with typically developing populations.

Axis labels:
- Top axis: Transfer effect size (Cohens d), with markings at -1, 0, 1, 2
- Left axis / bottom axis: Participant Age (years), with markings at 10, 20, 30, 40, 50, 60, 70
- Bottom left: Study Name

Study names (top to bottom):
- Thorell 09 (WM)
- St Clair Thom 10
- Holmes 09
- Klingberg 05
- Loosli et al 11
- Klingberg 02
- Roughan & Hadwin 11
- van der Mol 10
- Chein & Mor 10
- Jaeggi 10
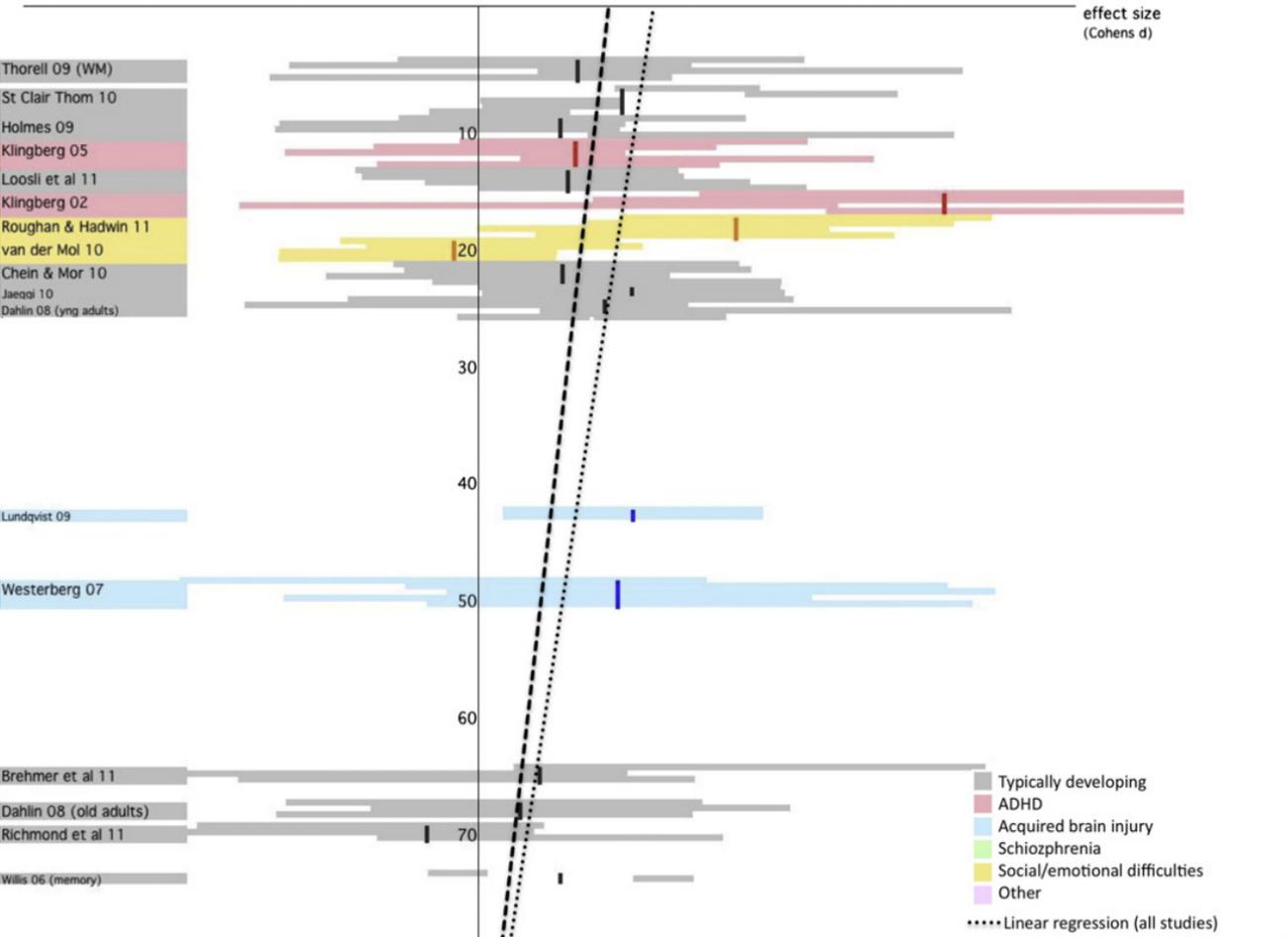- Dahlin 08 (yng adults)
- Lundqvist 09
- Westerberg 07
- Brehmer et al 11
- Dahlin 08 (old adults)
- Richmond et al 11
- Willis 06 (memory)

Legend:
- Typically developing
- ADHD
- Acquired brain injury
- Schizophrenia
- Social/emotional difficulties
- Other
- Linear regression (all studies)
- Linear regression (just studies with TD)

The first conclusion we wish to draw from this review is that substantially altering cognition is hard. The findings described here represent probably the most promising avenues of research in cognitive training of domain-general processes, yet negative findings (i.e. findings that training improvements fail to generalise, even to related tasks) are still frequent. For example even a well-replicated finding, such as that training working memory leads to improvements at reasoning tasks, is still observed only in about half of the studies that have examined this relationship (see Table 1). Key findings, such as the relationship between cognitive training and problem behaviours in ADHD (Klingberg et al., 2005) and math performance in schools (Holmes et al., 2009; Kerns et al., 1999) have yet to be widely replicated, although the number of negative findings in this area is also small. Only a small number of studies have examined the degree to which training improvements are maintained over longer time periods (see Column M of Table 1). It should be remembered, however, that the falsifiability of hypotheses within the training literature is different to that in other areas of cognitive science. The possibility of a type I error appears neglible in comparison to the possibility that a negative result arises from the fact that insufficient training was applied to reveal a genuine between-domain relationship. In this regard, it is important to remember that most of the studies we have included here have administered only one, intensive phase of training that lasts ca. 10 h (see Columns F and G of Table 1), which in terms of total waking hours is a relatively short exposure time.

Consistent with the predictions outlined in the introduction we have reviewed evidence suggesting that studies targeting younger participants are more likely to report transfer of training effects. This pattern is, however, only a relatively weak one. Although consistent relationships are found across all groups, the relationship between training transfer and age is significant when all studies are pooled together but not when WM and MA studies are considered individually. The fact that the observed relationship was stronger for WM than for MA studies may be because of the greater homogeneity of training within the WM group, which allows for the effect of individual variables such as age to be identified more easily.

The evidence we have reviewed certainly does not suggest that training applied to younger participants is the only means of demonstrating between-domain transfer following training: a number of studies targeting older individuals (both mature adults and older adults (aged 70+) have reported transfer of training effects to non-trained tasks (Brehmer et al., 2011; Dahlin et al., 2008; Richmond, Morrison, Chein, & Olson, 2011; Schmiedek et al., 2010; Wang, Chang, & Su, 2011; Zinke, Zeintl, Eschen, Herzog, & Kliegel, 2011). It can also be seen from Figs. 2 and 3 that studies targeting individuals with disorders such as ADHD (Klingberg et al., 2002, 2005) and acquired brain injury (e.g. Lundqvist et al., 2009; Westerberg et al., 2007) appear to report transfer of training effects that is above that reported by studies targeting typically developing individuals of a similar age.

If, however, we examine just those studies that have administered training to typically developing individuals, then we do observe a significant correlation between transfer reported and age. This correlation does not appear to be driven by any one particular age group, but to be continuous across developmental time. Given the theoretical considerations we discussed in the introduction, however, we might expect to find a stronger, non-linear (e.g. exponential) relationship – according to which training administered early in development reports much more widespread transfer than similar training administered to older participants. Why have no such relationships been observed?

Our own practical experience points to the vital importance of maintaining motivation and engagement whilst administering training to younger participants – particularly very young participants for whom metacognitive factors (i.e. an awareness of what they are doing and why) cannot be used to increase motivation. Studies targeting younger participants tend to adminster smaller doses of training (see Column F of table) – but what is not reported is differences in the level of engagement during training (e.g. frequency of off-task glances, etc.), which can also potentially influence the effectiveness of the training that is being applied. Future work should concentrate on addressing these goals by developing more sophisticated training methodologies that pay greater attention to the importance of maintaining motivation and to the limitations of traditional point-and-click interfaces with very young participants. The gaze-contingent training paradigms that we used to train attentional control in infants – according to which different events take place contingent on where the participant is looking – may represent fruitful avenues for further work here (Wass et al., 2011). Another factor that may mitigate the relationship observed between age and training transfer is that behavioural measures for

younger participants may be less reliable and/or less sensitive than those used with adults, due to the increased influence of factors such as short-term variability in mood on test performance in very young individuals (e.g. Akshoomoff, 2002).

An additional factor may have contributed to the fact that the relationship we found between transfer observed and training age was only a relatively weak one: namely that the majority of the studies we reviewed have administered just one short, intensive training phase and administered pre–post assessments immediately after the cessation of training (although see Column M of the Table for the small number of studies that did administer a longer-term follow-up). Many of the longer-term developmental interactions we postulated in the discussion would, therefore, not be measurable using these techniques. Future work should concentrate on administering training over longer time periods, and measuring systematically the degree to which transfer is observed to distal cognitive domains.

In the introduction we discussed longitudinal correlational evidence suggesting that attentional control may be important over long-term developmental time-frames for gating the subsequent acquisition of other skills in other non-contiguous domains – although we also pointed out that the evidence suggesting between-domain relationships is largely correlational (e.g. Cornish et al., 2012a, 2012b; Kannass & Oakes, 2008; Rose et al., 2009; Rose, Feldman, Jankowski, & Van Rossem, 2011; Scerif et al., 2012). Can findings from the training literature be used to demonstrate causal relationships here? From the studies we have reviewed, it would seem that no firm conclusions can be drawn. Although a variety of methods have been shown to be effective at training attentional control, the long-term relationships between attentional control and learning in other domains have only been very partially explored. This should not, though, be taken as evidence against the developmental hypotheses we discussed in the introduction – rather that the training studies currently in existence were inadequately designed to address these hypotheses. The amounts of training that have been administered remain small, and assessment of the longer-term effects of training remains limited.

With regard to both of the points we have raised – the problems of maintaining motivation over longer time frames, and the importance of assessing transfer over longer periods and following larger amounts of training, we believe that the non-traditional forms of cognitive training we have discussed (mindfulness meditation (MM), action computer game playing, behavioural interventions) can also make a substantial contribution. For example, whereas all of the WM papers we have reviewed have administered a single training phase (generally c. 10 h over c. 5 weeks – see Column F of the table) with follow-up assessments at intervals after the cessation of training (see Column M), MM studies generally (although not always – see e.g. Tang et al., 2007) administer a longer training phase. MM is designed to be practiced in an ongoing manner across the lifetime, which has allowed several studies to report on the effects of far larger doses of training (1000s of hours rather than 10s) (see Lutz et al., 2008; Tang & Posner, 2009).

One underacknowledged tension in this field is, we believe, that between theoretical and applied goals. Whereas some researchers have developed cognitive training from theoretical (basic science) perspectives (i.e. those applying training in order to help us better to understand the underlying cognitive processes), others have been motivated by applications of this work (i.e. improving outcome measures such as academic performance). These two approaches sometimes have conflicting goals. For example, administering a training regime that is directed at a single, clearly defined component of cognition such as WM (Klingberg, 2010) may be the best way of addressing theoretical questions, whereas applied goals (such as reducing symptom severity in ADHD) may be better accomplished with a more varied cognitive training regime. This is particularly true given that a number of researchers are reporting that transfer within executive attention training tasks tends to be fairly narrow (e.g. Diamond & Lee, 2011): it may that training a battery of partially complementary subcomponental processes may be more successful in this regard.

Concluding remarks

Substantially altering cognition is hard. We have reviewed considerable negative evidence of training transfer, and we have emphasised that many of the studies in this area that have reported positive findings are inadequately designed, in particular with regard to the vital importance of active control groups.

Consistent with our hypothesis, however, we have found that training studies targeting younger participants tend to report relatively more widespread transfer of training effects. This agrees with evidence discussed in the introduction that neural and behavioural plasticity ought to be greater earlier in development, and suggests that future applied cognitive training should concentrate on targeting younger participants. This is important not just because neural systems are more malleable earlier in development, but also because attentional control is thought to be required for the subsequent acquisition in other skills in other domains. In a number of developmental disorders for example, early deficits in attentional control are thought to lead, cascade-like, to subsequent impaired learning in other areas. Thus interventions targeted earlier in development potentially have the ability to minimize cascade-like deficits before they occur.

We have also suggested a number of practical ways in which training studies targeting younger participants might be optimised. First, we noted that different methods of training attentional control have been shown to lead to at least partially overlapping patterns of performance improvement, and suggested that one way of encouraging younger participants to remain engaged and motivated for longer training periods may be the development of more varied training regimes. Second, we noted that the studies we have reviewed have generally administered one, short (c. 10 h) training phase which in terms of waking hours is a small exposure time, and that the fact that training tends to be administered in one discrete training period may be non-optimal. (Compare the effect of one short phase of cardiovascular exercise on cardiovascular fitness over time.) Future studies should be designed not just to measure cross-sectional performance at short- and medium-term follow-up, but also to measure longitudinal changes, in particular with regard to assessing participants' subsequent ability to acquire new skills in other domains.

From an applied perspective, therefore, we recommend a number of goals for future work: (i) refinement of training techniques based on better integration of ideas from different traditions; (ii) longer-term and more regularly spaced training methods; (iii) more rigorous control groups and statistical analyses; (iv) intensive interventions targeting individuals very early in development.

Acknowledgments

References

Adams, C. E., Power, A., Frederick, K., & Lefebvre, C. (1994). An investigation of the adequacy of Medline searches for randomized controlled trials (RCTs) of the effects of mental-health care. *Psychological Medicine, 24*(3), 741–748.

Akshoomoff, N. (2002). Selective attention and active engagement in young children. *Developmental Neuropsychology, 22*(3), 625–642.

Annaz, D., Karmiloff-Smith, A., Johnson, M. H., & Thomas, M. S. C. (2009). A cross-syndrome study of the development of holistic face recognition in children with autism, Down syndrome, and Williams syndrome. *Journal of Experimental Child Psychology, 102*(4), 456–486.

Arns, M., de Ridder, S., Strehl, U., Breteler, M., & Coenen, A. (2009). Efficacy of neurofeedback treatment in ADHD: The effects on inattention, impulsivity and hyperactivity: A meta-analysis. *Clinical EEG and Neuroscience, 40*(3), 180–189.

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

Bangirana, P., Allebeck, P., Boivin, M. J., John, C. C., Page, C., Ehnvall, A., et al (2011). Cognition, behaviour and academic skills after cognitive rehabilitation in Ugandan children surviving severe malaria: A randomised trial. *BMC Neurology, 11.*

Barnett, W. S., Jung, K., Yarosz, D. J., Thomas, J., Hornbeck, A., Stechuk, R., et al (2008). Educational effects of the tools of the mind curriculum: A randomized trial. *Early Childhood Research Quarterly, 23*(3), 299–313.

Bavelier, D., Green, C. S., & Dye, M. W. G. (2010). Children, wired: For better and for worse. *Neuron, 67*(5), 692–701.

Beauregard, M. (2007). Mind does really matter: Evidence from neuroimaging studies of emotional self-regulation, psychotherapy, and placebo effect. *Progress in Neurobiology, 81*(4), 218–236.

Beauregard, M., & Levesque, J. (2006). Functional magnetic resonance imaging investigation of the effects of neurofeedback training on the neural bases of selective attention and response inhibition in children with attention-deficit/hyperactivity disorder. *Applied Psychophysiology and Biofeedback, 31*(1), 3–20.

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.

Bellander, M., Brehmer, Y., Westerberg, H., Karlsson, S., Furth, D., Bergman, O., et al (2011). Preliminary evidence that allelic variation in the LMX1A gene influences training-related working memory improvement. *Neuropsychologia, 49*(7), 1938–1942.

Bergman Nutley, S., Söderqvist, 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.

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology, 20*(3), 899–911.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129*(3), 387–398.

Brefczynski-Lewis, J. A., Lutz, A., Schaefer, H. S., Levinson, D. B., & Davidson, R. J. (2007). Neural correlates of attentional expertise in long-term meditation practitioners. *Proceedings of the National Academy of Sciences of the United States of America, 104*(27), 11483–11488.

Brehmer, Y., Rieckmann, A., Bellander, M., Westerberg, H., Fischer, H., & Backman, L. (2011). Neural correlates of training-related working-memory gains in old age. *NeuroImage, 58*(4), 1110–1120.

Brocki, K., Clerkin, S. M., Guise, K. G., Fan, J., & Fossella, J. A. (2009). Assessing the molecular genetics of the development of executive attention in children: Focus on genetic pathways related to the anterior cingulate cortex and dopamine. *Neuroscience, 164*(1), 241–246.

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.

Burger, K. (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early Childhood Research Quarterly, 25*(2), 140–165.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4*(6), 215–222.

Cahn, B. R., & Polich, J. (2006). Meditation states and traits: EEG, ERP, and neuroimaging studies. *Psychological Bulletin, 132*(2), 180–211.

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

Cannon, R., Congedo, M., Lubar, J., & Hutchens, T. (2009). Differentiating a network of executive attention: Loreta neurofeedback in anterior cingulate and dorsolateral prefrontal cortices. *International Journal of Neuroscience, 119*(3), 404–441.

Castel, A. D., Pratt, J., & Drummond, E. (2005). The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. *Acta Psychologica, 119*(2), 217–230.

Chan, D., & Woollacott, M. (2007). Effects of level of meditation experience on attentional focus: Is the efficiency of executive or orientation networks improved? *Journal of Alternative and Complementary Medicine, 13*(6), 651–657.

Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin and Review, 17*(2), 193–199.

Chiesa, A., Calati, R., & Serretti, A. (2010). Does mindfulness training improve cognitive abilities? A systematic review of neuropsychological findings. *Clinical Psychology Review, 31*(3), 449–464.

Chisholm, J. D., Hickey, C., Theeuwes, J., & Kingstone, A. (2010). Reduced attentional capture in action video game players. *Attention Perception and Psychophysics, 72*(3), 667–671.

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*(34), 283–322.

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

Cornish, K. M., Turk, J., Wilding, J., Sudhalter, V., Munir, F., Kooy, F., et al (2004). Annotation: Deconstructing the attention deficit in fragile X syndrome: A developmental neuropsychological approach. *Journal of Child Psychology and Psychiatry, 45*(6), 1042–1053.

Cornish, K., Cole, V., Longhi, E., Karmiloff-Smith, A., & Scerif, G. (2012a). Mapping developmental trajectories of attention and working memory in fragile X syndrome: Developmental freeze or developmental change? *Development and Psychopathology*.

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

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

Courage, M. L., Reynolds, G. D., & Richards, J. E. (2006). Infants' attention to patterned stimuli: Developmental change from 3 to 12 months of age. *Child Development, 77*(3), 680–695.

Dahlin, E., Nyberg, L., Backman, L., & Neely, A. S. (2008). Plasticity of executive functioning in young and older adults: Immediate training gains, transfer, and long-term maintenance. *Psychology and Aging, 23*(4), 720–730.

Davidson, M. C., Amso, D., Cruess Anderson, L., & 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.

Deoni, S. C. L., Mercure, E., Blasi, A., Gasston, D., Thomson, A., Johnson, M., et al (2011). Mapping infant brain myelination with magnetic resonance imaging. *Journal of Neuroscience, 31*(2), 784–791.

Diamond, A., Briand, L., Fossella, J., & Gehlbach, L. (2004). Genetic and neurochemical modulation of prefrontal cognitive functions in children. *American Journal of Psychiatry, 161*(1), 125–132.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333*(6045), 959–964.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). The early years - Preschool program improves cognitive control. *Science, 318*(5855), 1387–1388.

Dimitrov, D. M., & Rumrill, P. D. (2003). Pretest–posttest designs and measurement of change. *Work, 20*(2), 159–165.

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., et al (2006). A shift from diffuse to focal cortical activity with development. *Developmental Science, 9*(1), 1–8.

Dye, M. W., Green, C. S., & Bavelier, D. (2009). The development of attention skills in action video game players. *Neuropsychologia, 47*(8–9), 1780–1789.

Edelman, G. M. (1993). Neural Darwinism – Selection and reentrant signaling in higher brain-function. *Neuron, 10*(2), 115–125.

Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, Mass.: MIT Press.

Fair, D. A., Cohen, A. L., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., Barch, D. M., et al (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., Cohen, A. L., Power, J. D., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., et al (2009). Functional brain networks develop from a "local to distributed" organization. *PLoS Computational Biology, 5*(5).

Filippi, M., Riccitelli, G., Mattioli, F., Capra, R., Stampatori, C., Pagani, E., et al (2011). Multiple sclerosis: Effects of cognitive rehabilitation on structural and functional MR imaging measures – An explorative study. *Radiology, 262*(3), 932–940.

Gazzaley, A., & Nobre, A. C. (2011). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences, 16*(2), 129–135.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron, 14*(3), 477–485.

Gordon, E. M., Lee, P. S., Maisog, J. M., Foss-Feig, J., Billington, M. E., VanMeter, J., et al (2011). Strength of default mode resting-state connectivity relates to white matter integrity in children. *Developmental Science, 14*(4), 738–751.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534–537.

Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science, 18*(1), 88–94.

Green, C. S., & Bavelier, D. (2008). Exercising your brain: A review of human brain plasticity and training-induced learning. *Psychology and Aging, 23*(4), 692–701.

Greig, T. C., Zito, W., Wexler, B. E., Fiszdon, J., & Bell, M. D. (2007). Improved cognitive function in schizophrenia after one year of cognitive training and vocational services. *Schizophrenia Research, 96*, 156–161.

Grossmann, T., & Johnson, M. H. (2007). The development of the social brain in human infancy. *European Journal of Neuroscience, 25*(4), 909–919.

Halperin, J. M., & Healey, D. M. (2011). The influences of environmental enrichment, cognitive enhancement, and physical exercise on brain development: Can we alter the developmental trajectory of ADHD? *Neuroscience and Biobehavioral Reviews, 35*(3), 621–634.

Haut, K. M., Lim, K. O., & MacDonald, A. (2010). Prefrontal cortical changes following cognitive training in patients with chronic schizophrenia: Effects of practice, generalization, and specificity. *Neuropsychopharmacology, 35*(9), 1850–1859.

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

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando: Academic Press. ISBN 0-12-336380-2.

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.

Holmes, J., Gathercole, S. E., Place, M., Dunning, D. L., Hilton, K. A., & Elliott, J. G. (2010). Working memory deficits can be overcome: Impacts of training and medication on working memory in children with ADHD. *Applied Cognitive Psychology, 24*(6), 827–836.

Hunter, J., & Schmidt, F. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Beverly Hills, CA: Sage.

Hunter, J., Schmidt, F., & Jackson, G. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.

Iarocci, G., Enns, J. T., Randolph, B., & Burack, J. A. (2009). The modulation of visual orienting reflexes across the lifespan. *Developmental Science, 12*(5), 715–724.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America, 105*(19), 6829–6833.

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y. F., Jonides, J., & Perrig, W. J. (2010). The relationship between *n*-back performance and matrix reasoning – Implications for training and transfer. *Intelligence, 38*(6), 625–635.

Jamieson, J. (2004). Analysis of covariance (ANCOVA) with difference scores. *International Journal of Psychophysiology, 52*(3), 277–283.

Jang, J. H., Jung, W. H., Kang, D.-H., Byun, M. S., Kwon, S. J., Choi, C.-H., et al (2010). Increased default mode network connectivity associated with meditation. *Neuroscience Letters, 487*(3), 358–362.

Jha, A. P., Stanley, E. A., Kiyonaga, A., Wong, L., & Gelfand, L. (2011). Examining the protective effects of mindfulness training on working memory capacity and affective experience. *Emotion, 10*(1), 54–64.

Johnson, M. H. (2000). Functional brain development in infants: Elements of an interactive specialization framework. *Child Development, 71*(1), 75–81.

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

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.

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.

Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science, 12*(6), 978–990.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science.* Cambridge, Mass.: MIT Press/Bradford Books.

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.

Kelly, A. M. C., Di Martino, A., Uddin, L. Q., Shehzad, Z., Gee, D. G., Reiss, P. T., et al (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.

Kerns, K. A., MacSween, J., Wekken, S. V., & Gruppuso, V. (2010). Investigating the efficacy of an attention training programme in children with foetal alcohol spectrum disorder. *Developmental Neurorehabilitation, 13*(6), 413–422.

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., Dahlstrom, K., et al (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.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology, 24*(6), 781–791.

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.

Kolata, S., Light, K., Wass, C. D., Colas-Zelin, D., Roy, D., & Matzel, L. D. (2010). A dopaminergic gene cluster in the prefrontal cortex predicts performance indicative of general intelligence in genetically heterogeneous mice. *PLoS ONE, 5*(11).

Kouijzer, M. E. J., de Moor, J. M. H., Gerrits, B. J. L., Congedo, M., & van Schie, H. T. (2009). Neurofeedback improves executive functioning in children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 3*(1), 145–162.

Kray, J., Karbach, J., Haenig, S., & Freitag, C. (2012). Can task-switching training enhance executive control functioning in children with attention deficit/-hyperactivity disorder? *Frontiers in Human Neuroscience, 5.*

Laing, E., Butterworth, G., Ansari, D., Gsodl, M., Longhi, E., Panagiotaki, G., et al (2002). Atypical development of language and social communication in toddlers with Williams syndrome. *Developmental Science, 5*(2), 233–246.

Lawson, K. R., & Ruff, H. A. (2004a). Early attention and negative emotionality predict later cognitive and behavioural function. *International Journal of Behavioral Development, 28*(2), 157–165.

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

Lopez-Luengo, B., & Vazquez, C. (2003). Effects of attention process training on cognitive functioning of schizophrenic patients. *Psychiatry Research, 119*(1–2), 41–53.

Loevden, M., Backman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin, 136*(4), 659–676.

Loosli, S. V., Buschkuehl, M., Perrig, W. J., & Jaeggi, S. M. (2011). Working memory training improves reading processes in typically developing children. *Child Neuropsychology, 18*(1), 62–78.

Lundqvist, A., Grundstrom, K., Samuelsson, K., & Ronnberg, J. (2010). Computerized training of working memory in a group of patients suffering from acquired brain injury. *Brain Injury, 24*(10), 1173–1183.

Lutz, A., Greischar, L. L., Rawlings, N. B., Ricard, M., & Davidson, R. J. (2004). Long-term meditators self-induce high-amplitude gamma synchrony during mental practice. *Proceedings of the National Academy of Sciences of the United States of America, 101*(46), 16369–16373.

Lutz, A., Slagter, H. A., Dunne, J. D., & Davidson, R. J. (2008). Cognitive-emotional interactions – Attention regulation and monitoring in meditation. *Trends in Cognitive Sciences, 12*(4), 163–169.

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

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

Medalia, A., Aluma, M., Tryon, W., & Merriam, A. E. (1998). Effectiveness of attention training in schizophrenia. *Schizophrenia Bulletin, 24*(1), 147–152.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*(1), 40–48.

Mishra, J., Zinni, M., Bavelier, D., & Hillyard, S. A. (2010). Neural basis of superior performance of action videogame players in an attention-demanding task. *Journal of Neuroscience, 31*(3), 992–998.

Mulder, H., Pitchford, N. J., & Marlow, N. (2010). Inattentive behaviour is associated with poor working memory and slow processing speed in very pre-term children in middle childhood. *British Journal of Educational Psychology, 81*(1), 147–160.

Napoli, M., Krech, P. R., & Holley, L. C. (2005). Mindfulness training for elementary school students. *Journal of Applied School Psychology, 21*(1), 99–125.

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.

Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., et al (2010). Putting brain training to the test. *Nature, 465*(7299), U775–U776.

Paterson, S. J., Brown, J. H., Gsodl, M. K., Johnson, M. H., & Karmiloff-Smith, A. (1999). Cognitive modularity and genetic disorders. *Science, 286*(5448), 2355–2358.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience, 9*(2), 148–158.

Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology.*

Posner, M. I., & Rothbart, M. K. (2009). Toward a physical basis of attention and self-regulation. *Physics of Life Reviews, 6*(2), 103–120.

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

Rabiner, D. L., Murray, D. W., Skinner, A. T., & Malone, P. S. (2010). A randomized trial of two promising computer-based interventions for students with attention difficulties. *Journal of Abnormal Child Psychology, 38*(1), 131–142.

Razza, R. A., Martin, A., & Brooks-Gunn, J. (2010). Associations among family environment, sustained attention, and school readiness for low-income children. *Developmental Psychology, 46*(6), 1528–1542.

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.

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., & 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. (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., & 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.

Roughan, L., & Hadwin, J. A. (2011). The impact of working memory training in young people with social, emotional and behavioural difficulties. *Learning and Individual Differences, 21*(6), 759–764.

Rubia, K. (2009). The neurobiology of Meditation and its clinical effectiveness in psychiatric disorders. *Biological Psychology, 82*(1), 1–11.

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. (1990). In J. Colombo, & J. Fagen (Eds.), *Individual differences in sustained attention during infancy* (pp. 247–270).

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

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

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., Cornish, K., Wilding, J., Driver, J., & Karmiloff-Smith, A. (2007). Delineation of early attentional control difficulties in fragile X syndrome: Focus on neurocomputational mechanisms. *Neuropsychologia, 45*, 1889–1898.

Scerif, G., & Karmiloff-Smith, A. (2005). The dawn of cognitive genetics? Crucial developmental caveats. *Trends in Cognitive Sciences, 9*(3), 126–135.

Scerif, G., Karmiloff-Smith, A., Campos, R., Elsabbagh, M., Driver, J., & Cornish, K. (2005). To look or not to look? Typical and atypical development of oculomotor control. *Journal of Cognitive Neuroscience, 4*, 591–604.

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*.

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.

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

Schweizer, S., Hampshire, A., & Dalgleish, T. (2011). Extending brain-training to the affective domain: Increasing cognitive and affective executive control through emotional working memory training. *PLoS ONE, 6*(9).

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.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica, 50*(3–4), 245–276.

Shonkoff, J. P., & Levitt, P. (2011). 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., et al (2007). Mental training affects distribution of limited brain resources. *PLoS Biology, 5*(6), 1228–1235.

Sohlberg, M. M., McLaughlin, K. A., Pavese, A., Heidrich, A., & Posner, M. I. (2000). Evaluation of attention process training and brain injury education in persons with acquired brain injury. *Journal of Clinical and Experimental Neuropsychology, 22*(5), 656–676.

Sonuga-Barke, E. J. S., & Halperin, J. M. (2010). 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.

Sonuga-Barke, E. J. S., Koerting, J., Smith, E., McCann, D. C., & Thompson, M. (2011). Early detection and intervention for attention-deficit/hyperactivity disorder. *Expert Review of Neurotherapeutics, 11*(4), 557–563.

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.

Steele, A., Karmiloff-Smith, A., Cornish, K. M., & Scerif, G. (2012). The multiple sub-functions of attention: Differential developmental gateways to literacy and numeracy. *Child Development*.

Stevens, C., Lauinger, B., & Neville, H. (2009). Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: An event-related brain potential study. *Developmental Science, 12*(4), 634–646.

Tamm, L., Hughes, C., Ames, L., Pickering, J., Silver, C. H., Stavinoha, P., et al (2008). Attention training for school-aged children with ADHD: Results of an open trial. *Journal of Attention Disorders, 14*(1), 86–94.

Tang, Y.-Y., Lu, Q., Geng, X., Stein, E. A., Yang, Y., & Posner, M. I. (2010). Short-term meditation induces white matter changes in the anterior cingulate. *Proceedings of the National Academy of Sciences of the United States of America, 107*(35), 15649–15652.

Tang, Y.-Y., & Posner, M. I. (2009). Attention training and attention state training. *Trends in Cognitive Sciences, 13*(5), 222–227.

Tang, Y. Y., Ma, Y. H., Wang, J., Fan, Y. X., Feng, S. G., Lu, Q. L., et al (2007). Short-term meditation training improves attention and self-regulation. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 17152–17156.

Thomas, M., & Karmiloff-Smith, A. (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. *Behavioral and Brain Sciences, 25*(6). 727–+..

Thompson, M. J. J., Laver-Bradbury, C., Ayres, M., Le Poidevin, E., Mead, S., Dodds, C., et al (2009). A small-scale randomized controlled trial of the revised new forest parenting programme for preschoolers with attention deficit hyperactivity disorder. *European Child and Adolescent Psychiatry, 18*(10), 605–616.

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.

van de Weijer-Bergsma, E., Wijnroks, L., & Jongmans, M. J. (2008). Attention development in infants and preschool children born preterm: A review. *Infant Behavior and Development, 31*(3), 333–351.

van den Hurk, P. A. M., Giommi, F., Gielen, S. C., Speckens, A. E. M., & Barendregt, H. P. (2009). Greater efficiency in attentional processing related to mindfulness meditation. *Quarterly Journal of Experimental Psychology, 63*(6), 1168–1180.

Van der Molen, M. J., Van Luit, J. E. H., Van der Molen, M. W., Klugkist, I., & Jongmans, M. J. (2010). Effectiveness of a computerised working memory training in adolescents with mild to borderline intellectual disabilities. *Journal of Intellectual Disability Research, 54*, 433–447.

van Vugt, M. K., & Jha, A. P. (2011). Investigating the impact of mindfulness meditation training on working memory: A mathematical modeling approach. *Cognitive Affective and Behavioral Neuroscience, 11*(3), 344–353.

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.

Vernon, D., Egner, T., Cooper, N., Compton, T., Neilands, C., Sheri, A., et al (2003). The effect of training distinct neurofeedback protocols on aspects of cognitive performance. *International Journal of Psychophysiology, 47*(1), 75–85.

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., & Johnson, M. H. (2011). Training attentional control in infancy. *Current Biology, 21*(18), 1543–1547.

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.

Westerberg, H., Jacobaeus, H., Hirvikoski, T., Clevberger, P., Ostensson, M. L., Bartfai, A., et al (2007). Computerized working memory training after stroke – A pilot study. *Brain Injury, 21*(1), 21–29.

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., et al (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *JAMA – Journal of the American Medical Association, 296*(23), 2805–2814.

Zinke, K., Zeintl, M., Eschen, A., Herzog, C., & Kliegel, M. (2011). Potentials and limits of plasticity induced by working memory training in old–old age. *Gerontology, 58*(1), 79–87.

Zylowska, L., Ackerman, D. L., Yang, M. H., Futrell, J. L., Horton, N. L., Hale, T. S., et al (2008). Mindfulness meditation training in adults and adolescents with ADHD: A feasibility study. *Journal of Attention Disorders, 11*(6), 737–746.