Behavioural Processes 162 (2019) 205–214

Contents lists available at ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc

Models of attention-deficit hyperactivity disorder

Peter R. Killeen

Arizona State University, United States

ARTICLE INFO

Keywords:
ADHD
Attention deficit hyperactivity disorder
Diffusion-decision models
Lactate
Astrocyte-neuron lactate shuttle
Neuroenergetics

ABSTRACT

One of the most notable aspects of the behavior of individuals with Attention Deficit Hyperactivity Disorder (ADHD) is increased variability in many aspects of their behavior, including response times and attentional focus. Among the many theories of ADHD is one that identifies its material cause as phasic malnutrition of the neurons required to maintain constancy of performance. Of the diverse predictions issuing from this theory, one concerns ubiquitous data: response times and their variance in decision tasks. This paper reviews that behavioral neuroenergetics theory and model, shows how they predict representative data, and suggests their relevance to researchers studying animal models of ADHD.

1. Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is the classification of a common neurodevelopmental constellation of atypical behaviors. The Diagnostic and Statistical Manual (DSM-IV) uses five main criteria to define ADHD: Criterion A1 identifies 9 signs of inattention; A2, 9 signs of impulsivity/hyperactivity. Evidence of persistent and maladaptive manifestations of at least 6 of the former or 6 of the latter are required for "categorization." Criteria B (age of onset) and C (symptom presence in two or more settings) require temporal and contextual generality in the symptoms; Criterion D (domains of impairment) requires that the symptoms impede progress toward achievements valued by society; Criterion E (exclusionary diagnoses) lists diagnoses which trump ADHD in accounting for the symptoms. The latest version of the DSM juggles these criteria slightly. The prevalence of these symptoms is relatively constant across time and cultures (Polanczyk et al., 2014), despite different diagnostic criteria generating apparent disparities. The rate under the DSM-IV is 5–7% in the US, with about twice as many boys categorized as girls.

From the above it is obvious that children (and adults) are categorized by a check-list of signs, subject to the evaluations of parents and teachers (whose judgements are often disparate: Hartman et al., 2007; Mitsis et al., 2000), ideally after interviews with an expert diagnostician. Although there are objective psychomotor tests with some discriminant validity, they are not specific enough to stand alone, and the DSM has not included them as adjuncts to the judgmental criteria. Heritability of ADHD is high—almost as high as the heritability of height—and a number of genes have been implicated (Sharp et al., 2009; Cortese, 2012; Klein et al., 2017), but none of them account for more than a trivial amount of variance in the manifestation of ADHD. Unlike the situation for disease, where a pathogen can often be identified that plays a necessary (even if not sufficient) role in the illness, there are no biological markers of ADHD. Once individuals are categorized, physical concomitants can be found, for instance delayed myelination, reduced size of cerebellum, etc.; but again, these are not by themselves distinctive enough for categorization.

The lack of non-judgmental criteria for classification is due to at least two circumstances. The first is that the DSM is generally unable to identify "taxons"—that is, diagnoses as distinct as a broken leg vs. a broken arm. Psychiatric diagnoses blend into one-another, and presentation of one DSM category in an individual without any others is uncommon: comorbidities are the rule rather than the exception. The second reason is that there is no generally accepted theory of ADHD. Many extant theories are just names of cognitive processes believed to be at fault: Delay Aversion, for example, explains impulsiveness and steeper delay discounting as due to dislike of delays. Executive Function debility explains difficulty of maintaining attention as failure to reallocate resources due to insuffient cognitive energy. These are discussed further in Section 4. Another hypothesis followed from the immediate amelioration of the symptoms by methylphenidate (MPH; Ritalin), a dopamine reuptake inhibitor: It was that ADHD is caused by an insufficiency in dopamine. That insufficiency was never demonstrated, and amphetamines also ameliorate the symptoms with negligible impact on the dopamine system. A behavioral theory of ADHD proposed by Sagvolden and associates (Sagvolden et al., 2005; Catania, 2005) posited shorter steeper delay of reinforcement gradients as one of the factors that caused behavioral deficits in ADHD. There is some evidence for these in an animal model of ADHD (Johansen, Killeen, Russell, et al., 2009; Johansen et al., 2009a; Pellón et al., 2018), along with the predicted greater entropy of responding. But even if this is the

E-mail address: Killeen@asu.edu.

https://doi.org/10.1016/j.beproc.2019.01.001
Received 17 November 2018; Received in revised form 16 January 2019; Accepted 16 January 2019
Available online 21 January 2019
0376-6357/

case, what causes it?

It is the purpose of this paper to briefly contextualize the issues; outline a hypothesis that I have been involved in; generate predictions from that hypothesis, and review recent efforts to test that theory of ADHD. Those familiar with recent developments in the field should skip ahead.

2. The neuroenergetics theory (NeT) of ADHD

While on leave to the Institute for Advanced Studies in Oslo, an opportunity organized by Terje Sagvolden, my colleagues and I shuffled old ideas and new ones about ADHD, always discontent with those approaches and their rearrangements. On a winter's hike some of us came up with what we thought a novel approach to the disorder (Russell et al., 2006, although we soon found precedents). The thesis is that ADHD is a neuroenergetic disorder. The problem is not with dopamine supply, but with energy supply (which dopamine can release). The brain constitutes about 2% of the body's weight but uses 20% of its energy. Much of this energy goes to reestablishing ionic gradients after a neuron has discharged. The neurons themselves cannot store enough energy to fire fast reliably for more than a dozen seconds. Once they have begun spiking, the neurotransmitters they release at the synapse (in particular glutamate) also stimulate the nearby noradrenergic varicosities. These release norepinephrine, which stimulates the nearby astrocytes, one of several glial cells in the brain. Astrocytes play many roles (Kiray et al., 2016), an important one being the abstraction of glucose from the capillaries, conversion of it to glycogen for storage, or directly to glucose, and thence to lactate (Magistretti and Allaman, 2018). These processes are stimulated by the norepinephrine released by the varicosities. A thorough introduction to this model is found in Killeen et al. (2013).

The focus is on the Astrocyte-Neuron Lactate Shuttle (ANLS: Pellerin and Magistretti, 2011), involving the transport of lactate out of the astrocyte and into the neuron. We believe this to be the bottleneck in energy supply to the neurons—although, given the complexity of neuroenergetic processes (Pellerin et al., 2007), it could certainly be earlier in the supply chain. Lactate also modulates various genes, some of which are associated with hyperactivity (Margineanu et al., 2018). Like most mental disabilities ADHD is quite heterogeneous, and some of that may be due to different points at which the energy supply is bottlenecked (such as mitochondria: Kramer and Bressan, 2018). The much greater incidence of ADHD in boys might be due in part to problems with the Y-chromosome gene SRY, which encodes a protein controlling the concentration of dopamine (Pinares-Garcia et al., 2018). The processes of energy conversion in the astrocytes are quite complex. The diversity of potential bottlenecks makes some sense of the diversity of approaches that have had some success (e.g., Kaplan et al., 2015). Some things will work with some kids, and not with others, so that evaluating interventions in terms of group statistical differences is perilous (Williams, 2010). For most of the predictions here, the problem must reside either in the astrocytes or components of the ANLS. What are the behavioral implications if this hypothesis is correct? There are a number of qualitative ones listed below. There are also suggestions for clinical interventions. But it is the quantitative predictions that will be the focus of this paper. I distinguish the theory of energetic insufficiency (NeT) from the model of response times in psychometric tasks—the Neuroenergetics Model with Attention lapses (NeMA) that draws out some of NeTs implications.

2.1. Variability in ADHD

There are no well-defined distinctions among DSM categories (as noted above, they are not taxons); most individuals with ADHD have at least one comorbid condition. ADHD constitutes a range of scores in a multidimensional character state (Coghill and Sonuga-Barke, 2012), parts of which are also occupied by other psychiatric disorders (Polanczyk et al., 2007). There is a common theme to many of the associated syndromes: comorbidity among ADHD, reading disabilities and mathematical disabilities, for instance, "is due to common genetic risk factors that lead to slow processing speed" (Willcutt et al., 2010). This is typically accompanied by increased variability in attention and response characteristics, increased error rates, and other concomitants of that slowing (see the heroic meta-analysis of Kofler et al., 2013 showing an effect size for RT variability in ADHD of $g = 0.76$). For a more complete situating of the analysis than offered in the present paper, see Killeen et al. (2013). The focus of this paper is primarily on the symptoms of ADHD as manifest in neuropsychological tests.

Although familiar earmarks of ADHD are impulsivity and hyperactivity, another more easily measured characteristic is variability. ADHD motor behavior is more variable, spatial focus is more variable, and timing is more variable. Information processing also tends to be slower and more variable. One cause for this is the inadequate myelination of neurons, as myelin makes signal transmission both faster and more energy-efficient. Another may be the quick depletion of fuel and the slow reprovisioning of it. In tasks requiring sustained attention, therefore, performance decrements faster with time on task in ADHD than in Typically Developing controls (TD). Stimulants reenergize performance by impeding re-uptake of norepinephrine, whose increased level further encourage the astrocytes to release lactate. The astrocyte has substantial stores of energy in its glycogen, but it takes longer to convert this into lactate; and takes many hours to restock the glycogen—a task usually accomplished during sleep. (Thus the importance of adequate sleep for children, and the frequent need of individuals with ADHD to take time-outs in quick naps). As network hubs deplete their fuel, they become less able to maintain synchronous high-frequency firing across brain regions, which dulls their ability to process information. Responses take longer, their latency becomes more variable, transient targets are missed, and more time is spent off-task. Minds wander.

To better understand the nature of NeMA I address for it the five properties of computational models that require explication (Cowell et al., 2012): 1) *Biological Level*. NeT is concerned with the energetic supply of neurons, with neuroglia at center stage (Fig. 1); with diffusion upward toward neural hubs (Sporns, 2013) and systems; and downward toward genetic mechanisms. 2) *Problem space*: Cognitive impairment, as manifest in tasks that measure processing speed, seen in response-time distributions, their means and variances, attentional fugacity, impaired working memory capacity, coordination and timing. Other signs and comorbidities, such as hyperactivity and oppositional defiant disorder (which may be compensatory: Medin et al., 2018; Lea et al., 2018; Sarver et al., 2015), are outside the space. 3) *Biological plausibility*. A sustained argument for biological plausibility was writ in (Killeen et al., 2013). 4) *Parsimony*. NeMA is parsimonious to a fault. There are scores of parameters relevant to the neuroenergetic process (Cloutier et al., 2009); my early attempts to use these empirical models could not reduce their degrees of freedom sufficiently to support any predictions. NeMA could profitably add parameters to enable quantitative predictions of speed-accuracy trade-offs, but that work is now being better accomplished with more traditional DDMs. 5) *Predictions and postdictions* for humans diagnosed with ADHD, and criteria for animal models of it:

a) Rate of approach to asymptotic performance (such as accuracy, speed or time-on-task) depends on the sum of rates of supply and demand; where supply is compromised, the approach will be slow and the asymptote lower (Borger et al., 1999; Huang-Pollock et al., 2006); the same is the case when the demand is increased, as in poor myelination or more intense processing demands (see Huang-Pollock et al., 2012 for a meta-analysis of performance over time; Rapport et al., 2009, however found constant low levels of time on task for ADHD in a school setting). Stimulants, which increase energetic supplies, will ameliorate these effects, as will task time-outs.

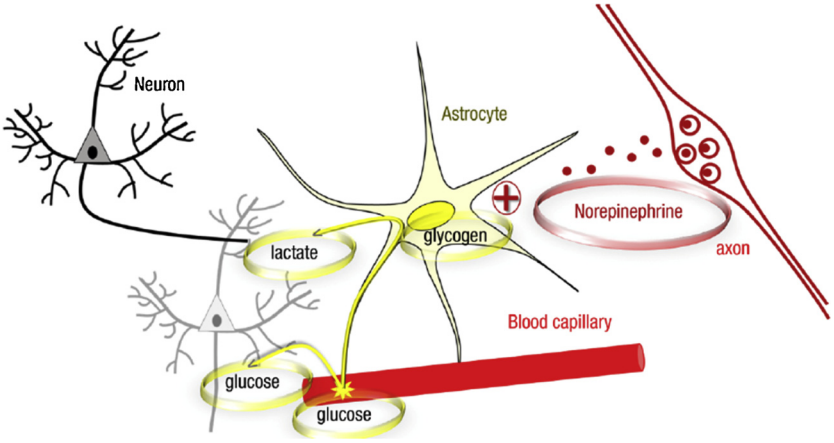

Fig. 1. Illustration of glucose and lactate provisioning neurons. Glucose is the brain's prime energy stock. Lactate is faster and less costly a fuel to convert to ATP than is glucose. Neurons release glutamate at synaptic connections to signal other neurons. The glutamate also signals astrocytes to absorb glucose from blood capillaries (a process generating the BOLD signal) and release lactate to neurons. Norepinephrine released from spiking axons further stimulates astrocytes to convert glucose stored as glycogen to lactate, releasing it to the hungry neurons.
Figure drawn by Russell, reprinted with permission from (Killeen et al., 2016). For a more detailed diagram, see Fig. 1 in (Killeen et al., 2013).

b) Whenever astrocytic adrenoceptors are compromised, or myelination faulty (e.g., multiple sclerosis, narcolepsy) symptoms similar to those of ADHD will result and similar medications are of some value. Astrocytes play an important role in myelination and remyelination after injury (Salmi et al., 2018), another reason to suspect their implication in ADHD.
c) All medications that are effective in ameliorating phasic symptoms of ADHD will do so by stimulating lactate release from astrocytes in relevant brain systems. Stimulants most effective in treating ADHD all do this (Jensen et al., 2005). Non-stimulants that activate the glia's adrenoreceptors (e.g., guanfacine, clonidine) are also effective.
d) Any intervention that stimulates the release of lactate into the extracellular space will ameliorate ADHD symptoms. Lactate stimulated by physical activity crossing the blood-brain barrier (Medin et al., 2018) is a potential instance not involving stimulation of glial adrenoreceptors.
e) Dopamine insufficiency plays no role in ADHD (Solanto, 2002). Monoamine reuptake inhibitors are effective because norepinephrine and dopamine stimulate astrocytes to release lactate.
f) Neural synchronization and maintenance of LTP are energetically demanding, yet essential for learning. Therefore individuals with ADHD will be disadvantaged in complex tasks involving learning (Huang-Pollock and Karalunas, 2010), memory (Karalunas and Huang-Pollock, 2013; Sjowall et al., 2013), or sustained attention, particularly at slow rates of stimulation.
g) Wandering of attention and discursive action will be especially pronounced for individuals with ADHD in monotone environments. It occurs because functional neural hubs fatigue more quickly, and other less fatigued hubs command attention (Salmi et al., 2018).
h) Group performances of ADHD vs. typically developing controls (TD) will diverge with time-through-trial and with time-on-task (Hervey et al., 2006). The same is true for ADHD on and off medication (Epstein et al., 2006; Scheres et al., 2003).
i) Slower presentations that require sustained attention will impair performance (van der Meere et al., 2009). Presentations that are slower because they permit relaxation of attention between trials should improve performance.
j) Power spectra of latencies that have not been detrended will show greatest discrepancies at low frequencies corresponding to the early moments on task as subjects transit to energy depletion (Castellanos et al., 2005). Detrended spectra will show no differences (Karalunas, Huang-Pollock, et al., 2013).
k) Quality of performance of ADHD will be improved in self-paced tasks, where, to a limited extent, slower processing would not incur an accuracy penalty. This prediction generally does not hold (Koschack et al., 2003; Dalby et al., 1989; Rommelse et al., 2007; Lee et al., 2008). Instead, ADHD individuals seldom slow down to permit more complete processing, even time allowing.
l) Wherever fast coordination of neural networks is required NeT predicts deficits in performance. This is found in ADHD who have significantly more speech and reading problems (Levy et al., 1996) motor coordination problems (Fliers et al., 2008) and cardiac markers (Beauchaine et al., 2013). Individuals with ADHD are more accident-prone (Swensen et al., 2004).
m) Sometimes quantitative predictions are also possible as described in the remainder of this paper.

3. DDMs

These ideas can be translated into predictions with a class of models called drift-decision (or sometimes drift-diffusion) models (DDMs). Simen et al. (Simen et al., 2013) provide a history of such models with application to interval timing, and Ratcliff, McKoon, and others (Ratcliff and McKoon, 2008; Ratcliff et al., 2015, 2016) provide authoritative overviews and recent status reports. In the simple case shown in Fig. 2, the rising traces represent the neural progress toward a decision as a random walk with upward drift. The parameters of all curves are the same (randomly two steps forward, one back, representing contradictory impulses to action). When they cross a threshold for action (here set at 75) a response occurs. The density shown in the top panel is a Wald (or Inverse Gaussian) distribution, the form that such processes yield (Chhikara and Folks, 1989). Imagine the random walks incrementing a bar in a relative frequency diagram each time they cross threshold; then the density in Panel A gives the long-run envelope of that diagram. In this case there are two parameters, the drift (the information-accumulation) rate v, and the criterion (here $c = 75$). The mean of the Wald is $m = c/v$ and its variance cv^{-3}. It follows that its coefficient of variation is $(cv)^{-1/2}$; the standard deviation is proportional to $m^{3/2}$. Variability therefore grows faster than the mean, faster than Weber's law.

NeT proposes that the velocity of conduction of information by neural assemblies is an increasing function of the energy available to the neurons, so should be lower for ADHD than for TD, and should be higher when on medication than off. Because variance grows as an inverse function of the cube of velocity in DDMs, it will be significantly larger when velocity is slow, as predicted for ADHD and routinely found. Since SD grows faster than the mean, variability differences should remain after controlling for mean RT, as they do (Kofler et al., 2013). The predictions of c depend on context. If time is of no consideration in the task either to the experimenter or to the subject, then values of c may be comparable in ADHD and TD, causing longer latencies without decreases in accuracy. If there is any explicit or implicit time urgency, the ADHD are forced toward a greater speed/accuracy tradeoff (Mulder et al., 2010; Seli et al., 2013; but see Karalunas et al., 2012; Losier et al., 1996; Metin et al., 2013), which they may achieve

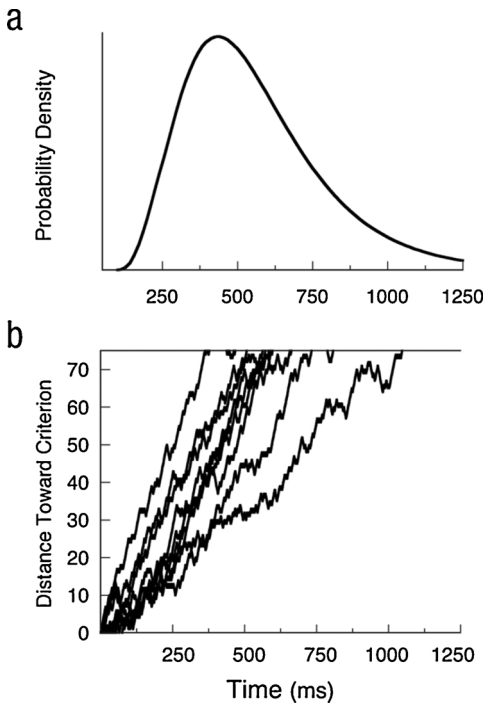

Fig. 2. A schematic illustration of the drift-decision model of response times. Panel b shows eight random walks with drift, representing the interplay of neural processes of excitation and inhibition that cumulate in a response when the criterion at the top of the panel (c = 75) is crossed. Simulation of the process many times results in the distribution of criterion crossings shown in Panel a.
Adapted from (Killeen, 2013).

by lowering their criterion for responding (*c*) in experimenter-paced tasks, or raising it in self-paced tasks, except when premature responding has become habitual. When a lowered criterion becomes habitual, it generates a trait of responding before information accumulation is complete—impulsivity. See Fig. 17 of Killeen et al. (2013), which displays a strong negative correlation between *c* and an independent measure of impulsivity.

Fig. 3 shows the fit of the Wald to response time data (Querne and Berquin, 2009). About every 3 s a light appeared on screen and subjects had to report its location (high or low) by pressing a key. The four groups of subjects are identified in the caption, and parameters for the averaged curves are shown in the right panel of the figure. As severity of symptomatology increased, processing speed decreased. The criterion was lowest for the hyperactive-impulsive, highest for the typically-developing (TD) group.

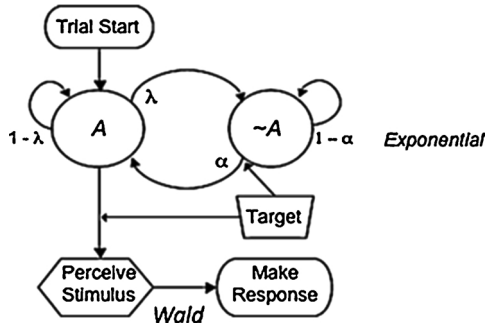

Fig. 4. The Neuroenergetics Model with Attentional lapses (NeMA). The start of trial puts the participant in the attentive state A If in the A State when a target stimulus appears, the subject responds according to the Wald distribution. Attention lapses with a probability λ in one time-step (say, 1 s). If the subject is inattentive (is in the ~A state), he must refocus attention to process the stimulus and initiate the response. In experiments with inconspicuous stimuli (vigilance or detection experiments), attention drifts back to A at a rate of α. In experiments using salient stimuli (as analyzed in this paper), those capture attention, driving the initiation of processing by greatly increasing α. A simple Wald process will control responding with probability of 1 - λ. The latency to recapture attention adds an exponential process with mean $\tau \approx 1/\alpha$ with probability governed by λ. In that case, a convolution of the Wald and exponential—the ex-Wald—will control latency. The final distribution of responses is a mixture of the Wald and the ex-Wald.

3.1. Attention

This model, the simple Wald, assumes that subjects are always "on task", but everyone's attention lapses (Schooler et al., 2011), with as much as 50% of people's thinking not under the control of external stimuli (Killingsworth and Gilbert, 2010). This is especially the case for individuals with ADHD, as maintaining attention requires focused neural energy, which they quickly deplete, causing their famously delinquent attention. An experimental paradigm that permits participants to replete energy (or at least not waste it) with time-outs between trials will help everyone, but especially those with ADHD. If, however, the task is slowed by increasing the inter-stimulus interval (ISI), while requiring attention throughout, this will hurt everyone, but especially those with ADHD (Tye et al., 2016). It will do so both by slowing rate of processing, and by making lapses of attention more likely. Such lapses will on the one hand give the neural hubs (Sporns, 2013) involved in the task time to replete, but on the other will increase response latency by a variable amount as attention is dragged back to the task, and may cause more frequent missing of targets if they are inconspicuous. Fig. 4 sketches a model of this process.

NeMA introduces a new distribution, the ex-Wald (Schwarz, 2001), a convolution of the exponential distribution and the Wald distribution,

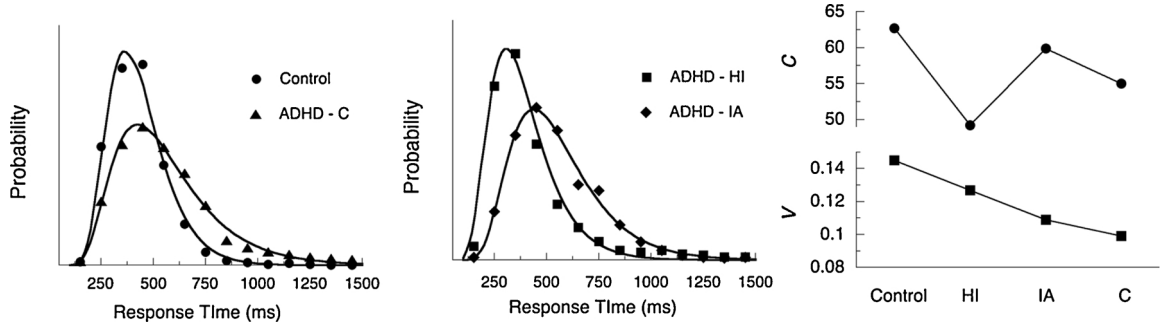

Fig. 3. Relative frequency distributions of response times on trials with a correct response for four groups of children (Querne and Berquin, 2009). The curves are Wald densities, with parameters displayed in the right panel. HI is Hyperactive–Impulsive; IA is Inattentive; and C is Combined types. As the groups are arrayed in the right panel, processing rate (*v*) decreased across groups from Control to Combined types (typically the most severe categorization). Criterion (*c*) was highest for the Control group, and lowest for the Hyperactive/Impulsive.
The figure is redrawn from (Killeen et al., 2013, Fig. 4) with permission.

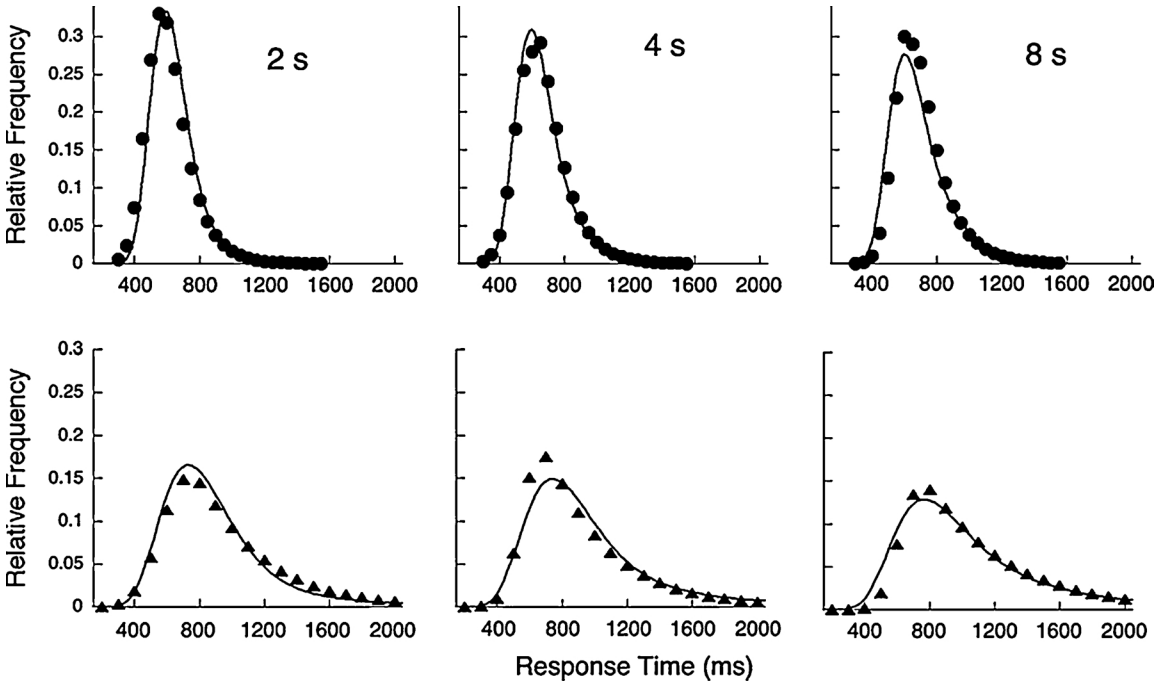

Fig. 5. The response-time distributions of 18 typically-developing (TD) boys (top panels) age-matched to 17 ADHD boys (bottom panels) on a continuous performance task (Leth-Steensen et al., 2000). The parameter in each column gives the ISI. The Neuroenergetics Model draws the curves. The parameters for (TD, ADHD) are: ν (0.24, 0.13); c (138, 96); λ (0.21, 0.27), and τ (0.20 s, 0.43 s).
Figure from (Killeen et al., 2013), and reprinted with permission.

that governs responses from the inattentive state. It introduces two additional parameters, λ, the probability of a lapse of attention during any time-step, and τ, the mean time to recover attention (the mean of the exponential part of the ex-Wald distribution). The predicted response latencies are a mixture of the Wald with probability 1 - λ, and the ex-Wald with probability λ. Because the Markov process shown in Fig. 4 is always running, it predicts increasing lapses of attention as ISI increases, as governed by λ, the probability of a lapse in any one time-step; see the appendix for details. Note that in such conditions, the subjects must maintain attention throughout the trial; if breaks are given between trials, it will allow them to restore a bit of the depleting energy. Thus there can be contrary effects of rate of presentation, depending on whether attention must be maintained or can be relaxed. It can be seen from the caption of Fig. 5, where the model is applied to the data of boys on a continuous performance task, that the parameters bear out the predictions of the theory: ν is approximately twice as fast for TD, c is greater for TD; the probability of a lapse of attention is smaller for TD, and the time to recover from a lapse is smaller for TD. Mean and variance of the distributions increase with ISI.

4. Better analyses, independent confirmation

The above is but a small sample of the maps between theory and data, and between model and data, issuing from NeT. There are two problems with these analyses: 1. They are based on group data, whose averaging may distort the resulting distributions, and afford no measures of variability in the parameter values; 2. All analyses were performed by an individual with a stake in this theory. Fortunately, the use of DDMs to study such developmental disorders has become currently appreciated. These DDMs are different from NeMA as they do not have mechanisms for lapses of attention, but do include two additional parameters omitted in NeMA: an initial minimal latency to begin processing and responding, and a bias to respond or not, or to respond for one alternative or the other. Some also contain a parameter for variability in drift rates from one trial to the next.

Karalunas and associates (Karalunas et al., 2014) conducted a meta-analysis of RT variability in ADHD. In many of the studies the data were fit by an ex-Gauss model, similar to an ex-Wald. For these data they concluded “RT variability in ADHD is primarily explained by very slow RTs on some trials (τ parameter) and by slow drift rate” (p. 37), consistent with our theoretical predictions and the data in Fig. 5. In 5 studies that employed a DDM, the effect size for slower drift rate was a substantial 0.63. They also found a smaller but significant effect for non-decisional times, with the ADHD group being faster. There was no reliable difference in criterion.

Huang-Pollack et al. (Huang-Pollock et al., 2017) tested the predictions of NeT. Approximately 50 children with ADHD and 20 TD in each condition were tested with fast and slow rate Go-NoGo discriminations. In the slow group, after a response was made, the stimuli remained on for 1.5 s, whereas in the fast group the program immediately moved to a 300 ms ITI and thence the next trial. Assuming that the subjects in the slow group could use the extra time to relax, perhaps briefly close their eyes, we would expect slower drift rates in the fast vs. slow condition for ADHD. “Contrary to longstanding belief [but consistent with NeT], we found that fast event rates slowed the rate at which children with ADHD accumulated evidence to make a decision to “no-go”, as indexed by drift rate” (p. 57). Drift (accumulation rate) on “go” trials was also slower than for controls, but not significantly so ($p < .06$). This difference may be due in part to the majority of the trials being “go”, and some of that speed may have been due to automatic responding.

Fosco and coauthors (Fosco et al., 2017) found information accumulation to be much slower in children with ADHD, and that both methylphenidate (MPH) and reinforcement sped up information accumulation. The former is predicted (because MPH stimulates the astrocytes to release lactate) but the effect of reinforcement, while reasonable in that the kids were differentially reinforced for responding quickly, is not in the domain of NeMA.

Weigard and associates (Weigard et al., 2018) conducted a choice-RT task on 80 children with ADHD and 32 age-matched controls. Children with ADHD were substantially slower in accumulating evidence on trials on which they were correct than the TD, but there was

no significant difference on error trials. The authors note that the first effect is consistent with NeT, but the second inconsistent. They speculated that neurons less relevant to the task that contribute noise (and errors) may be less heavily used than those contributing to a correct choice, and therefore not as depleted. Overall, however, drift rates were quite a bit slower on error trials, so other explanations are needed. Erroneous trials are by that fact quite difficult, and children confronting such perplexing stimuli may after a time just opt to respond randomly to escape them and move on. The authors felt that their data ruled against "acute attentional lapse" theories of ADHD; their measure of attentional lapse was increased variability in drift rates between trials. Very brief but frequent lapses would, however, merely contribute variance to drift rates. However, the variance in drift rates was apparently no greater for ADHD, and looked smaller, so this does not contravene their judgment against acute attentional lapses.

5. Comparison of NeT with other theories

There are numerous theories of ADHD; Johnson et al. (Johnson et al., 2009) review four major ones:

5.1. Executive function (EF) disorder

EF disorder is "used to explain deficits in "higher-order" cognitive processes, such as planning, sequencing, reasoning, holding attention to a task, working memory, inhibition of inappropriate and selection of appropriate behaviours" (Johnson et al., 2009, p. 1). Its physical substrate involves neural circuits that link the frontal cortices with the basal ganglia, thalamus and parietal cortices. An important meta-analysis (Willcutt et al., 2005), found significant impairments in executive functions, in particular in response inhibition, vigilance, working memory, and planning, that were associated with ADHD. But because of many exceptions, they concluded that "EF weaknesses are neither necessary nor sufficient to cause all cases of ADHD" (p. 1336). EF does not specify in what way the implicated neural circuits are weakened, nor why, nor predict the RT data that NeT predicts and NeMA maps.

5.2. State regulation theory/cognitive-energetic model (CEM)

"The cognitive-energetic model (CEM) proposes that the overall efficiency of information processing is determined by the interplay of three levels: computational mechanisms of attention, state factors, and management/executive function" (Sergeant, 2005, p. 1248). These all require optimal arousal/activation based on energetic pools, but does not identify what those are. The theory cites differential effectiveness of presentation rate, but does not predict or explain those effects. In some studies reinforcement contingencies are more effective for ADHD than for TD; that is interpreted not as a ceiling effect for TD, but rather as evidence that energy had not been correctly allocated in ADHD. It suggests that impulsivity and hyperactivity may be attempts to increase stimulation, moving individuals toward an optimal arousal level. Many of the shortcomings of CEM are resolved if the energetic pools are understood to be glycogen or lactate available to the neurons (Killeen et al., 2013).

5.3. The Dual-Pathway Model (DPM)

The DPM (Sonuga-Barke, 2005) is a combination of two hypotheses: delay aversion and inhibition failure. Delay aversion treats ADHD as a motivational style driven by distaste for delays of reinforcement, with attempts to avoid such delays generating some of the concomitant signs of the condition, such as impulsivity and hyperactivity. Delay aversion arises from disturbances in certain brain/reward centers. Inhibition failure, or "impulsive drive for immediate rewards" is an additional EF factor in impulsive choice, added in light of the failure of equalization of delays to completely remediate performance. Given choice between a delay of 2 s for 1 point (SS), and a delay of 30 s for 2 points (LL), both factors will move ADHD to prefer the short delay more than TD do. If a post reward delay of 28 s is added to the SS, thereby equating waiting times, preference should become the same as TD. They move in that direction but fall short, because ADHD are said to also have an impulsive drive for immediate rewards. Delay aversion is not always found (Sjowall et al., 2013), but was found in a large, multinational study (Marco et al., 2009): Without the post reward delay, ADHD more than TD maximized the overall rate of reinforcement by their greater preference for SS, and with the delay, reduced the extremity of that preference, but not to the level of TD. It should be noted that this constellation of choice patterns was available in the literature before the large test; they were not predicted from basic principles, but rather elevated to basic principles as fundamental characteristics of ADHD. Neither NeT nor NeMA (nor DPT) predicted the well-documented bias for SS shown by ADHD (see the meta-analysis of Patros et al., 2016). Delay aversion predicts that ADHD will find delays more aversive than TD, which has been found not to be the case (Scheres et al., 2013). It predicts that post-reward delay would normalize performance, a partially successful prediction. NeT predicted neither, but can rationalize the standard effect: Preferring the LL requires that goal be kept in mind over extended periods, which is something that undernourished neural hubs cannot do well, as they are quickly depleted. If this is the case stimulants should attenuate the discounting, which it does (Shiels et al., 2009). If the delays are merely hypothetical, there should be no difference, as has been found to be the case (e.g., Scheres et al., 2008; Wilson et al., 2011), and stimulants should have little effect, as sustained attention is not required; as demonstrated (Shiels et al., 2009). The post-reward delay gives some time out for repletion. DPM does not address RT distribution differences.

5.4. The dynamic developmental theory (DDT)

DDT (Sagvolden et al., 2005) claims that two factors underly ADHD: a steeper delay of reinforcement gradient, for which there is some evidence in a rat model of ADHD (e., g. Johansen et al., 2009b, Fig. 4); and an extinction deficit—slower cessation of responding when reward is withheld—which has also been found in an animal model (Johansen and Sagvolden, 2004). As Johnson et al. (2009) conclude: "The DDT is grounded in a well-articulated, scientific framework but needs to be extended further into the human experimental setting" (e.g., Johansen, Killeen, Russell, et al., 2009). NeT does not predict an extinction deficit. It did not predict a steeper delay of reinforcement gradient, but might have: Maintaining stimuli and responses in memory during a delay requires sustained activation of relevant neural centers, and that is the very activity that is compromised in the neuroenergetic theory. NeT and NeMA may be the best ways of further extending DDT into the human experimental setting.

6. Comparison of NeMA with other models

NeT predicts slowed processing of information due to insufficient sustained nourishment of active neurons. As realized in RT studies, the best way of implementing this prediction is with a DDM model, which is a gold-standard in modelling decision problems. The one crafted for this purpose is a Wald-type process with the possibility of attentional lapses, NeMA, detailed in the appendix. NeMA omits a non-decisional time to begin processing the stimuli and begin preparing a motor response. Models that include this parameter generally show that this process is faster for ADHD. It omits a bias (to respond one way or another) which means that individual differences in bias will essentially add noise to the precision of parameter estimates with NeMA. However, there is typically little systematic differences in bias between ADHD and TD. The inclusion of the attentional lapses configures the model as a mixture of simple Wald distributions and ex-Wald distributions, with slightly heavier tails than the pure Wald. A head-to-head comparison of

this and the more standard DDMs has not been made. Models are ways of connecting theories to data, and most of the DDMs can do an adequate job of that for NeT. The omission of attentional lapses, which are common in the population as a whole as well as ADHD (Smallwood and Schooler, 2015), may hobble those that do not build them in. Application of a DDM to trial-by-trial data from individual subjects is a more secure process than analyses of average distributions (Matzke and Wagenmakers, 2009). For a review of DDM models for ADHD, see Siegler et al. (2016).

7. Correlated brain mechanisms

As Karalunas et al. (2014, P. 15) prudently decided, "Given the vastness of the relevant literature that could be used to make circumstantial arguments related to brain mechanisms of RT variability, the wealth of existing reviews on ADHD and brain imaging generally, and the dearth of literature directly examining ADHD, the brain, and RT variability at the same time, we do not attempt comprehensive review of all brain findings related to RT variability." Nonetheless, after an in-depth study of a limited ambit of research, their conclusions were qualified, interpretations "not straightforward", conclusions in need of "careful interpretation", and "interactions poorly understood". Despite editorial encouragement, I venture not there. NeT is a distributed processing model (Anderson, 2010; Bartholomew et al., 2009; van der Maas et al., 2006). All NeT predicts is that the demand for lactate, as measured by BOLD, will not be so concentrated, but rather will be distributed over multiple hubs that are engaged to compensate while the focal ones replete. This has been found: "the variability in the size of the BOLD signal across individuals was significantly higher for the ADHD group than for the control group, and variability across the time series in individuals with ADHD was linked to symptom severity and behavioral performance" (Depue et al., 2010. p. 96; O'Halloran et al., 2018). For another review of DDM models for ADHD, see Ziegler et al. (2016).

8. Discussion

This paper introduced a neurobehavioral theory of the developmental disorder ADHD. It focused on the material cause—the neurophysiology—and formal models of it. Other kinds of explanations are available (Killeen et al., 2012). There is good evidence both in the original presentation of the theory (Russell et al., 2006; Killeen et al., 2013), and in independent tests of it, that NeT continues to provide a viable explanation of many of the qualitative and quantitative facts surrounding ADHD. As Weigard and Huang-Pollock note (2017), DDMs such as NeMA have the advantage of bypassing cognitive explanations involving homunculi—executive controllers, librarians, and such. This is a goal not only of behavioral psychologists, but, perhaps surprisingly, of cognitive ones as well: "A major item on the agenda of cognitive psychology is to banish the homunculus (i.e., the assumption of an intelligent agent residing elsewhere in the system, usually off stage, who does all the marvelous things that need to be done actually to generate the total behavior of the subject)" (Newell, 1980 p. 175, cited in Monsell and Driver, 2000, who see progress being made). In NeT, the prediction of slower processing speeds based on insufficient energization of the fatigued neural hubs is hypothetical; but has been productive of many validated predictions, and suggestions for therapy (Killeen et al., 2013), without the help of mediational homunculi.

There are several good choices of animal models of symptoms of ADHD (e.g., Sagvolden et al., 2009; Russell, 2011; Alsop, 2007; Sutherland et al., 2009; Sontag et al., 2010). Experimental paradigms such as delay discounting (Wickens et al., 2011) and fixed minimum interval (Watterson et al., 2015) can link the behavior of humans and rodent models. Several versions of DDMs are available online. We know that SHR rats are more hyperactive than control strains: their time on task is shorter than controls, with more frequent digressions away from responding (Íbias et al., 2015). Like ADHD individuals, the delay discounting of SHR is often steeper than controls (see the review and reanalysis in Killeen, 2015). Atomoxetine, a noradrenaline-specific reuptake inhibitor, improves the performance of rats with three types of impulsive behavior (Robinson et al., 2008). It is reasonable to proceed with analysis.

There have been many applications of DDMs to animal experiments. The typical animals are monkeys, and DDMs are typically used in coordination with neural recordings and models (e.g., Wang, 2008; Cassey et al., 2014; Hanes and Schall, 1996); for reviews, see (Gold and Shadlen, 2007; Forstmann et al., 2016). Cassey et al. (Cassey et al., 2016) offer a hierarchical Bayesian approach to joint analysis of neural and behavioral data; for a useful recent review of such attempts at unification, see (O'Connell et al., 2018). Luzardo and associates (Luzardo et al., 2017) have developed a DDM for use with interval timing, and shown how it can predict the results of peak procedures (also see Balcı and Simen, 2016). In a paradigm using rats in a standard experimental chamber, Scott et al. (Scott et al., 2015) had rats choose the dipper on the side where more LED flashes had occurred. They found that variability increased with flash number as Weber's law, more slowly than predicted by DDM. Hales et al. (Hales et al., 2016) used DDMs to study the effect of anxiety on rats' biases in ambiguous cue decision tasks. Brunton et al. (2013) showed that much of the noise in rats decisional processing arose from the stimuli, not the decision process. Finally, in a paradigm easily adopted to rat models of ADHD, Hales et al. (2017) used DDMs to study decisional effects and antidepressants. DDMs have even been applied to mate choice in insects. Female crickets are a critical audience for male's songs. The DDM showed that imprecise elements were weighted much more heavily than good ones: a few slurred notes and the courtship is over (Clemens et al., 2014). Yet, I have found no evidence for the deployment of DDMs to study animal models of ADHD. Until this challenge is engaged, many important questions will remain unanswered. Is the information processing of animal models of ADHD slower, and increasingly so as sessions are prolonged? Are they differentially worse than controls on continuous performance tasks, and is this alleviated by long ITIs? Do norepinephrine reuptake inhibitors speed information processing in animal models of ADHD? We don't know.

Appendix A

The Wald (Inverse-Gaussian) density is:

$$f(t) = \frac{c}{\sqrt{2\pi t^3}} e^{-\frac{(c-vt)^2}{2t}}$$

To render the parameters in a form more intuitive for present purposes, the traditional parameters (e.g., Evans et al., 1993) have been written as $c = \sqrt{\lambda}$ and $v = \sqrt{\lambda/\mu}$. For brief trials where attentional lapses are unlikely, this model can provide adequate fits to data.

In experimental paradigms where sustained attention is called for, the NeMA model (the Neuroenergetic Model with Attentional lapses) shown in Fig. 4 is necessary. This adds an exponentially distributed return time to the Wald. The model was developed by Schwarz (2001), which one may consult for the equation; it is too elaborate to be usefully given here. There are programs that fit it to RT data (Heathcote, 2004; Wagenmakers et al., 2007), but it may be the case that a simple shifted Wald (e.g., Anders et al., 2016; Heathcote et al., 2004) will suffice. This offsets the origin, and for

paradigms with conspicuous stimuli, assumes that attention gets pulled back to the target with a fixed latency. One then simply averages the density functions of a Wald and one with mean shifted by τ, with weights of p(A) and 1 - p(A). See (Matzke et al., 2009) for cautions in interpreting the parameters of these distributions; it is safer to fit one of the DDM models to individual data directly.

The mean of the RT given by NeMA is

$$\mu = \frac{c}{v} + p(\sim A)\tau^2$$

Where p(˜A) is the probability of being in the inattentive state, and τ (tau) is the time required to return to task. It holds for both the shifted-mean- and the ex-Wald. It gives the mean for the simple Wald by setting τ to 0.

The probability of being in one attentional state or another is:

$$p = \begin{vmatrix} 1-\lambda & \lambda \\ \alpha & 1-\alpha \end{vmatrix}$$

Here the first row and column correspond to the A state, and the second to ˜A. If one starts in the A state, the probability of moving to the ˜A state is λ. If in the ˜A state, the probability of staying there is 1 – α.

The probability of being in a state after t time steps is:

$$\mathrm{p_t} = \mathrm{p^t}$$

At equilibrium (after long t) the probability of attending is:

$$p(A)_\infty = \frac{\alpha}{\alpha + \lambda}$$

If alpha is small, but conspicuous target stimuli snap attention back to the task with a small (possibly fixed) delay, then the probability of being in the attentive state is:

$$p(A)_t = (1-\lambda)^t \approx \sim e^{-\lambda t}$$

This was the assumption in fitting the data in Fig. 5.

The variance of the Wald and the NeMA model using the shifted-Wald is

$$\sigma^2 = cv^{-3}$$

The variance of the NeMA model using the ex-Wald is

$$\sigma^2 = cv^{-3} + p(\sim A)\tau^2$$

References

Alsop, B., 2007. Problems with spontaneously hypertensive rats (SHR) as a model of attention-deficit/hyperactivity disorder (AD/HD). J. Neurosci. Methods 162, 42–48. https://doi.org/10.1016/j.jneumeth.2006.12.002.

Anders, R., Alario, F., Van Maanen, L., 2016. The shifted Wald distribution for response time data analysis. Psychol. Methods 21, 309. https://doi.org/10.1037/met0000066.

Anderson, M.L., 2010. Neural reuse: a fundamental organizational principle of the brain. Behav. Brain Sci. 33, 245–266. https://doi.org/10.1017/S0140525X10000853. discussion 266–313.

Balcı, F., Simen, P., 2016. A decision model of timing. Curr. Opin. Behav. Sci. 8, 94–101. https://doi.org/10.1016/j.cobeha.2016.02.002.

Bartholomew, D.J., Deary, I.J., Lawn, M., 2009. A new lease of life for Thomson's bonds model of intelligence. Psychol. Rev. 116, 567–579. https://doi.org/10.1037/a0016262.

Beauchaine, T.P., Gatzke-Kopp, L., Neuhaus, E., Chipman, J., Reid, M.J., Webster-Stratton, C., 2013. Sympathetic- and parasympathetic-linked cardiac function and prediction of externalizing behavior, emotion regulation, and prosocial behavior among preschoolers treated for ADHD. J. Consult. Clin. Psychol. 81, 481–493. https://doi.org/10.1037/a0032302.

Borger, N., van der Meere, J., Ronner, A., Alberts, E., Geuze, R., Bogte, H., 1999. Heart rate variability and sustained attention in ADHD children. J. Abnorm. Child Psychol. 27, 25–33. https://doi.org/10.1023/A:1022610306984.

Brunton, B.W., Botvinick, M.M., Brody, C.D., 2013. Rats and humans can optimally accumulate evidence for decision-making. Science 340, 95–98. https://doi.org/10.1126/science.1233912.

Cassey, P., Heathcote, A., Brown, S.D., 2014. Brain and behavior in decision-making. PLoS Comput. Biol. 10, e1003700. https://doi.org/10.1371/journal.pcbi.1003700.

Cassey, P.J., Gaut, G., Steyvers, M., Brown, S.D., 2016. A generative joint model for spike trains and saccades during perceptual decision-making. Psychon. Bull. Rev. 23, 1757–1778. https://doi.org/10.3758/s13423-016-1056-z.

Castellanos, F.X., Sonuga-Barke, E.J., Scheres, A., Di Martino, A., Hyde, C., Walters, J.R., 2005. Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. Biol. Psychiatry 57, 1416–1423. https://doi.org/10.1016/j.biopsych.2004.12.005.

Catania, A.C., 2005. Attention-deficit/hyperactivity disorder (ADHD): delay-of-reinforcement gradients and other behavioral mechanisms. Behav. Brain Sci. 28, 419–424. https://doi.org/10.1017/S0140525X05220071.

Chhikara, R.S., Folks, L., 1989. The Inverse Gaussian Distribution: Theory, Methodology, and Applications. CRC, Boca Raton, FL.

Clemens, J., Kramer, S., Ronacher, B., 2014. Asymmetrical integration of sensory information during mating decisions in grasshoppers. Proc. Natl. Acad. Sci. U. S. A. 111, 16562–16567. https://doi.org/10.1073/pnas.1412741111.

Cloutier, M., Bolger, F.B., Lowry, J.P., Wellstead, P., 2009. An integrative dynamic model of brain energy metabolism using in vivo neurochemical measurements. J. Comput. Neurosci. 27, 391–414. https://doi.org/10.1007/s10827-009-0152-8.

Coghill, D., Sonuga-Barke, E.J., 2012. Annual research review: categories versus dimensions in the classification and conceptualisation of child and adolescent mental disorders–implications of recent empirical study. J. Child Psychol. Psychiatry 53, 469–489. https://doi.org/10.1111/j.1469-7610.2011.02511.x.

Cortese, S., 2012. The neurobiology and genetics of attention-deficit/hyperactivity disorder (ADHD): what every clinician should know. Eur. J. Paediatr. Neurol. 16, 422–433. https://doi.org/10.1016/j.ejpn.2012.01.009.

Cowell, R.A., Bussey, T.J., Saksida, L.M., 2012. Empiricists are from Venus, modelers are from Mars: reconciling experimental and computational approaches in cognitive neuroscience. Neurosci. Biobehav. Rev. 36, 2371–2379. https://doi.org/10.1016/j.neubiorev.2012.08.008.

Dalby, J.T., Kinsbourne, M., Swanson, J.M., 1989. Self-paced learning in children with attention deficit disorder with hyperactivity. J. Abnorm. Child Psychol. 17, 269–275. https://doi.org/10.1007/BF00917398.

Depue, B.E., Burgess, G.C., Willcutt, E.G., Bidwell, L.C., Ruzic, L., Banich, M.T., 2010. Symptom-correlated brain regions in young adults with combined-type ADHD: their organization, variability, and relation to behavioral performance. Psychiatry Res. Neuroimaging 182, 96–102. https://doi.org/10.1016/j.pscychresns.2009.11.011.

Epstein, J.N., Keith Conners, C., Hervey, A.S., Tonev, S.T., Eugene Arnold, L., Abikoff, H.B., Elliott, G., Greenhill, L.L., Hechtman, L., Hoagwood, K., 2006. Assessing medication effects in the MTA study using neuropsychological outcomes. J. Child Psychol. Psychiatry 47, 446–456. https://doi.org/10.1111/j.1469-7610.2005.01469.x.

Evans, M., Hastings, N., Peacock, B., 1993. Statistical Distributions. Wiley, NY.

Fliers, E., Rommelse, N., Vermeulen, S., Altink, M., Buschgens, C., Faraone, S., Sergeant, J., Franke, B., Buitelaar, J., 2008. Motor coordination problems in children and adolescents with ADHD rated by parents and teachers: effects of age and gender. J. Neural Trans. 115, 211–220. https://doi.org/10.1007/s00702-007-0827-0.

Forstmann, B.U., Ratcliff, R., Wagenmakers, E.J., 2016. Sequential sampling models in cognitive neuroscience: advantages, applications, and extensions. Annu. Rev. Psychol. 67, 641–666. https://doi.org/10.1146/annurev-psych-122414-033645.

Fosco, W.D., White, C.N., Hawk Jr., L.W., 2017. Acute stimulant treatment and reinforcement increase the speed of information accumulation in children with ADHD. J. Abnorm. Child Psychol. 45, 911–920. https://doi.org/10.1007/s10802-016-0222-0.

Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision making. Annu. Rev. Neurosci. 30, 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038.

Hales, C.A., Robinson, E.S., Houghton, C.J., 2016. Diffusion modelling reveals the decision making processes underlying negative judgement Bias in rats. PLoS One 11, e0152592. https://doi.org/10.1371/journal.pone.0152592.

Hales, C.A., Houghton, C.J., Robinson, E.S., 2017. Behavioural and computational methods reveal differential effects for how delayed and rapid onset antidepressants effect decision making in rats. Eur. Neuropsychopharmacol. 27, 1268–1280. https://doi.org/10.1016/j.euroneuro.2017.09.008.
Hanes, D.P., Schall, J.D., 1996. Neural control of voluntary movement initiation. Science 274, 427–430. https://doi.org/10.1126/science.274.5286.427.
Hartman, C.A., Rhee, S.H., Willcutt, E.G., Pennington, B.F., 2007. Modeling rater disagreement for ADHD: are parents or teachers biased? J. Abnorm. Child Psychol. 35, 536–542. https://doi.org/10.1007/s10802-007-9110-y.
Heathcote, A., 2004. Fitting Wald and ex-Wald distributions to response time data: an example using functions for the S-PLUS package. Behav. Res. Methods Instrum. Comput. 36, 678–694. https://doi.org/10.3758/BF03195403.
Heathcote, A., Brown, S., Cousineau, D., 2004. QMPE: estimating Lognormal, Wald, and Weibull RT distributions with a parameter-dependent lower bound. Behav. Res. Methods Instrm. Comput. 36, 277–290. https://doi.org/10.3758/BF03195574.
Hervey, A.S., Epstein, J.N., Curry, J.F., Tonev, S., Eugene Arnold, L., Keith Conners, C., Hinshaw, S.P., Swanson, J.M., Hechtman, L., 2006. Reaction time distribution analysis of neuropsychological performance in an ADHD sample. Child Neuropsychol. 12, 125–140. https://doi.org/10.1080/09297040500499081.
Huang-Pollock, C.L., Karalunas, S.L., 2010. Working memory demands impair skill acquisition in children with ADHD. J. Abnorm. Psychol. 119, 174–185. https://doi.org/10.1037/a0017862.
Huang-Pollock, C.L., Nigg, J.T., Halperin, J.M., 2006. Single dissociation findings of ADHD deficits in vigilance but not anterior or posterior attention systems. Neuropsychology 20, 420. https://doi.org/10.1037/0894-4105.20.4.420.
Huang-Pollock, C.L., Karalunas, S.L., Tam, H., Moore, A.N., 2012. Evaluating vigilance deficits in ADHD: a meta-analysis of CPT performance. J. Abnorm. Psychol. 121, 360–371. https://doi.org/10.1037/a0027205.
Huang-Pollock, C., Ratcliff, R., McKoon, G., Shapiro, Z., Weigard, A., Galloway-Long, H., 2017. Using the diffusion model to explain cognitive deficits in attention deficit hyperactivity disorder. J. Abnorm. Child Psychol. 45, 57–68. https://doi.org/10.1007/s10802-016-0151-y.
Íbias, J., Pellón, R., Sanabria, F., 2015. A microstructural analysis of schedule-induced polydipsia reveals incentive-induced hyperactivity in an animal model of ADHD. Behav. Brain Res. 278, 417–423. https://doi.org/10.1016/j.bbr.2014.10.022.
Jensen, P.S., Garcia, J.A., Glied, S., Crowe, M., Foster, M., Schlander, M., Hinshaw, S., Vitiello, B., Arnold, L.E., Elliott, G., Hechtman, L., Newcorn, J.H., Pelham, W.E., Swanson, J., Wells, K., 2005. Cost-effectiveness of ADHD treatments: findings from the multimodal treatment study of children with ADHD. Am. J. Psychiatry 162, 1628–1636. https://doi.org/10.1176/appi.ajp.162.9.1628.
Johansen, E.B., Sagvolden, T., 2004. Response disinhibition may be explained as an extinction deficit in an animal model of attention-deficit/hyperactivity disorder (ADHD). Behav. Brain Res. 149, 183–196. https://doi.org/10.1016/S0166-4328(03)00229-8.
Johansen, E.B., Killeen, P.R., Russell, V.A., Tripp, G., Wickens, J.R., Tannock, R., Williams, J., Sagvolden, T., 2009a. Origins of altered reinforcement effects in ADHD. Behav. Brain Funct. 5, 7. https://doi.org/10.1186/1744-9081-5-7.
Johansen, E.B., Killeen, P.R., Sagvolden, T., 2009b. Behavioral variability, elimination of responses, and delay-of-reinforcement gradients in SHR and WKY rats. Behav. Brain Funct. 3. https://doi.org/10.1186/1744-9081-3-60.
Johnson, K.A., Wiersema, J.R., Kuntsi, J., 2009. What would Karl Popper say? Are current psychological theories of ADHD falsifiable? Behav. Brain Funct. 5 (15). https://doi.org/10.1186/1744-9081-5-15.
Kaplan, B.J., Rucklidge, J.J., Romijn, A., McLeod, K., 2015. The emerging field of nutritional mental health. Clin. Psychol. Sci. 3, 964–980. https://doi.org/10.1177/2167702614555413.
Karalunas, S.L., Huang-Pollock, C.L., 2013. Integrating impairments in reaction time and executive function using a diffusion model framework. J. Abnorm. Child Psychol. 41, 837–850. https://doi.org/10.1007/s10802-013-9715-2.
Karalunas, S.L., Huang-Pollock, C.L., Nigg, J.T., 2012. Decomposing attention-deficit/hyperactivity disorder (ADHD)-related effects in response speed and variability. Neuropsychology 26, 684–694. https://doi.org/10.1037/a0029936.
Karalunas, S.L., Huang-Pollock, C.L., Nigg, J.T., 2013. Is reaction time variability in ADHD mainly at low frequencies? J. Child Psychol. Psychiatry 54, 536–544. https://doi.org/10.1111/jcpp.12028.
Karalunas, S.L., Geurts, H.M., Konrad, K., Bender, S., Nigg, J.T., 2014. Annual Research Review: reaction time variability in ADHD and autism spectrum disorders: measurement and mechanisms of a proposed trans-diagnostic phenotype. J. Child Psychol. Psychiatry 55, 685–710. https://doi.org/10.1111/jcpp.12217.
Killeen, P.R., 2013. Absent without leave; a neuroenergetic theory of mind wandering. Front. Psychol. 4, 373–387. https://doi.org/10.3389/fpsyg.2013.00373.
Killeen, P.R., 2015. Models of ADHD: five ways smaller sooner is better. J. Neurosci. Methods 252, 2–13. https://doi.org/10.1016/j.jneumeth.2015.01.011.
Killeen, P.R., Tannock, R., Sagvolden, T., 2012. The Four causes of ADHD: a framework. In: Stanford, S.C., Tannock, R. (Eds.), Behavioral Neuroscience of Attention Deficit Hyperactivity Disorder and Its Treatment. Springer-Verlag, Berlin.
Killeen, P.R., Russell, V.A., Sergeant, J.A., 2013. A behavioral neuroenergetics theory of ADHD. Neurosci. Biobehav. Rev. 37, 625–657. https://doi.org/10.1016/j.neubiorev.2013.02.011.
Killeen, P.R., Russell, V.A., Tannock, R., 2016. Neuroenergetics. Curr. Direct. Psychol. Sci. 25, 124–129. https://doi.org/10.1177/0963721416628530.
Killingsworth, M.A., Gilbert, D.T., 2010. A wandering mind is an unhappy mind. Science 330https://doi.org/10.1126/science.1192439. 932–932.
Kiray, H., Lindsay, S.L., Hosseinzadeh, S., Barnett, S.C., 2016. The multifaceted role of astrocytes in regulating myelination. Exp. Neurol. 283, 541–549. https://doi.org/10.1016/j.expneurol.2016.03.009.
Klein, M., Onnink, M., van Donkelaar, M., Wolfers, T., Harich, B., Shi, Y., Dammers, J., Arias-Vásquez, A., Hoogman, M., Franke, B., 2017. Brain imaging genetics in ADHD and beyond–mapping pathways from gene to disorder at different levels of complexity. Neurosci. Biobehav. Rev. 80, 115–155. https://doi.org/10.1016/j.neubiorev.2017.01.013.
Kofler, M.J., Rapport, M.D., Sarver, D.E., Raiker, J.S., Orban, S.A., Friedman, L.M., Kolomeyer, E.G., 2013. Reaction time variability in ADHD: a meta-analytic review of 319 studies. Clin. Psychol. Rev. 33, 795–811. https://doi.org/10.1016/j.cpr.2013.06.001.
Koschack, J., Kunert, H.J., Derichs, G., Weniger, G., Irle, E., 2003. Impaired and enhanced attentional function in children with attention deficit/hyperactivity disorder. Psychol. Med. 33, 481–489. https://doi.org/10.1017/S0033291702007067.
Kramer, P., Bressan, P., 2018. Our (mother's) mitochondria and our mind. Perspect. Psychol. Sci. 13, 88–100. https://doi.org/10.1177/1745691617718356.
Lea, S.E., Matt Alderson, R., Patros, C.H.G., Tarle, S.J., Arrington, E.F., Grant, D.M., 2018. Working memory and motor activity: a comparison across attention-deficit/hyperactivity disorder, generalized anxiety disorder, and healthy control groups. Behav. Ther. 49, 419–434. https://doi.org/10.1016/j.beth.2017.08.009.
Lee, K.S., Osborne, R.E., Hayes, K.A., Simoes, R.A., 2008. The effects of pacing on the academic testing performance of college students with ADHD: a mixed methods study. J. Educ. Comput. Res. 39, 123–141. https://doi.org/10.2190/EC.39.2.b.
Leth-Steensen, C., King-Elbaz, Z., Douglas, V.I., 2000. Mean response times, variability, and skew in the responding of ADHD children: a response time distributional approach. Acta Psychol. (Amst) 104, 167–190 0.1016/S0001-6918(00)00019-6.
Levy, F., Hay, D., McLaughlin, M., Wood, C., Waldman, I., 1996. Twin sibling differences in parental reports of ADHD, speech, reading and behaviour problems. J. Child Psychol. Psychiatry 37, 569–578. https://doi.org/10.1111/j.1469-7610.1996.tb01443.x.
Losier, B.J., McGrath, P.J., Klein, R.M., 1996. Error patterns on the continuous performance test in non-medicated and medicated samples of children with and without ADHD: a meta-analytic review. J. Child Psychol. Psychiatry 37, 971–987. https://doi.org/10.1111/j.1469-7610.1996.tb01494.x.
Luzardo, A., Rivest, F., Alonso, E., Ludvig, E.A., 2017. A drift–diffusion model of interval timing in the peak procedure. J. Math. Psychol. 77, 111–123. https://doi.org/10.1016/j.jmp.2016.10.002.
Magistretti, P.J., Allaman, I., 2018. Lactate in the brain: from metabolic end-product to signalling molecule. Nat. Rev. Neurosci. 19, 235–249. https://doi.org/10.1038/nrn.2018.19.
Marco, R., Miranda, A., Schlotz, W., Melia, A., Mulligan, A., Muller, U., Andreou, P., Butler, L., Christiansen, H., Gabriels, I., Medad, S., Albrecht, B., Uebel, H., Asherson, P., Banaschewski, T., Gill, M., Kuntsi, J., Mulas, F., Oades, R., Roeyers, H., Steinhausen, H.C., Rothenberger, A., Faraone, S.V., Sonuga-Barke, E.J., 2009. Delay and reward choice in ADHD: an experimental test of the role of delay aversion. Neuropsychology 23, 367–380. https://doi.org/10.1037/a0014914.
Margineanu, M.B., Mahmood, H., Fiumelli, H., Magistretti, P.J., 2018. L-lactate regulates the expression of synaptic plasticity and neuroprotection genes in cortical neurons: a transcriptome analysis. Front. Mol. Neurosci. 11, 375–382. https://doi.org/10.3389/fnmol.2018.00375.
Matzke, D., Wagenmakers, E.J., 2009. Psychological interpretation of the ex-Gaussian and shifted Wald parameters: a diffusion model analysis. Psychon. Bull. Rev. 16, 798–817. https://doi.org/10.3758/PBR.16.5.798.
Medin, T., Medin, H., Hefte, M.B., Storm-Mathisen, J., Bergersen, L.H., 2018. Upregulation of the lactate transporter monocarboxylate transporter 1 at the blood-brain barrier in a rat model of attention-deficit/hyperactivity disorder suggests hyperactivity could be a form of self-treatment. Behav. Brain Res. https://doi.org/10.1016/j.bbr.2018.12.023.
Metin, B., Roeyers, H., Wiersema, J.R., van der Meere, J.J., Thompson, M., Sonuga-Barke, E., 2013. ADHD performance reflects inefficient but not impulsive information processing: a diffusion model analysis. Neuropsychology 27, 193. https://doi.org/10.1037/a0031533.
Mitsis, E.M., McKay, K.E., Schulz, K.P., Newcorn, J.H., Halperin, J.M., 2000. Parent-teacher concordance for DSM-IV attention-deficit/hyperactivity disorder in a clinic-referred sample. J. Am. Acad. Child Adolesc. Psychiatry 39, 308–313. https://doi.org/10.1097/00004583-200003000-00012.
Monsell, S., Driver, J., 2000. Banishing the Control Homunculus. Control of Cognitive Processes: Attention and Performance XVIII. pp. 3–32.
Mulder, M.J., Bos, D., Weusten, J.M.H., van Belle, J., van Dijk, S.C., Simen, P., van Engeland, H., Durston, S., 2010. Basic impairments in regulating the speed-accuracy tradeoff predict symptoms of attention-deficit/hyperactivity disorder. Biol. Psychiatry 68, 1114–1119. https://doi.org/10.1016/j.biopsych.2010.07.031.
O'Connell, R.G., Shadlen, M.N., Wong-Lin, K., Kelly, S.P., 2018. Bridging neural and computational viewpoints on perceptual decision-making. Trends Neurosci. 41, 838–852. https://doi.org/10.1016/j.tins.2018.06.005.
O'Halloran, L., Cao, Z., Ruddy, K., Jollans, L., Albaugh, M.D., Aleni, A., Potter, A.S., Vahey, N., Banaschewski, T., Hohmann, S., Bokde, A.L.W., Bromberg, U., Buchel, C., Quinlan, E.B., Desrivieres, S., Flor, H., Frouin, V., Gowland, P., Heinz, A., Ittermann, B., Nees, F., Orfanos, D.P., Paus, T., Smolka, M.N., Walter, H., Schumann, G., Garavan, H., Kelly, C., Whelan, R., 2018. Neural circuitry underlying sustained attention in healthy adolescents and in ADHD symptomatology. NeuroImage 169, 395–406. https://doi.org/10.1016/j.neuroimage.2017.12.030.
Patros, C.H., Alderson, R.M., Kasper, L.J., Tarle, S.J., Lea, S.E., Hudec, K.L., 2016. Choice-impulsivity in children and adolescents with attention-deficit/hyperactivity disorder (ADHD): a meta-analytic review. Clin. Psychol. Rev. 43, 162–174. https://doi.org/10.1016/j.cpr.2015.11.001.
Pellerin, L., Magistretti, P.J., 2011. Sweet sixteen for ANLS. J. Cereb. Blood Flow Metab. 32, 1152–1166. https://doi.org/10.1038/jcbfm.2011.149.
Pellerin, L., Bouzier-Sore, A.K., Aubert, A., Serres, S., Merle, M., Costalat, R., Magistretti, P.J., 2007. Activity-dependent regulation of energy metabolism by astrocytes: an update. Glia 55, 1251–1262. https://doi.org/10.1002/glia.20528.
Pellón, R., Íbias, J., Killeen, P.R., 2018. Delay gradients for spout-licking and magazine-entering induced by a periodic food schedule. Psychol. Rec. 68, 151–162. https://doi.org/10.1007/s40732-018-0275-2.
Pinares-Garcia, P., Stratikopoulos, M., Zagato, A., Loke, H., Lee, J., 2018. Sex: a significant risk factor for neurodevelopmental and neurodegenerative disorders. Brain

Sci. 8, 154. https://doi.org/10.3390/brainsci8080154.
Polanczyk, G., de Lima, M.S., Horta, B.L., Biederman, J., Rohde, L.A., 2007. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. Am. J. Psychiatry 164, 942–948. https://doi.org/10.1176/ajp.2007.164.6.942.
Polanczyk, G.V., Willcutt, E.G., Salum, G.A., Kieling, C., Rohde, L.A., 2014. ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. Int. J. Epidemiol. 43, 434–442. https://doi.org/10.1093/ije/dyt261.
Querne, L., Berquin, P., 2009. Distinct response time distributions in attention deficit hyperactivity disorder subtypes. J. Attent. Disord. 13, 66–77. https://doi.org/10.1177/1087054708323006.
Rapport, M.D., Kofler, M.J., Alderson, R.M., Timko Jr., T.M., Dupaul, G.J., 2009. Variability of attention processes in ADHD: observations from the classroom. J. Atten. Disord. 12, 563–573. https://doi.org/10.1177/1087054708322990.
Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. 20, 873–922. https://doi.org/10.1162/neco.2008.12-06-420.
Ratcliff, R., Smith, P.L., McKoon, G., 2015. Modeling response time and accuracy data. Curr. Direct. Psychol. Sci. 24, 458–470. https://doi.org/10.1177/0963721415596228.
Ratcliff, R., Smith, P.L., Brown, S.D., McKoon, G., 2016. Diffusion decision model: current issues and history. Trends Cogn. Sci. 20, 260–281. https://doi.org/10.1016/j.tics.2016.01.007.
Robinson, E.S., Eagle, D.M., Mar, A.C., Bari, A., Banerjee, G., Jiang, X., Dalley, J.W., Robbins, T.W., 2008. Similar effects of the selective noradrenaline reuptake inhibitor atomoxetine on three distinct forms of impulsivity in the rat. Neuropsychopharmacology 33, 1028–1037. https://doi.org/10.1038/sj.npp.1301487.
Rommelse, N.N., Altink, M.E., de Sonneville, L.M., Buschgens, C.J., Buitelaar, J., Oosterlaan, J., Sergeant, J.A., 2007. Are motor inhibition and cognitive flexibility dead ends in ADHD? J. Abnorm. Child Psychol. 35, 957–967. https://doi.org/10.1007/s10802-007-9146-z.
Russell, V.A., 2011. Overview of animal models of attention deficit hyperactivity disorder (ADHD). Curr. Protoc. Neurosci. https://doi.org/10.1002/0471142301.ns0935s54. Chapter 9: Unit9 35.
Russell, V.A., Oades, R.D., Tannock, R., Killeen, P.R., Auerbach, J.G., Johansen, E.B., Sagvolden, T., 2006. Response variability in Attention-Deficit/Hyperactivity Disorder: a neuronal and glial energetics hypothesis. Behav. Brain Funct. 2, 25. https://doi.org/10.1186/1744-9081-2-30.
Sagvolden, T., Johansen, E.B., Aase, H., Russell, V.A., 2005. A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) predominantly hyperactive/impulsive and combined subtypes. Behav. Brain Sci. 28, 397–419. https://doi.org/10.1017/S0140525X05000075. discussion 419-468.
Sagvolden, T., Johansen, E.B., Wøien, G., Walaas, S.I., Storm-Mathisen, J., Bergersen, L.H., Hvalby, Ø., Jensen, V., Aase, H., Russell, V.A., Killeen, P.R., DasBanerjee, T., Middleton, F.A., Faraone, S.V., 2009. The spontaneously hypertensive rat model of ADHD–The importance of selecting the appropriate reference strain. Neuropharmacology 57, 619–626. https://doi.org/10.1016/j.neuropharm.2009.08.004.
Salmi, J., Salmela, V., Salo, E., Mikkola, K., Leppamaki, S., Tani, P., Hokkanen, L., Laasonen, M., Numminen, J., Alho, K., 2018. Out of focus - brain attention control deficits in adult ADHD. Brain Res. 1692, 12–22. https://doi.org/10.1016/j.brainres.2018.04.019.
Sarver, D.E., Rapport, M.D., Kofler, M.J., Raiker, J.S., Friedman, L.M., 2015. Hyperactivity in attention-deficit/hyperactivity disorder (ADHD): impairing deficit or compensatory behavior? J. Abnorm. Child Psychol. 43, 1219–1232. https://doi.org/10.1007/s10802-015-0011-1.
Scheres, A., Oosterlaan, J., Swanson, J., Morein-Zamir, S., Meiran, N., Schut, H., Vlasveld, L., Sergeant, J.A., 2003. The effect of methylphenidate on three forms of response inhibition in boys with AD/HD. J. Abnorm. Child Psychol. 31, 105–120. https://doi.org/10.1023/A:1021729501230.
Scheres, A., Lee, A., Sumiya, M., 2008. Temporal reward discounting and ADHD: task and symptom specific effects. J. Neural Trans. 115, 221–226. https://doi.org/10.1007/s00702-007-0813-6.
Scheres, A., Tontsch, C., Lee Thoeny, A., 2013. Steep temporal reward discounting in ADHD-Combined type: acting upon feelings. Psychiatry Res. 209, 207–213. https://doi.org/10.1016/j.psychres.2012.12.007.
Schooler, J.W., Smallwood, J., Christoff, K., Handy, T.C., Reichle, E.D., Sayette, M.A., 2011. Meta-awareness, perceptual decoupling and the wandering mind. Trends Cogn. Sci. 15, 319–326. https://doi.org/10.1016/j.tics.2011.05.006.
Schwarz, W., 2001. The ex-Wald distribution as a descriptive model of response times. Behav. Res. Meth. Comput. 33, 457–469. https://doi.org/10.3758/BF03195403.
Scott, B.B., Constantinople, C.M., Erlich, J.C., Tank, D.W., Brody, C.D., 2015. Sources of noise during accumulation of evidence in unrestrained and voluntarily head-restrained rats. Elife 4, e11308. https://doi.org/10.7554/eLife.11308.
Seli, P., Jonker, T.R., Cheyne, J.A., Smilek, D., 2013. Enhancing SART validity by statistically controlling speed-accuracy trade-offs. Front. Psychol. 4, 265. https://doi.org/10.3389/fpsyg.2013.00265.
Sergeant, J.A., 2005. Modeling attention-deficit/hyperactivity disorder: a critical appraisal of the cognitive-energetic model. Biol. Psychiatry 57, 1248–1255. https://doi.org/10.1016/j.biopsych.2004.09.010.
Sharp, S.I., McQuillin, A., Gurling, H.M., 2009. Genetics of attention-deficit hyperactivity disorder (ADHD). Neuropharmacology 57, 590–600. https://doi.org/10.1016/j.neuropharm.2009.08.011.
Shiels, K., Hawk, L.W., Reynolds, B., Mazzullo, R.J., Rhodes, J.D., Pelham, W.E., Waxmonsky, J.G., Gangloff, B.P., 2009. Effects of methylphenidate on discounting of delayed rewards in attention deficit/hyperactivity disorder. Exp. Clin. Psychopharmacol. 17, 291–301. https://doi.org/10.1037/a0017259.
Simen, P., Rivest, F., Ludvig, E.A., Balci, F., Killeen, P.R., 2013. Timescale invariance in the pacemaker-accumulator family of timing models. Timing Time Percept. 1, 159–188. https://doi.org/10.1163/22134468-00002018.
Sjowall, D., Roth, L., Lindqvist, S., Thorell, L.B., 2013. Multiple deficits in ADHD: executive dysfunction, delay aversion, reaction time variability, and emotional deficits. J. Child Psychol. Psychiatry 54, 619–627. https://doi.org/10.1111/jcpp.12006.
Smallwood, J., Schooler, J.W., 2015. The science of mind wandering: empirically navigating the stream of consciousness. Annu. Rev. Psychol. 66, 487–518. https://doi.org/10.1146/annurev-psych-010814-015331.
Solanto, M.V., 2002. Dopamine dysfunction in AD/HD: integrating clinical and basic neuroscience research. Behav. Brain Res. 130, 65–71. https://doi.org/10.1016/S0166-4328(01)00431-4.
Sontag, T.A., Tucha, O., Walitza, S., Lange, K.W., 2010. Animal models of attention deficit/hyperactivity disorder (ADHD): a critical review. Atten. Defic. Hyperact. Disord. 2, 1–20. https://doi.org/10.1007/s12402-010-0019-x.
Sonuga-Barke, E.J., 2005. Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. Biol. Psychiatry 57, 1231–1238. https://doi.org/10.1016/j.biopsych.2004.09.008.
Sporns, O., 2013. Structure and function of complex brain networks. Dialogues Clin. Neurosci. 15, 247–262 PMID: 24174898.
Sutherland, K.R., Alsop, B., McNaughton, N., Hyland, B.I., Tripp, G., Wickens, J.R., 2009. Sensitivity to delay of reinforcement in two animal models of attention deficit hyperactivity disorder (ADHD). Behav. Brain Res. 205, 372–376. https://doi.org/10.1016/j.bbr.2009.07.011.
Swensen, A., Birnbaum, H.G., Hamadi, R.B., Greenberg, P., Cremieux, P.-Y., Secnik, K., 2004. Incidence and costs of accidents among attention-deficit/hyperactivity disorder patients. J. Adolesc. Health 35 (346). https://doi.org/10.1016/j.jadohealth.2003.12.003. e1-346. e9.
Tye, C., Johnson, K.A., Kelly, S.P., Asherson, P., Kuntsi, J., Ashwood, K.L., Azadi, B., Bolton, P., McLoughlin, G., 2016. Response time variability under slow and fast-incentive conditions in children with ASD, ADHD and ASD + ADHD. J. Child Psychol. Psychiatry 57, 1414–1423. https://doi.org/10.1111/jcpp.12608.
van der Maas, H.L., Dolan, C.V., Grasman, R.P., Wicherts, J.M., Huizenga, H.M., Raijmakers, M.E., 2006. A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. Psychol. Rev. 113, 842–861. https://doi.org/10.1037/0033-295X.113.4.842.
van der Meere, J.J., Shalev, R.S., Borger, N., Wiersema, J.R., 2009. Methylphenidate, interstimulus interval, and reaction time performance of children with attention deficit/hyperactivity disorder: a pilot study. Child Neuropsychol. 15, 554–566. https://doi.org/10.1080/09297040902758803.
Wagenmakers, E.J., Van Der Maas, H.L.J., Grasman, R.P.P.P., 2007. An EZ-diffusion model for response time and accuracy. Psychon. Bull. Rev. 14, 3–22.
Wang, X.J., 2008. Decision making in recurrent neuronal circuits. Neuron 60, 215–234. https://doi.org/10.1016/j.neuron.2008.09.034.
Watterson, E., Mazur, G.J., Sanabria, F., 2015. Validation of a method to assess ADHD-related impulsivity in animal models. J. Neurosci. Methods 252, 36–47. https://doi.org/10.1016/j.jneumeth.2015.03.020.
Weigard, A., Huang-Pollock, C., 2017. The role of speed in ADHD-related working memory deficits: a time-based resource-sharing and diffusion model account. Clin. Psychol. Sci. 5, 195–211. https://doi.org/10.1177/2167702616668320.
Weigard, A., Huang-Pollock, C., Brown, S., Heathcote, A., 2018. Testing formal predictions of neuroscientific theories of ADHD with a cognitive model–based approach. J. Abnorm. Psychol. 127, 529–539. https://doi.org/10.1037/abn0000357.
Wickens, J.R., Hyland, B.I., Tripp, G., 2011. Animal models to guide clinical drug development in ADHD: lost in translation? Br. J. Pharmacol. 164, 1107–1128. https://doi.org/10.1111/j.1476-5381.2011.01412.x.
Willcutt, E.G., Doyle, A.E., Nigg, J.T., Faraone, S.V., Pennington, B.F., 2005. Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. Biol. Psychiatry 57, 1336–1346. https://doi.org/10.1016/j.biopsych.2005.02.006.
Willcutt, E.G., Pennington, B.F., Duncan, L., Smith, S.D., Keenan, J.M., Wadsworth, S., Defries, J.C., Olson, R.K., 2010. Understanding the complex etiologies of developmental disorders: behavioral and molecular genetic approaches. J. Dev. Behav. Pediatr. 31, 533–544. https://doi.org/10.1097/DBP.0b013e3181ef42a1.
Williams, B.A., 2010. Perils of evidence-based medicine. Perspect. Biol. Med. 53, 106–120. https://doi.org/10.1353/pbm.0.0132.
Wilson, V.B., Mitchell, S.H., Musser, E.D., Schmitt, C.F., Nigg, J.T., 2011. Delay discounting of reward in ADHD: application in young children. J. Child Psychol. Psychiatry 52, 256–264. https://doi.org/10.1111/j.1469-7610.2010.02347.x.
Ziegler, S., Pedersen, M.L., Mowinckel, A.M., Biele, G., 2016. Modelling ADHD: a review of ADHD theories through their predictions for computational models of decision-making and reinforcement learning. Neurosci. Biobehav. Rev. 71, 633–656. https://doi.org/10.1016/j.neubiorev.2016.09.002.