The attentive brain: insights from developmental cognitive neuroscience

Dima Amso[1] and Gaia Scerif[2]

Abstract | Visual attention functions as a filter to select environmental information for learning and memory, making it the first step in the eventual cascade of thought and action systems. Here, we review studies of typical and atypical visual attention development and explain how they offer insights into the mechanisms of adult visual attention. We detail interactions between visual processing and visual attention, as well as the contribution of visual attention to memory. Finally, we discuss genetic mechanisms underlying attention disorders and how attention may be modified by training.

Working memory
A cognitive operation that involves manipulating the contents of short-term memory to direct goal-relevant action.

Attentional network task
(ANT). An attentional cueing paradigm designed to provide separable indices of alerting, orienting and executive attention.

[1]*Department of Cognitive, Linguistic & Psychological Sciences, Brown University, Providence, Rhode Island 02912, USA.*
[2]*Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK.*
Correspondence to G.S.
e-mail:
gaia.scerif@psy.ox.ac.uk

The world is cluttered with more information than can be processed at once. Attention is defined as a process or computation that is applied to competing environmental information, the result of which is to bias selection and action to one option while simultaneously filtering interference from the remaining alternatives[1–4]. Framing attention as a computation is useful because it explains how attention processes can be carried out on a range of sensory inputs, as well as on more-abstract representations. For example, visual attention can bias selection of information about objects, such as particular features or locations. Attention can also act to select goals for action from the contents of working memory. In all these cases, attention processes determine what information is selected for subsequent perception, action, learning and memory, imposing a crucial processing bottleneck. It is therefore one of the most-studied mechanisms in the adult cognitive neurosciences.

However, a complete understanding of attention processes must also include an understanding of their developmental origins. In this Review, we highlight how studying developing rather than developed attentional states broadens our understanding of attention mechanisms and forces a shift in focus from considering attention as an isolated process towards an understanding of its links with perception and memory, as well as its genetic constraints and malleability. We discuss studies of typical development of attention processes and studies of neurodevelopmental disorders in which attention processes are atypical. It is important to note that attention operates in various sensory modalities. Here, we focus largely on cortical mechanisms of visual attention development, but we suggest that the processes and approaches discussed here may operate in a similar way across other sensory modalities. Finally, we propose novel ideas for successful training of attention during development.

Attention processes in the adult

Posner and colleagues were the first to propose a model that described three separable attention processes — alerting, spatial orienting and executive attention — supported by different brain networks[1–3]. In this model, alerting is defined as the generating of a state of arousal or readiness elicited by an unexpected external cue. Orienting is defined as the shifting of attention to select information in the environment[4], and may be either overt (associated with an eye or head movement) or covert (not associated with eye or head movement). Executive attention is defined as a process that resolves conflict between competing inputs for the purpose of selecting goal-relevant action[5]. These three attention processes interact with sensorimotor processing systems: they may operate across different sensory modalities and therefore act together to regulate multi-sensory integration of information[6]. Attentional processes are also modulated by motor input and, although the interaction between motor input and attention is beyond the scope of this Review, it should be noted that these interactions have been studied extensively in the context of premotor theories of attention[7,8].

Spatial cueing tasks are often used to study alerting, orienting and executive attention processes. In the visual domain, a commonly used cueing task is the attentional network task (ANT)[1]. This task involves presenting the participant with an unexpected cue, often on a computer screen, to alert them that a stimulus is about to occur, and thus it is used to study alerting. The cue may also provide information about where the stimulus will

occur, allowing a participant to shift attention to that location before the stimulus appears, so it can also be used to study orienting. Alternatively, a central arrow cue may point to one side of the screen while other arrow cues flanking it point in either the same or the opposite direction; participants must resolve this competition to report the direction of the central item, providing an index of executive attention. Functional magnetic resonance imaging (fMRI) data obtained from adults during such visual cueing tasks have revealed that the networks supporting each of these attention processes are largely independent[1,9,10]. Alerting has been shown to involve the locus coeruleus, right parietal and frontal regions and is modulated by the neurotransmitter noradrenaline[1,2,11]. Orienting involves the activation of the frontal eye fields (FEF), superior parietal junction, superior temporal junction, superior colliculus and pulvinar and is modulated by acetylcholine[1,2,11]. Finally, executive attention is known to involve the anterior cingulate, anterior insula, frontal cortex and striatum and is modulated by dopamine[1,2,11]. Here, we review studies that focus on visual attention and that track the emergence, at separate developmental time points, of each attention process. A key observation is that temporal dissociations of attention processes are evident over the course of development, providing a powerful way of differentiating attention in developing and adult states[12,13].

Development of visual attention

Studies of visual attention alerting and orienting in infancy depend heavily on assessment of eye movement dynamics (BOX 1). Such studies have shown that even newborn babies have the capacity for alerting, in its most basic form[14]. However, the more-complex visual attention-orienting mechanism, which allows for suppression of competing information during attention-orienting shifts, becomes functional only between 4 and 6 months of age[14–19]. Before this age, attention orienting in infants primarily consists of simpler processes that facilitate the orienting of the infants' attention towards perceptually salient information[16,20–22]. For example, 3-month-old infants can quickly shift visual attention towards a particular location that is indicated to be important by a parent (facilitation-based orienting), but it is not until 4–6 months of age that they are able to suppress distracting information from the previously attended location when they make this attention-orienting shift. Both forms of orienting are likely to engage a similar cortical network that spans a caudal to rostral axis from the lateral occipital to the parietal and frontal cortices, including the prefrontal cortex (PFC) and FEF[23,24]. Studies of attention-orienting behaviour and neural activity in children from 6 years of age onwards show that there is continued development of visual orienting capacity during childhood and adolescence[12,25].

Top-down executive attention processes generally involve a rule that governs behaviour when stimuli compete or are in conflict. In addition to using the attentional network task described above, executive attention can be assessed using an antisaccade task in which, for example,

subjects are taught to always look away from a cue (antisaccades) rather than towards it (prosaccades). There is some evidence that this antisaccade function is present in infants as young as 4 months of age[26]. Adapting an antisaccade task that was first used with adult patients with prefrontal damage[27], Johnson[26] presented young infants with a dynamic and colourful target that appeared at a location opposite a cue preceding the target. The study found a reduction in prosaccades to the cue over a number of trials. Furthermore, toddlers and young children from 8 to 38 months of age have been shown to become increasingly competent at producing antisaccades[28]. Indeed antisaccade development continues well into adolescence and has been shown to become adult-like by approximately 14 years of age[29]. Moreover, top-down executive attention[30], in the form of frontoparietal engagement to select among competing or conflicting alternatives, also continues to develop into adolescence[31,32].

Visual attention-orienting mechanisms[2–4] and more-complex top-down executive control functions[5,33] have largely been treated independently in both the adult and the developmental literature. We incorporate data from both throughout this Review, motivated by their linked emergence over development, their relationships with education and their frequent disruption in developmental disorders. BOX 2 describes the relationship between executive attention and executive control functions, and reviews the literature on how these processes may affect educational achievement. Although the specific tasks used to assess these distinct attention processes may differ, the consensus is that attention, as a process of managing information in a cluttered environment, operates in similar ways at all stages along the information-processing hierarchy, from influencing perception to influencing information held in memory[34]. In the temporal dimension, attentional biases range from transient initial alerting to incoming stimuli to sustained focused attention over prolonged periods[35]. In spatial terms, selective-attention mechanisms bias incoming inputs to enhance the processing of target stimuli and suppress distractors[4]. These competitive biases extend to response-output systems to maintain task-relevant goals in working memory, to inhibit previous or now-irrelevant task goals and to flexibly shift attention across tasks[5,33]. Again, these general alerting, orienting and executive processes do not operate solely in the visual modality, but they are invoked by the need to modulate visual function. It is this specific interaction between vision and attention control over visual processing that we address in the proposed developmental framework below.

Evidence from structural and functional imaging studies focused on large populations or twin samples, as well as from computational modelling, has further described the childhood and adolescent trajectories of brain development, highlighting dramatic changes in local structure, connectivity and genetic and environmental influences on circuits that play a central part in efficient adult attentional states. For example, longitudinal structural neuroimaging data have been used to create four-dimensional quantitative maps of growth patterns in

---

Executive control functions
Functions deployed across modalities to implement task goals, including maintenance of working memory (also known as updating), inhibition of responses (also known as inhibitory control) and cognitive flexibility (also known as shifting).

Attentional biases
Processes by which rich sensory, motor or internally held information is modified by attention to enhance the processing of aspects that are relevant to the task at hand and to inhibit task-irrelevant dimensions.

Box 1 | **The relationship between visual attention development and eye movements**

The neural systems underlying visual attention and eye movements overlap[23,24]. Attentional control of saccadic eye movements involves connections between visual regions (primary visual cortex (V1), secondary visual area (V2), fourth visual area (V4), the parietal cortex and the frontal eye fields (FEF)), the superior colliculus and the basal ganglia[174]. Moreover, many of the neural mechanisms involved in alerting, orienting and executive attention are also implicated in the development of eye movements. This observation suggests that attentional biases operate within a dynamic brain, in which action and perception are closely linked[7,8].

Many aspects of oculomotor control show dramatic but temporally dissociated improvement between birth, 4 months of age[175] and beyond. A pathway from the retina to the superior colliculus is the first component of the system to be operational, with evidence of function of this pathway in newborns, whereas the development of projections from V1 and middle temporal area to the superior colliculus is completed later. An inability to disengage from salient stimuli, referred to as sticky fixation, is present in the first month of life in a typically developing infant and depends on the (predominantly inhibitory) input from the basal ganglia to the superior colliculus in combination with poor cortical control at that very early time point in infancy[175]. Input from the FEF allows for anticipatory looking after 3 months of age[176,177], although this ability continues to develop throughout infancy and can be influenced by experience and training[178]. Smooth pursuit, a process which reflects the tracking of moving visual stimuli, develops over the first 6 months and is related to patterns of sustained attention[179]. By contrast, the ability to suppress orienting towards salient peripheral stimuli emerges at approximately 4 months of age[26] but continues to develop during early childhood[28] and well into adulthood, as indexed by the increasing accuracy in producing antisaccades[180].

The differential maturation of layers of V1 and of projections from these layers to nodes of the oculomotor control network has been suggested to drive the characteristic onset of visual orienting behaviours in infancy[175]. However, this maturational progression does not mean that oculomotor control over the first year of life is entirely driven by changes in feedforward input from V1. Indeed, a role for the prefrontal cortex (PFC) in saccade planning emerges from 6 months of age[181]. Furthermore, functional connections between the PFC and parietal cortex have been measured from 6 months of age[66] and can influence oculomotor control well beyond infancy, as demonstrated by functional magnetic resonance imaging changes in activity from 8 to 30 years of age[182,183]. As we discuss in this Review, this functional connectivity forms the basis of the feedback circuitry that influences the function of lower-level visual cortices.

the developing human brain. These studies have shown that brain maturation is heterochronous and is particularly slow in the frontoparietal cortices[36,37], key regions involved in executive attention. Of note, protracted maturation also occurs in the temporal, occipital and subcortical areas, and in their white-matter connections with frontoparietal areas[38,39]. Additionally, recent studies have found a high degree of maturational coupling between frontal cortical thickness and global cortical thickness, perhaps because the frontal cortices subserve integrative functions that require coordination with a large proportion of the cortical sheet[40]. Furthermore, analyses of longitudinal imaging data collected from twins have shown that genetic and environmental contributions to the variance in cortical thickness change over the course of childhood, with more variance being accounted for by environmental factors early in childhood, especially for the dorsolateral PFC, and greater genetic contributions to variance being seen at later points[41]. Finally, the frontoparietal and cingulo-opercular networks associated with attentional control have distinct patterns of development during childhood and into adolescence and adulthood. In both networks, development involves decreases in local connectivity and increases in long-range connectivity[32,42,43]. In summary, the development of visual attention may occur alongside protracted and distributed changes in brain structure, function and connectivity.

A proposed framework for visual attention development.
As described above, much is known about the time course of development of brain regions and connectivity involved in vision, visual attention and the attentional modulation of visual systems. However, little is known about how

these processes interact as they develop and become fully functional. Here, we apply several principles described in the fields of visual neuroscience and computational vision to this problem. First, anatomical connections in the visual cortical hierarchy convey information forwards and backwards from one region to the next[44,45], so indirectly link the primary visual cortex (V1) to the PFC[46]. There are distinct hierarchical streams in the visual system for the analysis of motion (dorsal pathway) and colour (ventral pathway)[45] (FIG. 1). Second, computational models of vision and attention make use of the concept that rostral regions in the visual cortical hierarchy integrate inputs from caudal regions and thus process increasingly complex aspects of stimuli, space and even abstract rules for action[47–49]. Third, the development of behaviour, functional connectivity and grey-matter volume, as discussed in the previous section, approximately mirrors this cortical organization, with development beginning caudally and becoming increasingly rostral over time. For example, visual-processing cortical areas mature early in development, followed by maturation of parietal and temporal regions, which support spatial and object-based attention, respectively, and then by PFC regions, which are involved in executive attention. Together, these observations have led us to propose a framework that embeds visual attention development into the emerging functionality of this hierarchical architectural organization of visual pathways.

Connections among cortical areas of the dorsal and ventral visual pathways have been mapped in the non-human primate brain and found to be organized hierarchically[45]. Visual information enters the system via the lateral geniculate nucleus (LGN) and is first processed in

Box 2 | Linking attention development and education

Effective executive attention involves a series of processes commonly referred to as executive control functions, including working memory, inhibitory control and cognitive flexibility. A defining principle of executive attention is that behaviour is directed by a rule or the achievement of some goal. A simple example is one in which a parent may ask a small child to pick up the red ball, and a competing green ball may be nearby. The child would need to suppress the action of picking up the competing green ball (which shares its shape and function with the red ball) to complete the goal-oriented action. This task may be more difficult than if the competing toy were a stuffed teddy bear, for example. To act appropriately, the child must maintain the goal (to pick up the red ball) in working memory and suppress or inhibit distraction (from the competing green ball). In addition, over time the child might need to switch flexibly between rules (to sometimes pick up the red ball, and at other times pick up a different toy). This example reflects differing but overlapping kinds of conflict among percepts, responses or rules that need to be resolved by an executive attention system.

These executive processes are known to facilitate educational attainment. For example, multiple studies suggest that working memory skills, and in particular the contribution of executive attention to those skills, are a significant concurrent and longitudinal predictor of educational outcomes, especially in mathematics[93], that are independent of individual differences in intelligence[89]. Even before the onset of formal instruction, executive and attentional skills provide preschoolers with a head start when in school, especially in numeracy[184,185]. The stronger relationship between executive attention development and numeracy compared with executive attention development and literacy may be due, in part, to the visuo-spatial nature of numerical constructs acquired early in childhood. What mechanisms underlie the robust relationships between executive attentional control and educational outcomes? Growing evidence highlights the role of attentional biases in encoding and maintenance of information in working memory. For example, electrophysiological[91] markers of attentional biases in preparation for encoding into memory and resting-state functional connectivity[186] correlate with working memory capacity in 9–11-year-old children. Furthermore, magnetoencephalography shows that frontoparietal oscillations before encoding predict the accuracy of later memory in the same age group[90] (FIG. 2). These findings suggest that individual and developmental differences in the ability to deploy attention, and their neural correlates, constrain the efficiency of memory processes and, in turn, may influence classroom learning.

parallel by visual areas governing the processing of specific features (such as motion), then it feeds forward and converges on the other cortical nodes (FIG. 1) within the hierarchy[50]. For example, from a computational perspective, the role of the parietal cortex in the dorsal pathway is thought to be the integration of information from multiple feedforward or bottom-up local-processing visual streams and to resolve the competition between these inputs in a topographically meaningful manner to allow for selection and allocation of spatial attention[47]. Similarly, the ventral pathway from V1 to the inferior temporal cortex (IT) allows for increasingly complex object representation through pooling of inputs from lower levels along the pathway[48,49]. It is also known that the function of some cortical areas, including V1, is modulated by the top-down or feedback signals generated in relatively high-level regions[51,52]. This top-down modulation of V1 acts as a form of gain control over visual processing and results in improved quality of low-level vision, enhanced contrast sensitivity, improved acuity and improved perceptual processing of attended information[51]. Notably, the reciprocal feedforward and feedback connections are anatomically distinct, originating and terminating in different cortical layers[52]. In the stable adult state, the hierarchy is better described as a series of parallel loops reverberating across cortical circuits, with no obvious beginning or end[46,52].

Feedforward
Efferent flow of information away from a lower cortical region to a higher cortical region.

Feedback
Afferent flow of information from a higher cortical region to a lower cortical area.

We therefore propose a framework whereby this hierarchical organization[53] develops into this stable state over human ontogeny. Mechanistically, the cumulative development of visual areas feeding forward into higher-level regions may function as the catalyst for top-down attentional modulation of these same visual pathways. This top-down attentional modulation also functions as a form of gain control over visual processing and results in improved quality of early vision, enhanced contrast sensitivity, acuity and perceptual processing of attended information[51]. Any disruption to local visual organization, for example in the form of disruption to pyramidal cell or interneuron populations, may then disrupt sensory-driven dynamics and affect top-down or feedback-loop organization[54]. In turn, as feedback-loop integrity shapes perceptual learning through changes in expectation and attention[55], initial low-level changes would result in the local and long-range changes in functional connectivity and network integration that characterize typical visual attention development (from simple visual orienting to executive attention processes) and might also contribute to multiple developmental disorders.

This proposal is consistent with the timing of the development of vision, visual attention and executive attention modulation of visual processing. Recent studies have shown that visual-orienting processes that depend largely on simple feedforward and feedback architecture develop during the first postnatal year, and that their functionality is predicted by improvements in local visual-feature processing[56]. In addition, vision is poor in newborn humans, and most visual skills improve rapidly during the first 6 months. Acuity is estimated at 20/800 for most newborns[57] but improves quickly over the next few months, as does contrast sensitivity[58]. Full motion sensitivity is noted by 6 months[59], whereas it is known that infants are sensitive to orientation shifts as early as 6 weeks after birth[60]. It seems reasonable to suggest that the cumulative functional development of visual processing areas is a prerequisite for the development of higher-level visual computations, such as visual attention, that resolve competition between visual elements in a scene[4]. As an analogy, imagine that you must make a decision. When there is only one option available, there is no decision-making required and so there is no challenge to the decision-making circuitry. Imagine now that additional options are added. Now, the competition between these options needs to be resolved, which requires the use of the relevant decision-making circuitry, including long-range cortical top-down connections. Similarly, in the case of visual attention, in early postnatal life, vision is poor and feedforward visual information conveyed to higher cortical areas is minimal. We hypothesize that, with visual development, there is an increase in feedforward information competing for attention allocation in higher-level regions, thus linking top-down visual attention development with visual experience[56]. In turn, these regions, now engaged, send top-down signals to begin to tune local visual areas, setting the hierarchical loops in motion from very early in the first postnatal year. This developmentally timed
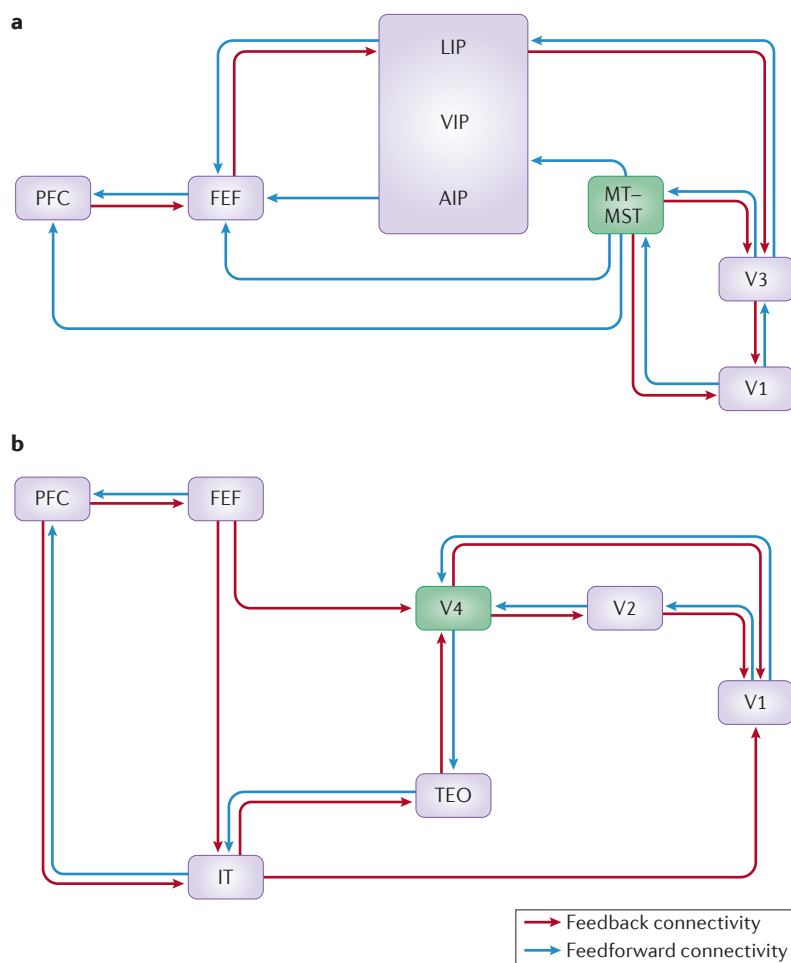
REVIEWS

Figure 1 | **Primate dorsal and ventral visual pathways and possible sites of disruption.** Disruption to the local architecture and organization of specific visual areas may have effects on circuit development. This figure shows a simplified overview of feedforward and feedback connectivity between visual areas and more-rostral cortical areas, including parts of the parietal, frontal and temporal cortices involved in visual attention processes[46,197]. For simplicity, we omit here the reciprocal connections between the dorsal and ventral pathways. **a** | Over the course of development and hierarchical cortical organization, disruption to the local organization of motion processing via the middle temporal area (MT; shown in green) may result in disrupted feedforward- and feedback-loop architecture integrity both in regions involved in executive attention, through weaker connections to the prefrontal cortex (PFC) and frontal eye fields (FEF), and in regions involved in visuo-spatial attention orienting, through the parietal cortex (the lateral intraparietal area (LIP)). **b** | Similarly, over the course of development, disruption at the level of the fourth visual area (V4; shown in green) could result in weaker long-range connectivity through the ventral visual pathway, disrupting the hierarchical feedforward and feedback organization of executive attention processes through the PFC and disrupting object-based visual attention through the inferior temporal area (IT). AIP, anterior intraparietal area; MST, medial superior temporal area; TEO, temporo-occipital cortex; V1, primary visual cortex. V2, secondary visual area; V3, third visual area; VIP, ventral intraparietal area.

cascade model may explain the emergence of visual attention in the 3–6 month age range[14,61], in the sense that it must occur after the critical first several weeks of cumulative visual development.

This ontogenetic model is also consistent with a recent phylogenetic description of cortical organization. Finlay and Uchiyama[62] present evidence that a central

principle in phylogenetic change is organization around a rostrocaudal axis. Moreover, the same authors have suggested that the convergence of inputs from hierarchically organized cortical areas onto the frontal lobes, which are also hierarchically organized[63], may be critical to executive control functions (including the executive control over visual processing discussed here) in species with larger brains[62]. We propose that visual attention development fits this framework very well.

A similar ontogenetic argument can be inferred from studies of functional neural network integration versus segregation in developmental connectomics[64]. This novel approach has shown that short-range sensorimotor connections best-characterize the infant brain early in postnatal life[65], whereas long-range functionality develops later. Of course, even within infancy, there is evidence of continued development of frontoparietal network connectivity between 6 and 12 months[66]. This development is followed by continued organization of cortical long-range connections involving increasingly rostral cortical areas into childhood and adolescence[67–69]. However, connectomics data from young subjects must be evaluated carefully, as there are concerns that head motion during MRI is greater at younger ages and thus that these studies may find greater short-range connections in younger children than there really are[70]. Nonetheless, there is growing evidence from studies controlling for these artefacts to suggest that there are increases in long-range region connectivity during development, including increased myelination and white-matter integrity that would facilitate longer-range communication[38–40]. With respect to executive control functions, studies of effective connectivity in children find that strengthening of long-range connectivity from parietal to frontal regions and decreases in short-range connectivity within parietal and frontal regions occur in parallel with improvements in executive regions[32].

Visual attention development and memory
Visual attention orienting is one of the first coordinated active exploration systems to develop in human postnatal life and serves several functions. It allows sequences of individual visual images, obtained across successive saccades, to merge for scene coherence[71,72]. It also determines which information is selected for processing from complex cluttered environments, so supports learning from the currently attended location based on task goals. In addition, visual attention processes support suppression of distraction from the location of the previous locus of attention, which is necessary when the previously attended location is still in the field of view or when there is interference from a lingering memory trace of the previous focus of attention[73,74]. It follows that visual attention deployment should have functional consequences on learning and memory. This is supported by developmental studies, which show that as distinct attention processes develop so too do separable learning and memory processes[75–77]. Attention mechanisms are involved in encoding visual short-term memory (VSTM)[75], maintenance in working memory[77] and long-term recognition memory[78]. These distinct forms of memory can be dissociated at

different developmental stages, both in cognitive and in neural terms. In turn, as memory traces become long-term memories, information held in memory influences attentional selection, as demonstrated by contextual cueing of visual attention, for example[79].

External cues can successfully direct infants' attention orienting to perceptually salient information from 3 months of age[16,19]. The emergence of this orienting mechanism is relevant to VSTM abilities[75]. VSTM allows for fluid integration of information across successive saccadic eye movements[80]. VSTM is known to have very stringent capacity limits in early infancy[81,82] (that is, as little as one item of information can be maintained in VSTM at one time), but this can be overcome if external cues are used to orient attention to the stimulus location before its onset[75]. By contrast, orienting to a stimulus while simultaneously suppressing previously attended competing distraction begins to emerge later, between 4 and 6 months of age[15,16,19], and it is coupled with robust encoding of attended items for subsequent recognition[76,83]. This has been shown in studies in which heart rate deceleration is used as an index of sustained focused attention[84], as well as in studies in which an inhibitory mechanism ('inhibition of return') is experimentally elicited by manipulating timing parameters of attention cues, to engage suppression of a previously cued location while objects are incidentally encoded in a currently attended location[76]. In 9-month-old infants, suppression of the previously cued location during object encoding enhanced subsequent recognition memory for objects placed in attended locations, whereas identical tasks that simply facilitated orienting to the cued location without concurrent suppression of the distractor did not[76]. Ostensibly, this balance between attentional enhancement and distractor suppression supports a more-robust visual signal for downstream encoding[51,85].

Furthermore, as discussed above, aspects of top-down executive attention continue to develop well into adolescence and adulthood[31,32], and this may have implications for working memory at these developmental stages. For example, neuroimaging data have shown that attention-related top-down frontoparietal modulation of visual regions was reduced in 8–12-year-old children relative to in adults, and this reduced modulation related to poorer working memory performance[86]. This suggests that the ability to maintain information in working memory, in a manner that is resistant to distraction, is supported by top-down prefrontal modulation of areas involved in stimulus processing. These data can be dissociated from attentional effects on encoding into VSTM[87,88]. Indeed, although directing attention at encoding provides developmentally robust advantages for encoding into both short-term[77,89,90] and long-term memory[78], attentional influences on maintenance in working memory are less efficient in 6- and 11-year-olds compared with in adults[77,87,89]. Furthermore, attending to the contents of working memory facilitates the accuracy of memory report both in children and in adults[77,89]. However, children deploy attention less efficiently in working memory. For example, 11-year-olds show differences in the dynamics of attentional influences on working

memory compared with adults[91] (FIG. 2). However, the observation of adult-like frontal electroencephalography (EEG) signatures when orienting attention to memory[91] and the engagement of a frontoparietal network at around the time of encoding, as measured with magnetoencephalography (MEG), predict the accuracy of their later memory[90]. The deployment of attention in function of maintenance in working memory is the aspect of attentional modulation that is most dependent on prefrontal engagement, is most protracted in its developmental course and has the most-direct links to educational attainment[92,93] (BOX 2).

Of note, the interaction between visual attention and higher-order control functions is bidirectional: the development of higher-order cognitive functions (for example, long-term memory formation) also influences the deployment of visual attention. It has long been known that information previously encoded into long-term memory and categorical knowledge can guide attention in adults[94–96], but recently it has been shown that this also occurs in childhood. For example, in a modification of the now classic contextual cueing paradigm[97], children as young as 5 years of age were found to direct attention more efficiently when guided by information held in long-term memory[79]. In this paradigm, visual search for targets embedded in repeated scenes is more efficient compared with visual search for a novel target not associated with a contextual memory trace. Furthermore, attention orienting is most effective for memoranda for which familiar representations are available in long-term memory, both in children and in adults[98].

Atypical visual attention development

Overview. Abnormal attention is a symptom of several disorders, including attention deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD) and fragile X syndrome (FXS). The developmental cognitive neuroscience of these and other attentional disorders has increasingly shifted from a focus on individual brain regions involved in attention towards the study of connections between these areas and how these connections develop. For example, in the case of ADHD, the initial targets of investigation were the frontostriatal circuitry[99] and the dopaminergic system, as well as its candidate genetic moderators[100]. This emphasis was justified by the fact that the majority of cases respond to methylphenidate, a dopamine-reuptake inhibitor[101]. Furthermore, structural imaging data had suggested that individuals with ADHD have reduced frontostriatal volumes[102]. However, focusing solely on frontostriatal areas in ADHD has been questioned as a result of findings from behavioural[103], systems neuroscience[104] and genetics studies[105]. For example, structural abnormalities observed in individuals with ADHD include not only frontal cortical thinning but also global cortical thinning[106,107]. Children with ADHD show decreased functional connectivity in frontoparietal networks and increased local efficiency within networks. The term 'neural efficiency' remains to be clearly operationalized[108], but these connectivity data seem to suggest that in ADHD there is increased network segregation and decreased long-range connectivity,

Connectomics
An emerging field that identifies functional coupling of brain regions to form networks by assessing correlated activity using functional magnetic resonance imaging analyses.

Contextual cueing
A visual search paradigm designed to improve attention selection of targets that appear repeatedly in the same scene (context) compared with attention directed towards targets that appear in novel contexts.
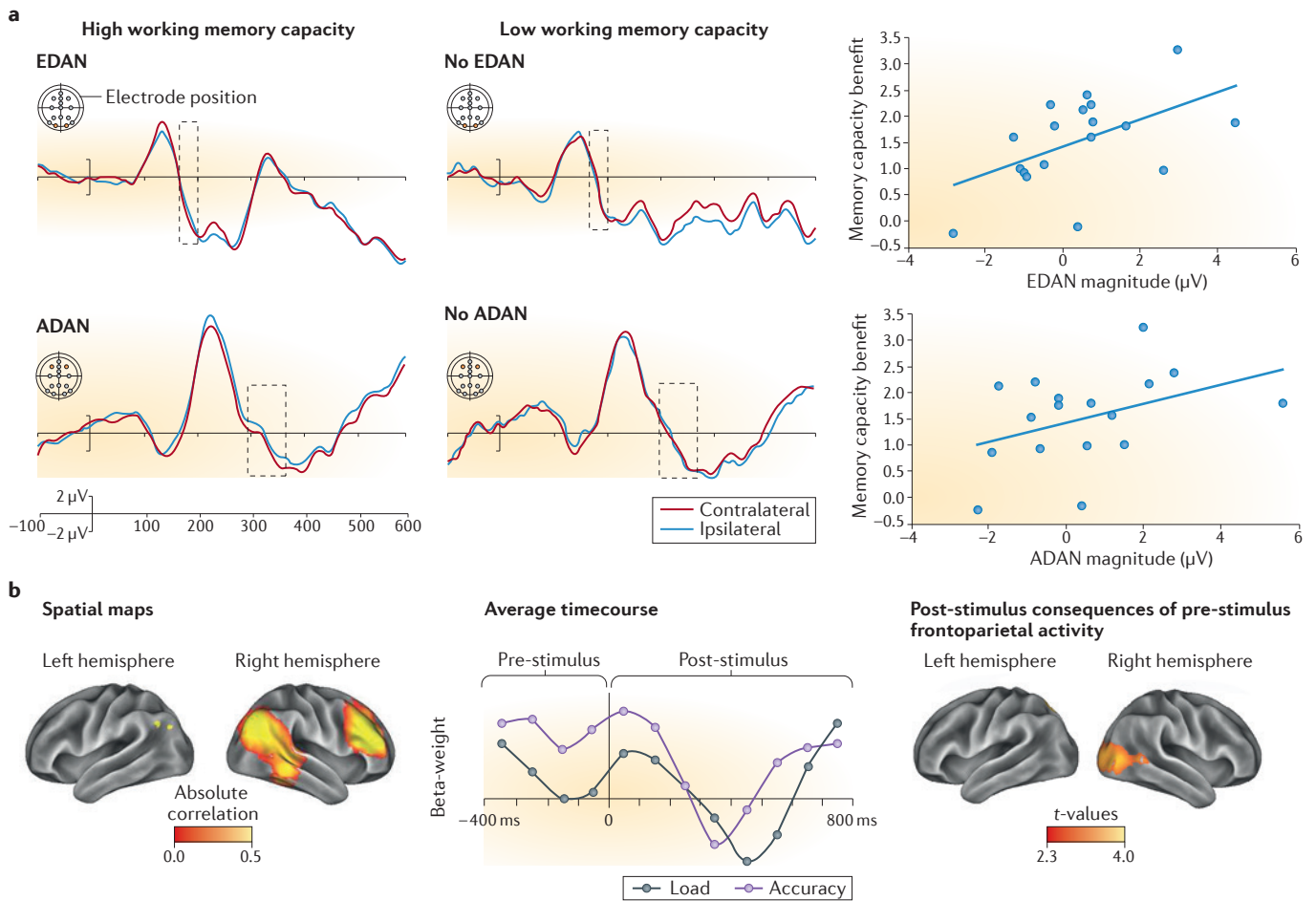
a

High working memory capacity | Low working memory capacity

b

Spatial maps | Average timecourse | Post-stimulus consequences of pre-stimulus frontoparietal activity

Figure 2 | **Visual attention correlates with working memory capacity.** Electrophysiological and magnetoencephalographic studies provide evidence that variation in the neural markers of attentional deployment correlate with individual differences in memory capacity. **a** | Electrophysiological markers of visuo-spatial attentional orienting in preparation for encoding information into memory distinguish 10-year-old children with higher or lower working memory capacity. The early directing attention negativity (EDAN) is an event-related potential locked to the onset of spatial cues that direct attention, and it is characterized by greater negativity at posterior scalp electrodes that are contralateral than at posterior scalp electrodes that are ipsilateral to the direction of the attention-orienting cue. EDAN is thought to indicate cue-processing. Another event-related potential, anterior directing attention negativity (ADAN), is also characterized by greater negativity at scalp electrodes contralateral to cue direction, but at electrodes that are more anterior than those used for EDAN. ADAN is associated with deployment of attentional control. The waveforms (left-hand and central panels) represent the average time course of these differences for children with high working memory capacity (who do show EDAN and ADAN); and children with low working memory capacity (who do not show EDAN and ADAN). The area marked with a dashed box highlights when the waveforms differ for contralateral and ipsilateral sites for adults and children with high working memory capacity, but not children with low working memory capacity. The scatterplots (right-hand panels) show significant correlations between the magnitudes of EDAN and ADAN and the benefits of cues for memory on this task. **b** | Magnetoencephalographic data suggest that the preparatory oscillations of a right frontoparietal network before encoding items into memory predict the accuracy of later memory recall and the activity of visual cortices when the memoranda are first encoded. Adults and 10-year-old children were asked to encode either two (low load) or four (high load) simultaneously presented items into memory and then to recall whether a probe item was among the memoranda. The children's spatial maps of right-frontoparietal network oscillations are shown on the left and the time course of the effect of these oscillations on memory accuracy are shown in the centre. Represented on the x-axis is time, with 0 indicating the time point at which the to-be-encoded items were presented. The y-axis represents beta weights. Pre-stimulus activity in this frontoparietal network in preparation of encoding successfully discriminates trials in which participants remember items accurately (purple line). By contrast, this frontoparietal network is not differentially engaged by memory load (an index of task difficulty; grey line). The panel on the right represents the area in the children's visual cortex whose activity after the onset of the to-be-encoded stimuli was significantly predicted by right frontoparietal network activity, illustrating the coupling that occurs between this network and the visual cortex. Part **a** is adapted from REF. 91. Reprinted by permission of MIT Press Journals, © 2014 Massachusetts Institute of Technology. Part **b** is adapted from Astle, D. E. et al., The neural dynamics of fronto-parietal networks in childhood revealed using magnetoencephalography, *Cereb. Cortex*, 2014, doi:10.1093/cercor/bhu271, by permission of Oxford University Press.

indicating poor integration[109]. Furthermore, a meta-analysis of over 50 functional neuroimaging studies[110] comparing children and adults with ADHD with neurotypical controls has revealed abnormalities in several areas, including the frontostriatal circuitry but also the visual and default mode network[111,112]. ASD, another developmental disorder, is also in part characterized by atypical attention and is now also considered a network disorder[90], although the networks disrupted in ASD may be different from those disrupted in ADHD[113–115]. Thus, our understanding of typical attention development is not complete if it centres entirely on individual areas or even isolated circuits, rather than on networks. In other words, studying just a single node in any network (for example, the PFC) would provide an incomplete picture of brain regions involved in attention.

A fuller picture emerges as we instead examine a developmentally titrated model. The organization of the developing brain begins to influence visual attention development from infancy, potentially through changes in the cortical hierarchical organization of the ventral and dorsal pathways described above (FIG. 1). By studying these pathways, one can make specific predictions about how disruptions to the development of brain regions involved in visual processing early in postnatal life may affect downstream attentional network function (FIG. 1). Indeed, impairments in visual attention and visual processing are a common feature of several neurodevelopmental disorders. Three hypotheses suggesting that attentional disruption occurs in these populations in association with atypical perceptual processing have been proposed: the hypothesis that atypical global motion processing characterizes various developmental disorders (the dorsal stream vulnerability hypothesis[116]), the enhanced perceptual functioning hypothesis[117,118] and the atypical neural noise accounts of autism[119]. However, particularly in the case of autism, these accounts have been challenged by other hypotheses that top-down and feedback influences are more important than feedforward influences in understanding atypical perception and attention[120]. Here, we point out that the origin of the disruption is difficult to disentangle by studying the adult state alone, because it could very well arise from earlier feedforward abnormalities, feedback abnormalities or both, even if in the adult only one of the two types of disruptions can be isolated. A resolution must come instead from considering the potential developmental origins of disruptions. The recent findings on early visual, perceptual and attentional development and their neural correlates in very young children with autism and infants at high familial risk for autism are beginning to address these questions[121–125].

Genes implicated in attention dysfunction. Research on the genetic basis of differences in attention in healthy individuals (both adults and children) initially focused on common polymorphisms for a small number of genes regulating the efficiency of neurochemical metabolism, such as variants of *DAT1* (also known as *SLC6A3*; which encodes the sodium- and chloride-dependent

dopamine transporter), *DRD4* (which encodes the D4 subtype of the dopamine receptor) and *COMT* (which encodes catechol-*O*-methyltransferase and is involved in monoamine synthesis)[126,127]. Similarly, investigations of the genetic basis of complex neurodevelopmental disorders that are diagnosed by their behavioural symptoms, such as ADHD and ASD, initially focused almost exclusively on variants in individual candidate genes regulating neurotransmitter efficiency[128,129].

Recently, a more-complex view has emerged owing to studies of attention during development. Large-scale studies of the genetic basis of individual variability in attention and clinical risk for ADHD have shown that the functional outcomes of distinct monoamine-related gene variants differ at various stages of development, both in the healthy population[130] and in individuals with ADHD[131]. For example, the *COMT* Met variant results in higher dopamine availability, but the benefits of carrying the Met variant compared with carrying the alternative *COMT* variant Val (better performance and reduced PFC activation during working memory) emerge only after 10 years of age[130]. Furthermore, although the 10/10 genotype of *DAT1* is thought to be a risk factor for ADHD in children, the 9/9 genotype is associated with persistent ADHD in adulthood[131], which again suggests a complex developmental picture. In addition, genome-wide association studies of risk for neurodevelopmental disorders like ADHD and ASD have shown that common variants of individual candidate genes have small influences[105,132], highlighting that genetic risk must instead be studied in the context of polygenic risk factors, both in individuals with attention disorders[133,134] and in the neurotypical population[135]. Furthermore, these studies do not identify genes directly associated with neurotransmitter regulation but rather genes that regulate the establishment of local and long-range connectivity over development, such as those implicated in dendritic neurite outgrowth[136]. Thus, genomic data highlight that a one-to-one mapping between the genetics of neurotransmitter function and attentional dysfunctions is unlikely[137] and that models of genetic influence on attention must address how genetic variability acts on the development of local and long-range connectivity, which, as discussed above, is central to attentional development.

Moreover, studying the genetics of neurodevelopmental disorders highlights further points to consider when investigating the influence of genetics on attentional development. First, it has become clear that the candidate gene approach is not helpful in the context of understanding risk for attentional disruptions or susceptibility to adverse environmental conditions, such as prenatal stress[138]. Instead, polygenic risk clusters in broad functional gene networks. For example, recent studies on copy number variation have identified multiple converging functional gene regulatory networks associated with risk for ASD[139,140]. These gene networks overlap somewhat but are still distinguishable from functional gene networks that confer risk for ADHD and schizophrenia[132,141]. Remarkably, these networks do not often directly implicate the neurotransmitters that

Polygenic risk
Genetic risk for a particular phenotype (for example, the likelihood of attention deficit hyperactivity disorder diagnosis) captured as the cumulative effect of differences at multiple genetic loci.

Functional gene networks
Genes operating in concert to regulate particular neural or developmental functions (for example, dendritic dynamics and receptor clustering, intracellular transport and regulation of gene transcription).

were initially the targets of the candidate gene studies (such as dopamine-related gene variants) but instead point to disruptions in gene networks implicated in setting up network dynamics and their vulnerability, a shared factor across these disorders. These studies highlight possible interactions between the expression of susceptibility genes and endogenous maturational changes in the availability of neurotransmitters like dopamine[142], with atypical dopamine availability putting certain populations at risk for atypical development of attention.

Powerful illustrations of this point come from the development of attentional difficulties in individuals with rare, high-penetrance genetic mutations associated with severe and complex neurodevelopmental disorders. These can be informative, because mechanisms can be studied from the single gene to the symptom level[143]. For example, the fragile X mental retardation 1 (*FMR1*) gene is silenced in FXS[144], an inherited condition associated with inattention and hyperactivity (BOX 3). The associated protein, FMR protein (FMRP), is a key regulator of glutamatergic[145] and GABAergic balance[146], as well as synaptic development and function, as demonstrated by dendritic spine dysmorphology and altered synaptic plasticity in FXS[147,148]. Although the primary effects are on intrinsic neurotransmitter regulation, not dopamine, the balance in extrinsic neurotransmitters (like monoamines) is also affected[149] and computational properties that are central to the development and function of frontoparietal connections[150] are compromised. One of the strengths of studying individuals with

genetically identified developmental disorders associated with high risk for attention difficulties, such as FXS, is that prospective longitudinal data can be gathered from infancy, in both humans and animals, long before ADHD (or ASD) diagnoses can be attained.

Changes in computational constraints on neural development and functioning like those characteristic of FXS (BOX 3) sit at the convergence of risk for disorders like ADHD[151] and ASD[140,152]. High-penetrance mutations may have converging (or diverging) effects on attentional functions, depending on the specific ways in which they regulate neural development, neurophysiological properties and network emergence, as has been proposed in the context of similarities and differences between tuberous sclerosis and FXS[153]. The body of information about how these individually rare but cumulatively rather common mutations affect attentional control networks is growing[143]. A fruitful approach to use this information is to group rare genetic differences that affect common networks and test their effects on attentional control skills in comparison with relatively well-understood abnormalities like those measured in FXS[154].

In summary, to date, the gene networks implicated in attention impairments seem to modulate the dynamic hierarchical organization of the cortex and connections that underlie the development of attention, rather than predetermining attentional control directly. Thus, such gene networks are best thought of as playing a part in heightened susceptibility to developing attention disorders.

Box 3 | Rare genotypes, gene functional networks and risk for attention disorders

Studying individuals with rare but highly penetrant genetic variants associated with high risk of attention difficulties from early childhood can provide insight into the genetic, cellular and systems mechanisms of risk, because these individuals can be studied at multiple levels, from genetics to behaviour[143]. For example, fragile X syndrome (FXS) is a monogenic neurodevelopmental disorder affecting 1 in 4,000 males and 1 in 6,000 females[187] and is associated with a very high risk for attention deficits. The gene that is silenced in FXS, fragile X mental retardation 1 (*FMR1*), encodes FMR protein (FMRP), which regulates glutamatergic[145] and GABAergic[146] balance, suggesting that the attention impairments observed in FXS are not directly related to dopaminergic signalling. At the systems level, function of frontostriatal and frontoparietal cortices is atypical in individuals with FXS[188,189], but several other networks are also affected[190]. Studies using *Fmr1*-knockout mice also suggest that the embryonic development of broad neuronal networks is atypical in FXS[191,192]. FMRP is expressed widely across cortical and subcortical circuits, an observation that would predict global impairments. At the cognitive level, individuals with FXS show difficulties in attentional control and working memory from childhood, both cross-sectionally[193] and longitudinally[194,195]. How could these relatively specific cognitive level deficits emerge, in contrast with broad and global impairments in neural function? As outlined above, *FMR1* silencing affects synaptic development and results in immature dendritic spine development. These changes in turn may alter a computational property that is essential to the development of higher-order circuits involving the parietal and prefrontal cortices[150]. Of note, developmental time is an essential factor to consider in understanding pathways to attention risk: FXS can be diagnosed early in infancy or childhood, making it possible to study trajectories from an early age, including early perceptual processing abnormalities in spatiotemporal integration[196], impairments in basic eye-movement control[28] and difficulties in basic attention processes that predict later attention deficit hyperactivity disorder (ADHD) symptoms[194].

Functional gene networks implicated in pathways such as those dysregulated in FXS are complex and have multiple components. Different genetic mutations may have converging (or diverging) effects on attention, depending on the specific ways in which they regulate network development. A fruitful approach is to group rare mutations associated with attention deficits according to their putative affected networks and test their effects on attentional control skills. For example, individuals with mutations in genes encoding membrane-associated guanylate kinases (which regulate synaptic plasticity function) display hyperactivity and autistic-like symptoms that are similar to some of the symptoms of FXS, but these individuals seem to have attention profiles distinct from those observed in individuals with FXS[154]. These findings suggest that the mechanisms underlying attention function and dysfunction can be understood by studying distinguishable molecular pathways disrupted in people with rare mutations associated with attention disorders.

Environmental influences and training. Several lines of evidence suggest that the developmental architecture of attentional processes is plastic. Here, we discuss these complementary bodies of work, as well as novel opportunities and caveats for attention training and intervention, in the context of both typically developing individuals and individuals with developmental disorders affecting attention. The potential for effective training of attention has attracted much interest from initially rather different groups with distinct agenda: practitioners focused on improvements in attention outcomes in real-world environments and neuroscientists interested in the mechanisms of attention plasticity. Initial excitement about the modifiability of attentional processes emerged through diverse but complementary 'natural experiments' charting the effects of environmental differences on attention. Pioneering studies demonstrated that congenitally deaf individuals have better peripheral visual attention than those with hearing[155,156] and that variation in executive attention in healthy individuals is associated with socioeconomic status, which may incorporate a number of environmental factors[157–159]. These findings suggested that some attentional mechanisms, namely top-down executive attention processes, are heavily shaped by the environment. However, it is difficult to directly attribute these effects on executive attention to plasticity induced by altered environmental exposure because the target populations (for example, congenitally deaf individuals) might also be characterized by other neural or cognitive differences.

The effects of attention-training regimes are better studied by randomly allocating individuals to distinct exposure regimes. The finding that expert adult video-game players differed in cognitive and neural markers of executive and spatial visual attention compared with non-expert players[160,161] led to additional studies of the effects of video-game exposure on attention in naive players. When naive individuals were first exposed to a gaming regime, improvements in low-level visual processing and spatial attention were observed, although effects on executive attention were weaker[162]. Recently, similar training experiments have begun to study attention-training regimes to investigate the malleability of attentional mechanisms and their neural correlates from childhood[163,164]. The training regimes used in these studies typically involve prolonged exposure to computerized attention games, which aim to stretch the level of ability of individual participants by becoming increasingly challenging[164]. These training programs have been particularly successful in training executive attention and related functions, such as working memory[164,165], in neurotypical adults and young children[166], and in children with neurodevelopmental disorders such as ADHD[167], although there have been failures to replicate these studies[168].

Given the interaction between attention and learning and memory processes, one would expect that training attentional processes would transfer organically to associated learning and memory systems. However, surprisingly, transfer of training benefits to untrained neurocognitive processes (for example, mathematical

ability or intelligence) or behaviours (for example, hyperactivity and inattention in the classroom) that are known to relate to attention processes has been difficult to demonstrate convincingly, as there have been conflicting findings in these studies, as indicated by recent meta-analyses and systematic reviews[169–171]. Why is transfer of attention training to related functions such as mathematical ability or behaviour in the classroom ineffective? We suggest that, in addition to possible methodological limitations to existing training regimes (including limited attempts to follow-up training benefits over time, difficulties in choosing pre- and post-training assessment measures and difficulties in designing an active control regime against which to compare training effects), these failures lie in an incorrect core assumption: that a repetitive 'diet' of attention tasks, training certain processes through repetition, will automatically generate transfer. In this Review, we suggest that the emerging efficiency of connections between executive attention control regions and more-specialized regions supporting the specific tasks towards which transfer is aimed is an important part of the development of adult attention. If this is correct, attention training that is devoid of a focus on its relationships with specific processes (for example, numerical processing, visuo-spatial processing and perceptual processing) will not transfer easily to these skills. Perhaps this flaw of attention-training regimes is best epitomized by an analogy: current attempts focus on training attention as if attention was a specific muscle, or set or muscles, so they train attention as a body-building regime might train a specific muscle. Instead attention-training regimes should aim to be analogous to training a dancer, who successfully coordinates skill interplay across specialized and general systems.

One consequence of the proposed framework is that it suggests novel strategies for improving outcomes in children with neurodevelopmental disorders of attention. If attention-training regimes place an emphasis on visual feedforward processes as well as on low-level orienting mechanisms, it may be possible to subsequently achieve better cortical integration and network connectivity. Studies of attention during ageing — which is associated with a decline in perceptual and working memory processes — suggest that such a training approach may be fruitful. Gazzaley and colleagues have shown that, in older adults, training interventions to improve the perceptual precision of stimulus representations also resulted in improvements in working memory[172].

We suggest that repetitive attention training does not automatically improve long-range corticocortical connectivity and integration, which are necessary for plasticity and transfer to untrained functions, both in childhood and in the ageing. Indeed, following intensive working memory training in neurotypical children, individual differences in the magnitude of transfer to untrained tasks correlate with changes in network connectivity at rest, and in particular connections between attentional frontoparietal and specialized processing areas in the IT[173]. Perhaps the best way to facilitate transfer of attention training is to first understand better the

---

Transfer
The outcome of a cognitive or neural training regime that may improve untrained tasks that use the specific skill being trained (such as attention), improve closely related functions (referred to as narrow transfer) or improve more-distally related system functions (referred to as wide or far transfer; for example, mathematical achievement or intelligence improving after attention training).

mechanisms through which functional connections between attentional networks and specialized networks are modified by training.

Conclusions and future directions

The findings on attention development discussed in this Review highlight how studying attentional circuits or processes in isolation is not sufficient. This is because reaching the efficient adult attentive state involves the coordination of perceptual development, the strengthening of functional connections and interactions with memory processes. A full understanding of attention in the adult therefore requires an understanding of developmental trajectories. Here, we have highlighted how attention influences memory processes over the course of development and vice versa. Furthermore, our overview of visual attention in neurodevelopmental disorders highlights the interplay of genetic and environmental influences on visual attention mechanisms. This type of interdisciplinary approach is critical to understanding visual attention and ultimately developing treatments for disorders in which visual attention is impaired, including effective attention-training protocols.

An emerging future direction for the cognitive neuroscience of attention may therefore be the identification of the developmental origins of attention dysfunction, in the hope of rehabilitating a less-than-efficient system through the strengthening of relevant network connectivity from the ground up. This strategy has the potential to improve visual attention network dynamics and thus the learning and memory mechanism with which they are coupled. Beyond attention training specifically, we argue more broadly that the development of attention function occurs through the functional coupling with perceptual and memory systems. Cognitive neuroscience may therefore benefit from focusing on the coupling of systems rather than on treating these processes separately.

1. Fan, J., McCandliss, B. D., Sommer, T., Raz, A. & Posner, M. I. Testing the efficiency and independence of attentional networks. *J. Cogn. Neurosci.* **14**, 340–347 (2002).
2. Petersen, S. E. & Posner, M. I. The attention system of the human brain: 20 years after. *Annu. Rev. Neurosci.* **35**, 73–89 (2012).
3. Posner, M. I. & Petersen, S. E. The attention system of the human brain. *Annu. Rev. Neurosci.* **13**, 25–42 (1990).
4. Desimone, R. & Duncan, J. Neural mechanisms of selective visual-attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
5. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
6. Matusz, P. J. *et al.* Multi-modal distraction: insights from children's limited attention. *Cognition* **136**, 156–165 (2015).
7. Rizzolatti, G., Riggio, L., Dascola, I. & Umilta, C. Reorienting attention across the horizontal and vertical meridians — evidence in favor of a premotor theory of attention. *Neuropsychologia* **25**, 31–40 (1987).
8. Pineda, J. A. Sensorimotor cortex as a critical component of an 'extended' mirror neuron system: does it solve the development, correspondence, and control problems in mirroring? *Behav. Brain Funct.* **4**, 47 (2008).
9. Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).
10. Fan, J. *et al.* Testing the behavioral interaction and integration of attentional networks. *Brain Cogn.* **70**, 209–220 (2009).
11. Green, A. E. *et al.* Using genetic data in cognitive neuroscience: from growing pains to genuine insights. *Nat. Rev. Neurosci.* **9**, 710–720 (2008).
12. Rueda, M. R. *et al.* Development of attentional networks in childhood. *Neuropsychologia* **42**, 1029–1040 (2004).
13. Scerif, G. Attention trajectories, mechanisms and outcomes: at the interface between developing cognition and environment. *Dev. Sci.* **13**, 805–812 (2010).
14. Amso, D. & Johnson, S. P. Learning by selection: visual search and object perception in young infants. *Dev. Psychol.* **42**, 1236–1245 (2006).
15. Amso, D. & Johnson, S. P. Development of visual selection in 3- to 9-month-olds: evidence from saccades to previously ignored locations. *Infancy* **13**, 675–686 (2008).
16. Butcher, P. R., Kalverboer, A. F. & Geuze, R. H. Infants' shifts of gaze from a central to a peripheral stimulus: a longitudinal study of development between 6 and 26 weeks. *Infant Behav. Dev.* **23**, 3–21 (2000).
17. Hood, B. M. Inhibition of return produced by covert shifts of visual-attention in 6-month-old infants. *Infant Behav. Dev.* **16**, 245–254 (1993).
18. Johnson, M. H. in *Attention and Performance XV: Conscious and Nonconscious Information Processing*

(eds Umiltà, C. & Moscovitch, M.) 291–310 (MIT Press, 1994).
19. Johnson, M. H. & Tucker, L. A. The development and temporal dynamics of spatial orienting in infants. *J. Exp. Child Psychol.* **63**, 171–188 (1996).
20. Johnson, M. H., Posner, M. I. & Rothbart, M. K. Components of visual orienting in early infancy — contingency learning, anticipatory looking, and disengaging. *J. Cogn. Neurosci.* **3**, 335–344 (1991).
21. Atkinson, J., Hood, B., Wattambell, J. & Braddick, O. Changes in infants ability to switch visual-attention in the first 3 months of life. *Perception* **21**, 643–653 (1992).
22. Atkinson, J. & Braddick, O. From genes to brain development to phenotypic behavior: "dorsal-stream vulnerability" in relation to spatial cognition, attention, and planning of actions in Williams syndrome (WS) and other developmental disorders. *Prog. Brain Res.* **189**, 261–283 (2011).
23. Corbetta, M. *et al.* A common network of functional areas for attention and eye movements. *Neuron* **21**, 761–773 (1998).
24. Nobre, A. C., Gitelman, D. R., Dias, E. C. & Mesulam, M. M. Covert visual spatial orienting and saccades: overlapping neural systems. *Neuroimage* **11**, 210–216 (2000).
25. Konrad, K. *et al.* Development of attentional networks: an fMRI study with children and adults. *Neuroimage* **28**, 429–439 (2005).
26. Johnson, M. H. The inhibition of automatic saccades in early infancy. *Dev. Psychobiol.* **28**, 281–291 (1995).
27. Guitton, D., Buchtel, H. A. & Douglas, R. M. Frontal-lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Exp. Brain Res.* **58**, 455–472 (1985).
28. Scerif, G. *et al.* To look or not to look? Typical and atypical development of oculomotor control. *J. Cogn. Neurosci.* **17**, 591–604 (2005).
29. Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A. & Sweeney, J. A. Maturation of cognitive processes from late childhood to adulthood. *Child Dev.* **75**, 1357–1372 (2004).
30. Davidson, M. C., Amso, D., Cruess Anderson, L. & Diamond, A. Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* **44**, 2037–2078 (2006).
31. Crone, E. A. Executive functions in adolescence: inferences from brain and behavior. *Dev. Sci.* **12**, 825–830 (2009).
32. Hwang, K., Velanova, K. & Luna, B. Strengthening of top-down frontal cognitive control networks underlying the development of inhibitory control: a functional magnetic resonance imaging effective connectivity study. *J. Neurosci.* **30**, 15535–15545 (2010).
   Describes a study using Granger causality analysis to test developmental changes in effective connectivity underlying inhibitory control (using an antisaccade task) compared with reflexive responses (using a prosaccade task). In early

childhood, few top-down connectivities were evident with increased parietal interconnectivity; however, by adolescence, connections from the PFC were increased and parietal interconnectivity was decreased.
33. Miyake, A. & Friedman, N. P. The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Direct. Psychol. Sci.* **21**, 8–14 (2012).
34. Chun, M. M., Golomb, J. D. & Turk-Browne, N. B. A taxonomy of external and internal attention. *Annu. Rev. Psychol.* **62**, 73–101 (2011).
35. Robertson, I. H., Mattingley, J. B., Rorden, C. & Driver, J. Phasic alerting of neglect patients overcomes their spatial deficit in visual awareness. *Nature* **395**, 169–172 (1998).
36. Gogtay, N. *et al.* Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl Acad. Sci. USA* **101**, 8174–8179 (2004).
37. Sowell, E. R. *et al.* Longitudinal mapping of cortical thickness and brain growth in normal children. *J. Neurosci.* **24**, 8223–8231 (2004).
38. Raznahan, A. *et al.* Longitudinal four-dimensional mapping of subcortical anatomy in human development. *Proc. Natl Acad. Sci. USA* **111**, 1592–1597 (2014).
39. Vandekar, S. N. *et al.* Topologically dissociable patterns of development of the human cerebral cortex. *J. Neurosci.* **35**, 599–609 (2015).
40. Raznahan, A. *et al.* Patterns of coordinated anatomical change in human cortical development: a longitudinal neuroimaging study of maturational coupling. *Neuron* **72**, 873–884 (2011).
41. Schmitt, J. E. *et al.* The dynamic role of genetics on cortical patterning during childhood and adolescence. *Proc. Natl Acad. Sci. USA* **111**, 6774–6779 (2014).
42. Fair, D. A. *et al.* Development of distinct control networks through segregation and integration. *Proc. Natl Acad. Sci. USA* **104**, 13507–13512 (2007).
   Uses resting-state fMRI to show that development of adult control networks involves both segregation (that is, decreased short-range connections) and integration (that is, increased long-range connections) of the brain regions that comprise them.
43. Luna, B. & Sweeney, J. A. in *Adolescent Brain Development: Vulnerabilities and Opportunities* (eds Dahl, R. E. & Spear, L. P.) 296–309 (New York Academy of Sciences, 2004).
44. Maunsell, J. H. R. & Vanessen, D. C. The connections of the middle temporal visual area (mt) and their relationship to a cortical hierarchy in the macaque monkey. *J. Neurosci.* **3**, 2563–2586 (1983).
45. Vanessen, D. C. & Maunsell, J. H. R. Hierarchical organization and functional streams in the visual-cortex. *Trends Neurosci.* **6**, 370–375 (1983).
46. Gilbert, C. D. & Li, W. Top-down influences on visual processing. *Nat. Rev. Neurosci.* **14**, 350–363 (2013).
47. Itti, L. & Koch, C. Computational modelling of visual attention. *Nat. Rev. Neurosci.* **2**, 194–203 (2001).

48. Serre, T., Oliva, A. & Poggio, T. A feedforward architecture accounts for rapid categorization. *Proc. Natl Acad. Sci. USA* **104**, 6424–6429 (2007).

49. Zhang, Y. *et al.* Object decoding with attention in inferior temporal cortex. *Proc. Natl Acad. Sci. USA* **108**, 8850–8855 (2011).

50. Maunsell, J. H. R. & Newsome, W. T. Visual processing in monkey extrastriate cortex. *Annu. Rev. Neurosci.* **10**, 363–401 (1987).

51. Carrasco, M., Ling, S. & Read, S. Attention alters appearance. *Nat. Neurosci.* **7**, 308–313 (2004).

52. Gilbert, C. D. & Sigman, M. Brain states: top-down influences in sensory processing. *Neuron* **54**, 677–696 (2007).

53. Batardiere, A. *et al.* Early specification of the hierarchical organization of visual cortical areas in the macaque monkey. *Cereb. Cortex* **12**, 453–465 (2002).

54. Moore, C. I., Carlen, M., Knoblich, U. & Cardin, J. A. Neocortical interneurons: from diversity, strength. *Cell* **142**, 184–188 (2010).

55. Summerfield, C. & de Lange, F. P. Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* **15**, 745–756 (2014).

56. Amso, D., Haas, S. & Markant, J. An eye tracking investigation of developmental change in bottom-up attention orienting to faces in cluttered natural scenes. *PLoS ONE* **9**, e85701 (2014).

57. Brown, A. M. & Yamamoto, M. Visual-acuity in newborn and preterm infants measured with grating acuity cards. *Am. J. Ophthalmol.* **102**, 245–253 (1986).

58. Atkinson, J., Braddick, O. & Moar, K. Development of contrast sensitivity over first 3 months of life in human infant. *Vision Res.* **17**, 1037–1044 (1977).

59. Banton, T. & Bertenthal, B. I. Multiple developmental pathways for motion processing. *Optom. Vision Sci.* **74**, 751–760 (1997).

60. Braddick, O. J., Wattambell, J. & Atkinson, J. Orientation-specific cortical responses develop in early infancy. *Nature* **320**, 617–619 (1986).

61. Amso, D. & Johnson, S. P. Selection and inhibition in infancy: evidence from the spatial negative priming paradigm. *Cognition* **95**, B27–B36 (2005).

62. Finlay, B. L. & Uchiyama, R. Developmental mechanisms channeling cortical evolution. *Trends Neurosci.* **38**, 69–76 (2015).

63. Badre, D. & D'Esposito, M. Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *J. Cogn. Neurosci.* **19**, 2082–2099 (2007).

64. Sporns, O. The human connectome: origins and challenges. *Neuroimage* **80**, 53–61 (2013).

65. Fransson, P. *et al.* Resting-state networks in the infant brain. *Proc. Natl Acad. Sci. USA* **104**, 15531–15536 (2007).

66. Pruett, J. R. Jr *et al.* Accurate age classification of 6 and 12 month-old infants based on resting-state functional connectivity magnetic resonance imaging data. *Dev. Cogn. Neurosci.* **12**, 123–133 (2015).

67. Fair, D. A. *et al.* Functional brain networks develop from a "local to distributed" organization. *PLoS Comput. Biol.* **5**, e1000381 (2009).

68. Supekar, K., Musen, M. & Menon, V. Development of large-scale functional brain networks in children. *PloS Biol.* **7**, e1000157 (2009).

69. Dosenbach, N. U. F. *et al.* Prediction of individual brain maturity using fMRI. *Science* **329**, 1358–1361 (2010).

70. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* **59**, 2142–2154 (2012).

71. Luck, S. J. & Vogel, E. K. The capacity of visual working memory for features and conjunctions. *Nature* **390**, 279–281 (1997).

72. Sperling, G. The information available in brief visual presentations. *Psychol. Monographs* **74**, 1–29 (1960).

73. Hollingworth, A., Williams, C. C. & Henderson, J. M. To see and remember: visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonom. Bull. Rev.* **8**, 761–768 (2001).

74. Golomb, J. D., Chun, M. M. & Mazer, J. A. The native coordinate system of spatial attention is retinotopic. *J. Neurosci.* **28**, 10654–10662 (2008).

75. Ross-Sheehy, S., Oakes, L. M. & Luck, S. J. Exogenous attention influences visual short-term memory in infants. *Dev. Sci.* **14**, 490–501 (2011).

Shows that infants as young as 5 months of age can encode information in VSTM from multiple-object arrays, and that attention-directing cues influence both perceptual-memory and VSTM encoding of stimuli in infants, as they do in adults.

76. Markant, J. & Amso, D. Selective memories: infants' encoding is enhanced in selection via suppression. *Dev. Sci.* **16**, 926–940 (2013).
This study showed that 9-month-old infants have better recognition memory for category exemplars encoded in the context of an attention-orienting mechanism involving suppression of distractor information in contrast with a condition in which such suppression is not engaged.

77. Shimi, A., Nobre, A. C., Astle, D. & Scerif, G. Orienting attention within visual short-term memory: development and mechanisms. *Child Dev.* **85**, 578–592 (2014).

78. Markant, J. & Amso, D. Leveling the playing field: attention mitigates the effects of intelligence on memory. *Cognition* **131**, 195–204 (2014).

79. Dixon, M. L., Zelazo, P. D. & De Rosa, E. Evidence for intact memory-guided attention in school-aged children. *Dev. Sci.* **13**, 161–169 (2010).

80. Hollingworth, A., Richard, A. M. & Luck, S. J. Understanding the function of visual short-term memory: transsaccadic memory, object correspondence, and gaze correction. *J. Exp. Psychol. General* **137**, 163–181 (2008).

81. Ross-Sheehy, S., Oakes, L. M. & Luck, S. J. The development of visual short-term memory capacity in infants. *Child Dev.* **74**, 1807–1822 (2003).

82. Kaldy, Z. & Leslie, A. M. A memory span of one? Object identification in 6.5-month-old infants. *Cognition* **97**, 153–177 (2005).

83. Wu, R. & Kirkham, N. Z. No two cues are alike: depth of learning during infancy is dependent on what orients attention. *J. Exp. Child Psychol.* **107**, 118–136 (2010).
Demonstrates that attention-directing social cues have powerful influences on young infants' ability to learn about features of their visual world.

84. Richards, J. E. & Casey, B. J. Heart-rate-variability during attention phases in young infants. *Psychophysiology* **28**, 43–53 (1991).

85. Markant, J., Worden, M. S. & Amso, D. Not all attention orienting is created equal: recognition memory is enhanced when attention orienting involves distractor suppression. *Neurobiol. Learn. Memory* **120**, 28–40 (2015).

86. Crone, E. A., Wendelken, C., Donohue, S., van Leijenhorst, L. & Bunge, S. A. Neurocognitive development of the ability to manipulate information in working memory. *Proc. Natl Acad. Sci. USA* **103**, 9315–9320 (2006).

87. Wendelken, C., Baym, C. L., Gazzaley, A. & Bunge, S. A. Neural indices of improved attentional modulation over middle childhood. *Dev. Cogn. Neurosci.* **1**, 175–186 (2011).
Shows that children's reduced ability to maintain items in working memory, especially in the presence of distraction, is driven by weaker top-down modulation of activity in areas involved in stimulus processing.

88. Olesen, P. J., Macoveanu, J., Tegner, J. & Klingberg, T. Brain activity related to working memory and distraction in children and adults. *Cereb. Cortex* **17**, 1047–1054 (2007).

89. Astle, D. E., Nobre, A. C. & Scerif, G. Attentional control constrains visual short-term memory: insights from developmental and individual differences. *Q. J. Exp. Psychol. (Hove)* **65**, 277–294 (2012).

90. Astle, D. *et al.* The neural dynamics of fronto-parietal networks in childhood revealed using magnetoencephalography. *Cereb. Cortex* http://dx.doi.org/10.1093/cercor/bhu271 (2014).
Uses magnetoencephalography to show that, in children, slow frequency-theta (4–7 Hz) activity within a right-lateralized frontoparietal network in anticipation of memoranda being encoded into VSTM predicts the accuracy with which those memory items were subsequently retrieved, as well as activity associated with early visual processing of the memoranda.

91. Shimi, A., Kuo, B.-C., Astle, D. E., Nobre, A. C. & Scerif, G. Age group and individual differences in attentional orienting dissociate neural mechanisms of encoding and maintenance in visual STM. *J. Cogn. Neurosci.* **26**, 864–877 (2014).

Uses electroencephalography to show that adults, but not children, elicit a set of neural markers that are broadly similar in preparation for encoding and during maintenance in VSTM.

92. Best, J. R., Miller, P. H. & Naglieri, J. A. Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learn. Individ. Differ.* **21**, 327–336 (2011).

93. Bull, R. & Scerif, G. Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory. *Dev. Neuropsychol.* **19**, 273–293 (2001).

94. Chun, M. M. & Jian, Y. H. Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cogn. Psychol.* **36**, 28–71 (1998).

95. Summerfield, J. J., Lepsien, J., Gitelman, D. R., Mesulam, M. M. & Nobre, A. C. Orienting attention based on long-term memory experience. *Neuron* **49**, 905–916 (2006).

96. Wu, R. *et al.* Searching for something familiar or novel: top-down attentional selection of specific items or object categories. *J. Cogn. Neurosci.* **25**, 719–729 (2013).

97. Chun, M. M. & Jiang, Y. H. Top-down attentional guidance based on implicit learning of visual covariation. *Psychol. Sci.* **10**, 360–365 (1999).

98. Shimi, A. & Scerif, G. The interplay of spatial attentional biases and mnemonic codes in VSTM: developmentally informed hypotheses. *Dev. Psychol.* **51**, 731–743 (2015).

99. Casey, B. J. *et al.* Implication of right frontostriatal circuitry in response inhibition and attention-deficit/hyperactivity disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **36**, 374–383 (1997).

100. Durston, S. *et al.* Differential effects of DRD4 and DAT1 genotype on fronto-striatal gray matter volumes in a sample of subjects with attention deficit hyperactivity disorder, their unaffected siblings, and controls. *Mol. Psychiatry* **10**, 678–685 (2005).

101. Volkow, N. D., Wang, G. J., Fowler, J. S. & Ding, Y. S. Imaging the effect of methylphenidate on brain dopamine: new model on its therapeutic actions for attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **57**, 1410–1415 (2005).

102. Castellanos, F. X. *et al.* Quantitative brain magnetic resonance imaging in attention-deficit hyperactivity disorder. *Arch. Gen. Psychiatry* **53**, 607–616 (1996).

103. Castellanos, F. X., Sonuga-Barke, E. J. S., Milham, M. P. & Tannock, R. Characterizing cognition in ADHD: beyond executive dysfunction. *Trends Cogn. Sci.* **10**, 117–123 (2006).

104. Castellanos, F. X. & Proal, E. Large-scale brain systems in ADHD: beyond the prefrontal-striatal model. *Trends Cogn. Sci.* **16**, 17–26 (2012).

105. Neale, B. M. *et al.* Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **49**, 884–897 (2010).

106. Shaw, P. *et al.* Longitudinal mapping of cortical thickness and clinical outcome in children and adolescents with attention-deficit/hyperactivity disorder. *Arch. Gen Psychiatry* **63**, 540–549 (2006).

107. Batty, M. J. *et al.* Cortical gray matter in attention-deficit/hyperactivity disorder: a structural magnetic resonance imaging study. *J. Am. Acad. Child Adolesc. Psychiatry* **49**, 229–238 (2010).

108. Poldrack, R. A. Is "efficiency" a useful concept in cognitive neuroscience? *Dev. Cogn. Neurosci.* **11**, 12–17 (2015).

109. Wang, L. *et al.* Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder. *Hum. Brain Mapp.* **30**, 638–649 (2009).

110. Cortese, S. *et al.* Toward systems neuroscience of ADHD: a meta-analysis of 55 fMRI Studies. *Am. J. Psychiatry* **169**, 1038–1055 (2012).

111. Liddle, E. B. *et al.* Task-related default mode network modulation and inhibitory control in ADHD: effects of motivation and methylphenidate. *J. Child Psychol. Psychiatry* **52**, 761–771 (2011).

112. Fair, D. A. *et al.* Atypical default network connectivity in youth with attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **68**, 1084–1091 (2010).

113. Hahamy, A., Behrmann, M. & Malach, R. The idiosyncratic brain: distortion of spontaneous connectivity patterns in autism spectrum disorder. *Nat. Neurosci.* **18**, 302–309 (2015).

114. Ray, S. *et al.* Structural and functional connectivity of the human brain in autism spectrum disorders and attention-deficit/hyperactivity disorder: a rich club-organization study. *Hum. Brain Mapp.* **35**, 6032–6048 (2014).

REVIEWS

115. Di Martino, A. *et al.* Unraveling the miswired connectome: a developmental perspective. *Neuron* **83**, 1335–1353 (2014).
This review highlights the excitement and caveats associated with the new field of developmental connectomics and their implications for understanding neurodevelopmental disorders.

116. Braddick, O., Atkinson, J. & Wattam-Bell, J. Normal and anomalous development of visual motion processing: motion coherence and 'dorsal-stream vulnerability'. *Neuropsychologia* **41**, 1769–1784 (2003).
Summarizes the wealth of evidence for dorsal stream vulnerability across a large number of neurodevelopmental disorders.

117. Mottron, L., Dawson, M., Soulieres, I., Hubert, B. & Burack, J. Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. *J. Autism Dev. Disord.* **36**, 27–43 (2006).

118. O'Riordan, M. & Plaisted, K. Enhanced discrimination in autism. *Q. J. Exp. Psychol. A* **54**, 961–979 (2001).

119. Davis, G. & Plaisted-Grant, K. Low endogenous neural noise in autism. *Autism* **19**, 351–362 (2015).

120. Pellicano, E. & Burr, D. When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends Cogn. Sci.* **16**, 504–510 (2012).

121. Shen, M. D. *et al.* Early brain enlargement and elevated extra-axial fluid in infants who develop autism spectrum disorder. *Brain* **136**, 2825–2835 (2013).

122. Nordahl, C. W. *et al.* Brain enlargement is associated with regression in preschool-age boys with autism spectrum disorders. *Proc. Natl Acad. Sci. USA* **108**, 20195–20200 (2011).

123. Keehn, B., Mueller, R.-A. & Townsend, J. Atypical attentional networks and the emergence of autism. *Neurosci. Biobehav. Rev.* **37**, 164–183 (2013).

124. Jones, W. & Klin, A. Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature* **504**, 427–431 (2013).

125. Elsabbagh, M. *et al.* Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. *Biol. Psychiatry* **74**, 189–194 (2013).

126. Posner, M. I., Rothbart, M. K. & Sheese, B. E. Attention genes. *Dev. Sci.* **10**, 24–29 (2007).

127. Fossella, J. *et al.* Assessing the molecular genetics of attention networks. *BMC Neurosci.* **3**, 14 (2002).

128. Brookes, K. *et al.* The analysis of 51 genes in DSM-IV combined type attention deficit hyperactivity disorder: association signals in *DRD4*, *DAT1* and 16 other genes. *Mol. Psychiatry* **11**, 934–953 (2006).

129. Faraone, S. V., Doyle, A. E., Mick, E. & Biederman, J. Meta-analysis of the association between the 7-repeat allele of the dopamine D₄ receptor gene and attention deficit hyperactivity disorder. *Am. J. Psychiatry* **158**, 1052–1057 (2001).

130. Dumontheil, I. *et al.* Influence of the *COMT* genotype on working memory and brain activity changes during development. *Biol. Psychiatry* **70**, 222–229 (2011).

131. Franke, B. *et al.* Multicenter analysis of the *SLC6A3/DAT1* VNTR haplotype in persistent ADHD suggests differential involvement of the gene in childhood and persistent ADHD. *Neuropsychopharmacology* **35**, 656–664 (2010).

132. Stergiakouli, E. *et al.* Investigating the contribution of common genetic variants to the risk and pathogenesis of ADHD. *Am. J. Psychiatry* **169**, 186–194 (2012).

133. Yang, L. *et al.* Polygenic transmission and complex neuro developmental network for attention deficit hyperactivity disorder: genome-wide association study of both common and rare variants. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **162B**, 419–430 (2013).

134. Hamshere, M. L. *et al.* Shared polygenic contribution between childhood attention-deficit hyperactivity disorder and adult schizophrenia. *Br. J. Psychiatry* **203**, 107–111 (2013).

135. Martin, J., Hamshere, M. L., Stergiakouli, E., O'Donovan, M. C. & Thapar, A. Genetic risk for attention-deficit/hyperactivity disorder contributes to neurodevelopmental traits in the general population. *Biol. Psychiatry* **76**, 664–671 (2014).

136. Poelmans, G., Pauls, D. L., Buitelaar, J. K. & Franke, B. Integrated genome-wide association study findings: identification of a neurodevelopmental network for attention deficit hyperactivity disorder. *Am. J. Psychiatry* **168**, 365–377 (2011).
One of the first studies to discuss the failures of genome-wide studies of ADHD risk to identify single-candidate variants and argue that such studies should be replaced with analyses that instead focus on neurodevelopmental pathways.

137. Scerif, G. & Karmiloff-Smith, A. The dawn of cognitive genetics? Crucial developmental caveats. *Trends Cogn. Sci.* **9**, 126–135 (2005).

138. Rice, F. *et al.* The links between prenatal stress and offspring development and psychopathology: disentangling environmental and inherited influences. *Psychol. Med.* **40**, 335–345 (2010).

139. Hall, J., Trent, S., Thomas, K. L., O'Donovan, M. C. & Owen, M. J. Genetic risk for schizophrenia: convergence on synaptic pathways involved in plasticity. *Biol. Psychiatry* **77**, 52–58 (2015).

140. Pinto, D. *et al.* Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am. J. Hum. Genet.* **94**, 677–694 (2014).

141. Lee, S. H. *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **45**, 984–994 (2013).

142. Wahlstrom, D., Collins, P., White, T. & Luciana, M. Developmental changes in dopamine neurotransmission in adolescence: Behavioral implications and issues in assessment. *Brain Cogn.* **72**, 146–159 (2010).

143. Scerif, G. & Baker, K. Annual Research Review: rare genotypes and childhood psychopathology — uncovering diverse developmental mechanisms of ADHD risk. *J. Child Psychol. Psychiatry* **56**, 251–273 (2015).

144. Verkerk, A. *et al.* Identification of a gene (*FMR-1*) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile-X syndrome. *Cell* **65**, 905–914 (1991).

145. Bear, M. F., Huber, K. M. & Warren, S. T. The mGluR theory of fragile X mental retardation. *Trends Neurosci.* **27**, 370–377 (2004).

146. D'Hulst, C. *et al.* Decreased expression of the GABA_A receptor in fragile X syndrome. *Brain Res.* **1121**, 238–245 (2006).

147. Bagni, C. & Greenough, W. T. From mRNP trafficking to spine dysmorphogenesis: The roots of fragile X syndrome. *Nat. Rev. Neurosci.* **6**, 376–387 (2005).

148. Huber, K. M., Gallagher, S. M., Warren, S. T. & Bear, M. F. Altered synaptic plasticity in a mouse model of fragile X mental retardation. *Proc. Natl Acad. Sci. USA* **99**, 7746–7750 (2002).

149. Zhang, Y. Q. *et al.* Protein expression profiling of the *Drosophila* fragile X mutant brain reveals up-regulation of monoamine synthesis. *Mol. Cell. Proteom.* **4**, 278–290 (2005).

150. Elston, G. N., Oga, T., Okamoto, T. & Fujita, I. Spinogenesis and pruning from early visual onset to adulthood: an intracellular injection study of layer III pyramidal cells in the ventral visual cortical pathway of the macaque monkey. *Cereb. Cortex* **20**, 1398–1408 (2010).

151. Elia, J. *et al.* Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Mol. Psychiatry* **15**, 637–646 (2010).

152. Darnell, J. C. *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247–261 (2011).

153. Auerbach, B. D., Osterweil, E. K. & Bear, M. F. Mutations causing syndromic autism define an axis of synaptic pathophysiology. *Nature* **480**, 63–68 (2011).

154. Baker, K., Scerif, G., Astle, D. E., Fletcher, P. C. & Raymond, F. L. Psychopathology and cognitive performance in individuals with membrane-associated guanylate kinase mutations: a functional network phenotyping study. *J. Neurodev. Disord.* **7**, 8 (2015).

155. Bavelier, D. & Neville, H. J. Cross-modal plasticity: where and how? *Nat. Rev. Neurosci.* **3**, 443–452 (2002).

156. Bavelier, D. *et al.* Visual attention to the periphery is enhanced in congenitally deaf individuals. *J. Neurosci.* **20**, RC93 (2000).

157. Mezzacappa, E. Alerting, orienting, and executive attention: developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. *Child Dev.* **75**, 1373–1386 (2004).

158. Hackman, D. A., Farah, M. J. & Meaney, M. J. Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nat. Rev. Neurosci.* **11**, 651–659 (2010).

159. Amso, D., Markant, J. & Haas, S. Agents of developmental change in orienting to faces in cluttered natural scenes. *PLoS ONE* **9**, e85701 (2014).

160. Bejjanki, V. R. *et al.* Action video game play facilitates the development of better perceptual templates. *Proc. Natl Acad. Sci. USA* **111**, 16961–16966 (2014).

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

162. Green, C. S. & Bavelier, D. Effect of action video games on the spatial distribution of visuospatial attention. *J. Exp. Psychol. Hum. Percept. Perform.* **32**, 1465–1478 (2006).

163. Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L. & Posner, M. I. Training, maturation, and genetic influences on the development of executive attention. *Proc. Natl Acad. Sci. USA* **102**, 14931–14936 (2005).

164. Klingberg, T. Training and plasticity of working memory. *Trends Cogn. Sci.* **14**, 317–324 (2010).

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

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

167. Klingberg, T. *et al.* Computerized training of working memory in children with ADHD — a randomized, controlled trial. *J. Am. Acad. Child Adolesc. Psychiatry* **44**, 177–186 (2005).

168. Chacko, A. *et al.* A randomized clinical trial of Cogmed Working Memory Training in school-age children with ADHD: a replication in a diverse sample using a control condition. *J. Child Psychol. Psychiatry* **55**, 247–255 (2014).

169. Shipstead, Z., Redick, T. S. & Engle, R. W. Is working memory training effective? *Psychol. Bull.* **138**, 628–654 (2012).

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

171. Wass, S. V., Scerif, G. & Johnson, M. H. Training attentional control and working memory — is younger, better? *Dev. Rev.* **32**, 360–387 (2012).

172. Anguera, J. A. *et al.* Video game training enhances cognitive control in older adults. *Nature* **501**, 97–101 (2013).

173. Astle, D. E. *et al.* Cognitive training enhances intrinsic brain connectivity in childhood. *J. Neurosci.* **35**, 6277–6283 (2015).

174. Schiller, P. H. in *Models of the Visual Cortex* (eds Rose, D. & Dobson, V. G.) 62–70 (Wiley, 1985).

175. Johnson, M. H. Cortical maturation and the development of visual attention in early infancy. *J. Cogn. Neurosci.* **2**, 81–95 (1990).

176. Canfield, R. L. & Haith, M. M. Young infants visual expectations for symmetrical and asymmetric stimulus sequences. *Dev. Psychol.* **27**, 198–208 (1991).

177. Haith, M. M. & McCarty, M. E. Stability of visual expectations at 3.0 months of age. *Dev. Psychol.* **26**, 68–74 (1990).

178. Johnson, S. P., Amso, D. & Slemmer, J. A. Development of object concepts in infancy: evidence for early learning in an eye-tracking paradigm. *Proc. Natl Acad. Sci. USA* **100**, 10568–10573 (2003).

179. Richards, J. E. & Holley, F. B. Infant attention and the development of smooth pursuit tracking. *Dev. Psychol.* **35**, 856–867 (1999).

180. Luna, B., Velanova, K. & Geier, C. F. Development of eye-movement control. *Brain Cogn.* **68**, 293–308 (2008).

181. Csibra, G., Tucker, L. A. & Johnson, M. H. Neural correlates of saccade planning in infants: a high-density ERP study. *Int. J. Psychophysiol.* **29**, 201–215 (1998).

182. Luna, B. *et al.* Maturation of widely distributed brain function subserves cognitive development. *Neuroimage* **13**, 786–793 (2001).

183. Ordaz, S. J., Foran, W., Velanova, K. & Luna, B. Longitudinal growth curves of brain function underlying inhibitory control through adolescence. *J. Neurosci.* **33**, 18109–18124 (2013).

184. Bull, R., Espy, K. A. & Wiebe, S. A. Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Dev. Neuropsychol.* **33**, 205–228 (2008).

185. Steele, A., Karmiloff-Smith, A., Cornish, K. & Scerif, G. The multiple subfunctions of attention: differential developmental gateways to literacy and numeracy. *Child Dev.* **83**, 2028–2041 (2012).

186. Barnes, J. J. M., Woolrich, M. W., Baker, K., Colclough, G. L. & Astle, D. E. Electrophysiological measures of resting state functional connectivity and their relationship with working memory capacity in childhood. *Dev. Sci.* http://dx.doi.org/10.1111/desc.12297 (2015).

187. Hagerman, P. J. The fragile X prevalence paradox. *J. Med. Genet.* **45**, 498–499 (2008).
188. Hoeft, F. *et al.* Fronto-striatal dysfunction and potential compensatory mechanisms in male adolescents with fragile X syndrome. *Hum. Brain Mapp.* **28**, 543–554 (2007).
189. Kwon, H. *et al.* Functional neuroanatomy of visuospatial working memory in fragile X syndrome: relation to behavioral and molecular measures. *Am. J. Psychiatry* **158**, 1040–1051 (2001).
190. Hoeft, F. *et al.* Region-specific alterations in brain development in one- to three-year-old boys with fragile X syndrome. *Proc. Natl Acad. Sci. USA* **107**, 9335–9339 (2010).
191. La Fata, G. *et al.* FMRP regulates multipolar to bipolar transition affecting neuronal migration and cortical circuitry. *Nat. Neurosci.* **17**, 1693–1700 (2014).
192. Goncalves, J. T., Anstey, J. E., Golshani, P. & Portera-Cailliau, C. Circuit level defects in the developing neocortex of Fragile X mice. *Nat. Neurosci.* **16**, 903–909 (2013).
193. Scerif, G., Cornish, K., Wilding, J., Driver, J. & Karmiloff-Smith, A. Visual search in typically developing toddlers and toddlers with Fragile X or Williams syndrome. *Dev. Sci.* **7**, 116–130 (2004).
194. Scerif, G., Longhi, E., Cole, V., Karmiloff-Smith, A. & Cornish, K. Attention across modalities as a longitudinal predictor of early outcomes: the case of fragile X syndrome. *J. Child Psychol. Psychiatry* **53**, 641–650 (2012).
195. Cornish, K., Cole, V., Longhi, E., Karmiloff-Smith, A. & Scerif, G. Mapping developmental trajectories of attention and working memory in fragile X syndrome: developmental freeze or developmental change? *Dev. Psychopathol.* **25**, 365–376 (2013).
196. Farzin, F., Rivera, S. M. & Whitney, D. Resolution of spatial and temporal visual attention in infants with fragile X syndrome. *Brain* **134**, 3355–3368 (2011).
197. Gilbert, C. B. in *Principles of Neural Science* 5th edn (eds Kandel, E. R., Schwartz, J., Jessel, T., Siegelbaum, S. A. & Hudspeth, A. J.) 623 (McGraw-Hill Education, 2013)