How the development of executive function influences our moment-by-moment interactions with the real-world environment

Wass, Samuel V

Correspondence address: Dr Sam Wass, University of East London, London E15 4LZ.
Telephone: +44(0)7725369189. Email: s.v.wass@uel.ac.uk

Acknowledgements: This research was funded by ESRC grant number ES/N017560/1, by ERC grant number ONACSA 853251 and by a Project Grant RPG-2018-281 from the Leverhulme Trust. Thanks to Marta Perapoch and Tom Northrop for reading the manuscript and giving comments, and to members of the UEL BabyDev Lab for useful discussions.
Historically, the study of executive function (EF) development has relied on using experimental paradigms to assess EFs as abstract, time-invariant properties of individual brains. Here, we discuss new research that moves away from studying EFs purely as internal mental constructs, towards an approach that aims to understand how EFs are expressed through the inter-relationship between an individual’s brain and the world around them. We offer three illustrative examples of this approach. The first looks at how we learn to make predictions and anticipations based on different types of regularity in our early social and physical environment. The second looks at how we learn to correct, moment-by-moment, for changes in the outside world to maintain stability in the face of change. The third looks at how we allocate our attention on a moment-by-moment basis, in naturalistic settings. We discuss potential new therapeutic avenues for improving EFs arising from this research.

Keywords: Executive function, volitional control, self regulation, naturalistic, development
“There is more pleasure to building castles in the air than on the ground” Edward Gibbon

Driven by calls for a more empirical science of the mind and brain that first started to emerge in the 18th century, psychologists and (more recently) cognitive neuroscientists often present standardised simulacra intended to mimic real-world cognitive operations 1,2. Although they have been well discussed 3-8, it is always worth remembering what some of the inherent assumptions underlying this approach are. First, the experimental simulacra often differ in a number of ways from the real-world cognitive operations that they are intended to mimic 7. For example, changes in behaviour, physiology and neural activity are often measured relative to the sudden appearance and disappearance of objects, and relative to the repeated presentation of exact sequences of events that reoccur – both of which rarely if ever occur in the real world. Second, the claim that the experimental simulacra do actually mimic the real-world cognitive operation that they were designed to imitate is rarely, if ever, tested 8. Third, these approaches rely on the assumption that cognitive operations can be abstracted – i.e. that, for example, cognitive control measured at one time and using one paradigm relates meaningfully to cognitive control measured using a different type of paradigm 9-13. In fact, research suggests that both individual differences 13,14 and transfer effects following cognitive training 15 are remarkably specific to minor details of the experimental paradigm used.

One further assumption underlying these approaches arises because both the events themselves, and their exact timings, are decided by the experimenter, and not the participant. In the real world, though, behaviour does not happen just through passive, serial-order responses to external stimuli 16-22. As Dewey first noted over hundred years ago “[w]hat we have is a circuit, not an
Running head: EF AND THE ENVIRONMENT

arc or broken segment of a circle. [ . . ] The motor response determines the stimulus, just as truly as sensory stimulus determines movement.” (Dewey, 1896, p. 365).

Reflecting these concerns, theoretical developments in other fields have seen a shift away from approaches that locate learning in individual brains, via internal representation and computation, towards approaches that emphasise how intelligence emerges through the interaction of an agent with their environment 20,23,24. More recently in social cognition there has been a shift away from studying how our brains process a one-way flow of information – e.g. from speaker to listener - towards approaches that examine how social interaction changes the inter-relationship between two brains through bidirectional exchanges 25–29.

Because EFs, defined as the ‘active, volitional component of cognition’ 30, are located within the individual by definition, the majority of research into EF, and how EFs develop, has concentrated on measuring abstract, internal mental constructs 31–35. Within cognitive neuroscience, research has concentrated almost exclusively on understanding how EFs are instantiated through changing patterns of communication and connectivity within an individual, between behaviour, physiology and neural activity 36–40.

Here, we review recent research that has taken a different approach. Rather than studying EFs as abstract, time-invariant mental constructs, it explores instead how EFs are expressed through the inter-relationship between an individual and the world around them 41,42. In other words, they examine how EFs – the active, volitional component of cognition - guide our moment-by-moment interactions with the world outside us.
We explore this idea in three sections. The first looks at how EFs develop through our interactions with our early environment – looking at how our brains learn to exploit rhythms in the outside world to learn to make predictions and anticipations. The second looks at how we use behaviours to correct, moment-by-moment, for changes in the outside world in order to maintain stability in the face of change. The third looks at how we decide how to allocate our attention on a moment-by-moment basis.

We concentrate on understanding how these develop through the first few years of life. This is because, although some disagree\textsuperscript{43,44}, it is still widely thought that the capacity for active, volitional control is largely absent at birth and develops through the early years\textsuperscript{45–49} – a conclusion drawn both from behavioural\textsuperscript{45,46} and neuroimaging\textsuperscript{36} studies. Researching this period affords us, therefore, the opportunity to observe how our moment-by-moment interactions with our environment change during this transition.

**Example 1: learning to make predictions and anticipations based on our early social and physical environment**

One metaphor sometimes applied to the child, learning to use executive control to influence how they respond and act in their environment, is that of the ‘ghost’ (the volitional component of cognition) learning to ‘drive the machine’\textsuperscript{50,51}. But how exactly does this process work? When they (the ghost) take charge of the machine, are they taking charge of an inanimate hunk of metal, that is lying silent on the tarmac until they step into it, turn the ignition, and start to navigate around their environment?
In fact, we know now that that is definitely not the case. The environments that we live in from birth are a dense, complex mishmash of information $^{23,52}$. Our brains $^{53-56}$ and our bodies $^{57}$ respond to changes in our environment even when we are not effortfully directing attention to them. So the machine (our brain) is already moving - reflexively bumping along, driven passively by information in our environment – by the time when the ghost first starts trying to take charge.

So if the machine is already moving when I start to try to take control, then how do I learn to steer it? Some of the ways in which we do this are, perhaps, similar to what we would actually do if this metaphor were instantiated as an actual, real-world scenario. (Terrifying as that might seem.)

First, I might notice that, every time that one thing happens (e.g. the steering wheel is nudged to the left, by chance), something else happens (e.g. the machine swerves left). This would allow me to gain self-awareness (‘I caused that’) and to make predictions (‘if I do that again, the same thing will happen’) – both of which are precursors to actually being able to steer. In the first part we discuss this mechanism: how contingent responsiveness in a child’s early interactions with their environment might influence the development of EF.

Second, I might notice - if, for example, I were driving over sand dunes, or skiing moguls - that certain temporal regularities exist in how the machine bumps along. Over time, by learning to predict when the peaks and troughs will occur, and nudging on the pedals or turning the steering wheel slightly, we can help to gain control. In the subsequent part we discuss how temporal regularities might influence the development of EFs.
Much early behaviour is partially stochastic during the first months of life, including eye gaze movements, affective states, vocal production and physical movements. However, many behaviours, even unintended ones, have consequences: for example, banging a table reliably makes a noise, and looking at someone reliably leads them to return their gaze. Although all physical environments are equally contingent (all tables make a noise when banged), some social environments are more contingent than others (because some social partners are more responsive).

Social contingencies, when a given behaviour reliably elicits a given response from an adult caregiver, have both an immediate and long-term effect on infants’ attention. Responding contingently to an infant’s gestures immediately improves the quality and quantity of the attention that they pay to objects; when caregivers behave redirectively (i.e. non-contingently), infants’ visual attention durations immediately decrease. There is also evidence that parental contingent social responsiveness has a long-term effect on EF development. For example, child→parent influences in facial affect during tabletop play at 3 and 9 months predicted child self-control at 2 years after temperament, IQ, and maternal style were partialled.

Research into action-oriented predictive processing suggests that motor intentions actively elicit active predictions about the ongoing results of the consequences of our actions. Because of this, anticipated events evoke greater neural responsiveness than unanticipated ones. Less well understood, however, is how action predictions mediate neural responsiveness across an interacting dyad. We know that mutual prediction drives brain activity during a social...
interaction 76,78–82; but we understand little about how the contingency of a partner’s responses to my actions influence my own neural responses to those actions. Intuitively, though, it seems plausible that similar mechanisms might operate across a dyad: i.e., that when I initiate, and my partner responds contingently, my own responses to that action are greater than if my partner performs an equivalent action non-contingently.

Two recent studies have reported results that may be consistent with this. One used EEG to record infant neural responses to an adult either turning to look at the same object as the infant (congruent actor) or turning to look at a different object (incongruent actor). Neural responses (alpha desynchronisation) was greater for the congruent actor 83 (see also 84). The second study recorded dual EEG during naturalistic child-parent play, and found that, when a parent responded contingently to an infant-led attention shift towards a new play object, the infant showed increased alpha desynchronisation, along with increased attention to that object 85. These findings may be attributable to infants predicting their social partner’s responses, moment-by-moment, and showing greater neural responsivity to anticipated responses as a consequence.
Figure 1: a) example of a tabletop play setting in which dual EEG is recorded. b) example of an interactive contingency. Time 0 shows moments during tabletop play when the infant looks to their mother. Three coloured lines show how the mother’s gaze behaviour changes relative to these moments. c) Topoplots showing how the child’s brain activity changes relative to mutual attention episodes initiated by the child (top) and by the mother (bottom). Greater alpha desynchronisation after the look onset is observed during child-led mutual attention episodes. From Phillips et al., 2021.

Temporal regularities

In addition to being contingent on my own behaviours, another way for our environment to be predictable is for it to have its own intrinsic patterns, or temporal regularities. Research suggests that, when we are exposed to environments with intrinsic patterns, periodic activity patterns
within an individual become coupled with periodic activity patterns in the environment – which is known as oscillatory entrainment. It is thought oscillatory entrainment may be a mechanism that facilitates sensory processing – for example, by helping to ensure that sensory information arrives at times of peak neuronal excitability and sensitivity.

Top-down modulation of oscillatory entrainment has been shown in adults using versions of the ‘cocktail party effect’ - in which multiple sensory streams are presented and the participant is asked to direct attention to one of them. In adults, greater low-frequency neural entrainment is observed to the attended-to stream (e.g. ). This is thought to be driven by top-down modulation of activity in primary sensory cortices (e.g. ). The mechanisms which drive top-down modulation to periodic stimuli may be quite similar to those involved in generating predictions for aperiodic stimuli, too.

Research into how, and when, oscillatory entrainment develops during early life is still scant. Based on other evidence, however, it seems probably that oscillatory entrainment driven by top-down modulation develops relatively slowly, through infant and childhood. Importantly, though, evidence suggests that some of the facilitatory effects of oscillatory entrainment on sensory processing may still be present even when no top-down modulation of oscillatory entrainment is taking place at all - in cases where the stimulus is regularly spaced (see Figure 2): we know that contingent, evoked responses to a stimulus occur even in the absence of prediction or anticipation (e.g. for unknown languages); and we know that these bottom-up contingent evoked responses affect the phase of brain activity, which in turn affects sensory processing. Therefore, when the stimulus is regularly spaced, bottom-up contingent evoked responses may also facilitate sensory processing, even in the absence of top-down modulation.
Figure 2: Illustrating three different mechanisms through which information presented might be optimised to arrive at times of peak neuronal excitability. a) a classical 'top-down' entrainment model: top-down modulation changes the phase of the underlying neuronal activity to align the sensory input with the stimulus.
neuronal oscillatory activity with the attended-to sensory stream; b) top-down modulation in anticipation to aperiodic stimuli; c) an illustration of how some of the facilitatory effects of oscillatory entrainment on sensory processing may still be present for periodic stimuli, even when no top-down modulation of oscillatory entrainment is taking place. Phase changes evoked by the previous stimulus might continue to ‘reverberate’ via a damped response, thus ensuring that the subsequent stimulus also arrives at an optimal phase for encoding.

It is possible, then, that early-life exposure to regularly spaced, periodic early-life environments may have two effects. First, in the short term, it may immediately facilitate sensory processing in the absence of top-down predictions and anticipations. In other words, the predictability of the environment may have a facilitatory effect on processing that is comparable to the effect of top-down control, even in the absence of top-down control.

Second, in the long term, periodic early-life environments may facilitate anticipations and predictions, leading to better top-down modulated entrainment, manifesting behaviourally as better executive control. There is ample evidence that temporal regularities are important during development – for example, child-directed speech is more periodic (97,98 (see also 99–101)), and that aperiodic environments may associate with worse EF outcomes. However, direct associations between periodicities in the early-life environment and short- and long-term facilitatory effects on neural processing of environmental information remain untested.

Over time, then, it is possible that these two different types of predictability – interactive contingencies (‘every time X happens, Y happens afterwards’) and temporal regularities - might help the ‘ghost’ to learn how to steer the machine. Regularities in the environment might directly
facilitate the neural processing of information, mimicking the facilitatory effects of top-down modulation on sensory perception even before top-down modulation develops.

Example 2: learning to correct for changes in the outside world to maintain stability in the face of change

As the capacity for executive control starts to develop, how then do we use that ability to influence how we interact with the environment from one moment to the next? In the next two sections we discuss two approaches to answering that question. The first looks at how we correct for changes in the outside world to maintain stability in the face of change. And the second looks at how we use executive control to decide how we allocate our attention in complex, dynamic naturalistic settings.

Our environment is in constant, dynamic fluctuation. Our stress response (originally called general adaptation syndrome 106), is the dynamical system through which we adaptively respond to external change in order to maintain internal constancy. Although our stress systems are multifarious 107,108, here we concentrate on the autonomic nervous system and the network of brain regions and neurotransmitters involved in controlling arousal and regulatory function 109–113. Optimal ANS arousal (henceforth arousal) lies at an intermediate point between over- and under-arousal 49,114–116.

Allostasis describes the active, dynamical process through which internal equilibrium (homeostasis) is achieved and maintained (Cannon, 1929; McEwen & Wingfield, 2003; Ramsay & Woods, 2014; Sterling, 2012 Selye, 1951). For example, when something occurs that
exogenously increases arousal, allostatic mechanisms would generally involve behaviours that seek out lower stimulation, thereby decreasing arousal. And when something occurs to decrease arousal, allostatic mechanisms would involve behaviours that seek out higher stimulation, thereby increasing arousal. This process is known as self-regulation, which is “the ongoing, dynamic, and adaptive modulation of internal state (emotion, cognition) or behaviour, mediated by central and peripheral physiology” \(^{117}\). Because different children have different levels of ‘optimal’ arousal (cf \(^{118}\)), it is possible that a given arousal level might elicit down-regulation in one child (because that arousal level if above ‘optimum’ level for that child) but up-regulation in another child (because that arousal level is above their ‘optimum’ level) \(^{49,119}\).

Recent research has used dynamical modelling to examine the continuous relationship between children’s overt displays of emotions (facial and vocal affect) and their use of executive processes (e.g. thumb-sucking as self-soothing) during a frustration-eliciting task \(^{120–122}\). Results showed coupling between the two variables, such that executive processes had a direct influence on changes in emotional displays at all ages. The strength of this coupling was stable between 24 months and 5 years. When examining coupling in the opposite direction – how emotional displays affect executive processes – they also found that emotions tended to inhibit the use of executive processes \(^{121,123}\). This is consistent with other research which suggests that executive processes are impaired at temporarily elevated arousal states \(^{49}\).

Other research has specifically looked for predicted allostatic behaviours following an experimental manipulation. For example, infants were more likely to show gaze aversion, which down-regulates arousal \(^{124}\), following toy removal, which up-regulates arousal \(^{125–127}\) (see also \(^{117,128,129}\)). And highly aroused infants preferentially look to less arousing, low-frequency stimuli; whereas less aroused infants prefer more arousing, high-frequency stimuli (see Figure 3) \(^{130–132}\).
In real-world settings, we know that other types of events, such as loud noises and intense parent vocalisations also trigger arousal increases, but little research has investigated how techniques such as gaze aversion are used to instantiate allostasis in real-world settings.

We also know that some children are exposed to less predictable and more chaotic environments than others. Presumably, allostatic corrections would be required more often in children living in more chaotic environments. Overwhelmingly, the evidence suggests that household chaos associates negatively with a wide range of executive control outcomes as assessed via experiments and questionnaires, even after confounding factors such as SES are accounted for. We return to this point below.

Figure 3. (a) Illustrations of the checkerboards of varying frequency used by. (b) Highly aroused infants prefer to look at less arousing, low-frequency stimuli; whereas less aroused infants prefer more arousing, high-frequency stimuli.
Figure 4. a) Example of arousal fluctuations in a 12-month-old infant in home settings, over an 8 hour recording. b) Plot in which arousal levels are categorised into 5 bins, and the arousal level at time t is plotted against the arousal level at time t+1. For points on the 1:1 line (highlighted in red), arousal at time t is the same as arousal at time t+1. c) Schematic showing the predicted and expected outcomes. From 138

Other research has questioned whether allostatic mechanisms actually influence real-world arousal patterns at all during early development. One study examined day-long recordings of arousal fluctuations in infants 138 (see also 139 (see Figure 4)). Based on allostasis, they predicted that, if fluctuations above and below the mean are corrected for via self-regulation, then high or low arousal states should be more short-lived than intermediate arousal states. In fact, they found the opposite: across multiple time-scales, high and low arousal states were more long-lasting than intermediate arousal 138,139. In other words, extreme high and low levels of arousal appear ‘sticky’.
It appears reasonable to suppose that real-world allostatic mechanisms are relatively rudimentary during infancy, given that executive control is rudimentary overall; but this alone would not explain why extreme high and low levels of arousal lasted longer than predicted by chance (Figure 4). One possible explanation is that the opposite to allostatic processes also exist – termed ‘metastatic’, from the Greek word ‘meta’ meaning beyond – which actively amplify, and prolong, naturally occurring increases and decreases in arousal (see Figure 5). Intuitively, this does seem possible – for example, a hyper-aroused child seems to be more likely to bang their cutlery on a table at dinner time, which seems to increase their arousal still further; and a hyper-aroused child also appears more likely to be oppositional and get told off, which again increases their arousal still further. Similar self-amplifying maintenance factors have been identified, for example, in cognitive models of adult anxiety. But, because these are so hard to study in the lab, they are under-studied in child development.

Figure 5: Schematic illustrating negative and positive feedback loops, as commonly used in electronics, which can be used as a metaphor for understanding allostatic (dynamic self-regulatory processes) and metastatic (dynamic dysregulatory processes). The circuit represents a system with gain (G) and feedback (β). $V_{in}$ and $V_{out}$ show the input and output. The summing junction at its input subtracts the feedback signal from the input signal to form the error signal $V_{in} - \beta G$, which drives the system. In an allostatic system, the feedback term $\beta$ is negative. This is also known as negative, or degenerative feedback – similar to allostasis. In a positive feedback system, the feedback term is positive and so feedback increases the
overall gain of a system. This is also known as positive, or amplificatory feedback – similar to metastasis.

Two possible models present themselves through which we could understand how elevated arousal states can become self-sustaining. The first model holds that executive control and allostasis operate, as described above, to correct for minor increases and decreases in arousal; but, when excessive increases in arousal states take place, executive processes fail and are replaced by perseverative behaviours characterised by dynamical failures in executive function. If true, however, then we need to explain what it is that draws us towards elevated arousal states when executive processes fail: what makes them ‘salient’, or automatically attention-eliciting. To our knowledge, no research hitherto has addressed this question.

An alternative possibility, however, is that seeking out high arousal states is not a consequence of failures of self-regulation at all – but are, rather, a consequence of executive processes driving us to seek out a high arousal state to which we have become accustomed. Unpredictable environments are known to lead to long-term increases in tonic arousal. Because of this, it is likely that childrens’ ‘optimal’ arousal level (i.e. the level to which they are accustomed, to which they dynamically ‘try’ to return via allostasis) would change over time. Children accustomed to high tonic arousal may be using executive control to return, effortfully, to that state.

To differentiate between these hypotheses it is necessary to study how changes in long-term, tonic arousal affect which allostastic behaviours are used at any given level of arousal. To our knowledge, no research so far has examined this. This point could also, potentially, affect how we understand the results of experimental EF assessments, as most sustained attention paradigms
measure children's responses in a low-arousal setting. For children accustomed to low arousal, this might be a stable state; but for children accustomed to higher arousal this might be an unstable state from which they would seek to escape, via allostasis.

Overall, then, it seems that evidence exists that children use allostatic behaviours in real-world settings to maintain equilibrium across externally triggered increases and decreases in arousal. However, there may be evidence for opponent processes as well: metastatic or dysregulatory behaviours which actively amplify increases and decreases in arousal. However, a number of key questions remain unanswered. First, do different children have different levels of “optimal” arousal (cf.) – so that a given arousal level might elicit downregulation in one child but upregulation in another? Second, how does this “optimal” arousal level change over time, and how do these changes affect self-regulatory behaviours?

**Example 3: deciding how we allocate our attention in complex, dynamic, naturalistic settings**

During the first few months of life, infants’ attention allocation is driven largely by whatever in their environment is most salient (i.e., automatically attention-eliciting) as they develop, they gain the capacity to choose to reorient attention. This affects their behaviours when they are alone and during social interaction.

Shorter spontaneous attention durations in naturalistic settings are one of the diagnostic criteria for conditions such as Attention Deficit Hyperactivity Disorder in children. During early life, the Infant Behavior Questionnaire asks parents questions such as: ‘how often does your child
look at pictures in books and/or magazines for 5 minutes or longer at a time?” as part of parental observations that are predictive of later psychopathology \(^{153}\). Several studies looking at micro-level attention behaviours have found, for example, that infants who show shorter time intervals between refoveating eye movements during infancy show poorer effortful control during later development \(^{154–156}\).

But what is the exact role of executive functions in guiding how we allocate our attention on a moment-by-moment basis? As we discussed in the previous section, one role is that executive functions are used in arousal regulation, through behaviours such as gaze aversion. But what are others? This, of course, would differ between different types of naturalistic setting \(^{6}\); here we are considering attention durations during free play whilst playing alone with toys on a tabletop \(^{157,158}\) or rug \(^{159}\). Importantly, in these settings there tends to be little externally generated movement present in the child’s field of view, which is the strongest driver of exogenous salience \(^{160,161}\). We will also concentrate purely on overt attention - although the distinction between overt and covert attention is important and well studied in adults \(^{162–165}\) and children \(^{166–170}\). And we concentrate on attention in solo settings, although the influence of social context on child attention is also well researched \(^{171–173}\). Finally, we will also concentrate on the durations of individual attention episodes (i.e. how long we look for), rather than the equally valid and well-studied question of where we look in naturalistic settings \(^{160,174,175}\).

Approaches that have modelled attention durations in naturalistic settings have, broadly, concentrated on two factors that operate in interaction. The first is an oscillatory ‘foraging’ component which triggers reorientations at regular time intervals. Evidence for such a component has been shown using modelling work looking at saccade generation timing with adults \(^{176,177}\) and infants \(^{178}\) as well as at a larger temporal scale, between attention episodes on a
time-scale of seconds. These periodic attention shifts, which are often associated with oscillatory fluctuations in perceptual sensitivity, are comparable to animals’ rhythmic sampling of the environment through whisking (Kleinfeld, Deschénes, & Ulanovsky, 2016), sniffing (Ahrens & Kleinfeld, 2004) and licking (Amarante, Caetano, & Laubach, 2017).

Of course, if a periodic reorientation timer were the only factor determining attention durations, then all attention durations would be equally long. Modelling work suggests that the factor that best captures how naturalistic attention durations vary between individual attention episodes is hysteresis, or attention inertia: the longer an attention episode lasts, the more its likelihood of ending during the next successive time interval diminishes. Attention durations in a variety of different naturalistic attention contexts show hysteresis, such as during free play, reading and looking towards a screen.

EFs may play a role in both of these factors, in different ways. First, there are some circumstances in which phasic decreases in EF should associate with longer look durations. First, and particularly during early development, it is thought that longer attention durations can arise from failures to disengage, known as ‘sticky fixation’ (although see). This may be more influential during situations where the child’s field of view is more exogenously salient (e.g. during TV viewing).

Second, attention lapses can take place, when overt and covert attention become decoupled – known as mind-wandering. Again, this leads to longer overt attention duration patterns. This is well studied in adults where studies have shown, for example, that mind-wandering is more common in individuals who score more poorly on tests of ‘traditional’
executive functions such as working memory and inhibition\textsuperscript{162,192,193}. However, it is relatively under-investigated in children\textsuperscript{156,170}.

Some work has examined how periodic ‘foraging’ relates to EF development. For example, one study looked at periodic fluctuations in children’s looking behaviour to static pictures, and found that children whose attention patterns were more periodic showed faster learning and discrimination\textsuperscript{179} (see also\textsuperscript{194}). This may be because internal oscillators help maintain stable inhibitory/excitatory balances that are crucial for information processing as they organise the information in regular periods of information intake and states of limitation on input\textsuperscript{50}. Studies that look at spontaneous attention behaviours in altered naturalistic environments, where the information is evenly distributed and so differences in salience at the point of regard are less likely to influence attention durations, would help understand this further.

Overall, however, the largest strand of work examines the positive association between EFs and attention hysteresis – consistent with the behavioural findings noted above that both ADHD diagnoses and lower scores on executive control/orienting during infancy arise from shorter naturalistic attention durations. Attention hysteresis increases with increasing age\textsuperscript{183} and is lower in children with ADHD\textsuperscript{195}. We also know that attention hysteresis correlates with phasic increases in EF\textsuperscript{183,196,197}. Experimental work has shown, for example, that distractibility shows phasic decreases during an attention episode\textsuperscript{198,199}. Evoked neural responses to unattended objects also phasically decrease during an attention episode\textsuperscript{197}. Finally infants’ ability to subsequently recognise information increases during an attention episode\textsuperscript{200,201}.

But, although it is well documented that EFs and attention hysteresis tend to co-fluctuate, the question of what causes what remains unclear\textsuperscript{197}. Many researchers would, implicitly or
explicitly, hold that the cognitive faculty of sustained attention \textit{causes} attention hysteresis \textsuperscript{202}. But there is little direct evidence to suggest this. Experimentally assessed sustained attention declines over time (Esterman & Rothlein, 2019), whereas attention hysteresis (by definition) increases. And cognitive training studies found that training attention in infants, which increases looking duration to pictures, leads to \textit{shorter} attention durations in one naturalistic setting \textsuperscript{203,204}.

It is possible that both EFs and attention hysteresis might be jointly influenced by other factors. One possibility is the physical position of the play: longer attention episodes might afford more possibility to ‘create’ a situation in which one object is more salient than its competitors – for example, by moving it or bringing it closer \textsuperscript{205,206}. Another possibility is that both are influenced by some third cognitive factor, such as comprehensibility. We know that comprehensibility influences attention durations: the likelihood of an attention episode ending increases around natural event boundaries \textsuperscript{207,208}; more attention hysteresis is observed for comprehensible TV clips (in older but not younger infants) \textsuperscript{209}; and during joint play parents can extend infants’ fixation durations by presenting semantic information about the object being viewed \textsuperscript{210}. In children, distractibility decreases during semantically critical points during a story \textsuperscript{195,211}, but the same effect is not observed in children with ADHD \textsuperscript{195}. Future work should investigate this question further.

Overall, then, it seems that executive control influences naturalistic looking behaviours in a variety of ways, but that our current understanding is limited. One way to improve this would be to record neuroimaging during everyday attention episodes \textsuperscript{8,212} to examine how mechanisms identified as top-down control mechanisms (including phase modulation, cross-frequency coupling and intra-brain entrainment) dynamically fluctuate during real-world attention \textsuperscript{37,40}. This would allow us to understand how these fluctuations associate with changes in everyday
behaviour \textsuperscript{150} (see also \textsuperscript{213}). Experimental manipulations could also help address the question of whether phasic increases in EF and attention hysteresis are directly causally connected, or whether both are consequences of some third factor. This will help address the theoretically important question of whether EFs ‘drive’ real-world attention behaviours, or whether they are instead better seen as correlates of other factors, such as comprehension.

**Conclusion**

In this article, we have discussed three strands of research that examine how EFs are expressed through our moment-by-moment interactions with the environment. We have concentrated on how these interactions change around the moment, during early development, when it is thought that the capacity for volitional control of behaviour starts to emerge during the first years of life.

In part we used the metaphor of the ghost in the machine to describe how the active, volitional component of cognition (the ghost) steers and guides the machine (our brains and bodies). The first important point to have emerged is that the machine isn’t sitting idle on the tarmac at the time when the ghost first starts to take control: it is already moving, driven along passively by multiple streams of sensory information in the environment. The ghost learns to steer the machine by starting to spot patterns: using temporal predictability and contingencies in the environment to learn first to make predictions, and then to decide when behaviours might be optimally timed to navigate through the changing environment. These predictabilities and contingencies can be understood as emerging both between the child and its physical environment (e.g. if I bang this table it will make a noise) and between the child and its social environment (e.g. if I smile at my father he will smile back).
Temporal regularities also guide how we voluntarily reorient our attention in the real world: our shifts of both overt and covert attention show periodic, oscillatory properties. In addition, though, our environmental interactions also show contingencies: we use active, volitional control in our interactions with our environment to reactively respond to change, through allostatic corrective mechanisms that allow us to maintain equilibrium in the face of external change. Although there is evidence that these self-regulatory mechanisms are present even from an early age, much remains to be discovered about how we self-regulate, moment by moment, through our interactions with the environment.

During early life, there is also evidence that perseverative behaviours can develop, that take on a self-sustaining, amplificatory characteristic – such that an increase in arousal can trigger a chain reaction of events that leads to progressively greater increases in arousal. These chain reactions can take place through interactions both between the child and the physical environment and between the child and the social environment (people around them). Again, though, there is much still to be discovered about what gives these amplificatory cascades their ‘sticky’ nature, i.e. understanding what causes perseveration.

Practical implications

What, though, are the practical implications of this work? I have a child who is showing atypical attention patterns at school, and I want to understand better why. At the moment, our first approach would be to administer a battery of tests to measure their executive function ability. In other words we replace one score – for their academic marks – with another score, of a static,
time-invariant, abstract mental construct. We do this even though the relationship of one to
the other are not straightforward. Arguably, this approach is proven now to be of little
therapeutic benefit: for example, we know that training abstract mental constructs causes little
changes in actual real-world behaviours.

In the not-too-far future (say, twenty or thirty years from now), it may be that another child in a
similar situation would be met with a very different response. We might record multi-modal
quantitative observations of their behaviours in naturalistic settings – the behaviours that led to
them getting flagged in the first place – and then use computer modelling to try to understand the
cause for why their behaviours have been flagged as atypical. This might reveal that, for
example, their self-regulation is effective, but that the average level to which they are trying to
return is different to other children (section 3); or uncover a variety of reasons why attention
episodes tend to go on either too long, or not long enough (section 4). It might also lead to
different therapeutic solutions, such as increasing the predictability of their home environment
(section 2), or focusing on increasing the role of comprehension in sustaining attention (section
4).

We started this article with the simile that studying EFs as purely internal mental constructs is
like ‘building castles in the sky’. We have argued that, on its own, this approach is insufficient;
and that the question of how EFs are expressed through the inter-relationship between an
individual and the world around them remains under-explored. As Henry David Thoreau put it:
“If you have built castles in the air, your work need not be lost; there is where they should be.
Now put foundations under them”.
References

1. Danziger, K. *Constructing the subject: Historical origins of psychological research.* (Cambridge University Press, 1994).


50. Perapoch, M. Wass, S.V. How the ghost learns to drive the machine: oscillatory entrainment to our early social or physical environment and the emergence of volitional control. *Preprint.*
Running head: EF AND THE ENVIRONMENT


Running head: EF AND THE ENVIRONMENT


Running head: EF AND THE ENVIRONMENT

158. Wass, S. V. et al. Infants’ visual sustained attention is higher during joint play than solo play: is this due to increased endogenous attention control or exogenous stimulus capture? Dev. Sci. (2018).


206. Schroer, S. E. & Yu, C. Multimodal attention creates the visual input for infant word learning. in 1–6 (IEEE, 2021).


