A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) predominantly hyperactive/impulsive and combined subtypes

Terje Sagvolden
Department of Physiology, University of Oslo, NO-0317 Oslo, Norway
terje.sagvolden@medisin.uio.no http://folk.uio.no/terjesa/

Heidi Aase
Norwegian Centre for the Studies of Behavioural Problems and Innovative Practice, Ltd., University of Oslo, NO-0118 Oslo, Norway
heidiaa@atferd.unirand.no

Espen Borgå Johansen
Department of Physiology, University of Oslo, NO-0317 Oslo, Norway
e.b.johansen@medisin.uio.no

Vivienne Ann Russell
Department of Human Biology, University of Cape Town, ZA-7925 South Africa
russell@curie.uct.ac.za

http://www.uct.ac.za/

Abstract: Attention-deficit/hyperactivity disorder (ADHD) is currently defined as a cognitive/behavioral developmental disorder where all clinical criteria are behavioral. Inattentiveness, overactivity, and impulsiveness are presently regarded as the main clinical symptoms. The dynamic developmental behavioral theory is based on the hypothesis that altered dopaminergic function plays a pivotal role by failing to modulate nondopaminergic (primarily glutamate and GABA) signal transmission appropriately. A hypofunctioning mesolimbic dopamine branch produces altered reinforcement of behavior and deficient extinction of previously reinforced behavior. This gives rise to delay aversion, development of hyperactivity in novel situations, impulsiveness, deficient sustained attention, increased behavioral variability, and failure to “inhibit” responses (“disinhibition”). A hypofunctioning mesocortical dopamine branch will cause attention response deficiencies (deficient orienting responses, impaired saccadic eye movements, and poorer attention responses toward a target) and poor behavioral planning (poor executive functions). A hypofunctioning nigrostriatal dopamine branch will cause impaired modulation of motor functions and deficient nondeclarative habit learning and memory. These impairments will give rise to apparent developmental delay, clumsiness, neurological “soft signs,” and a “failure to inhibit” responses when quick reactions are required.

Hypofunctioning dopamine branches represent the main individual predispositions in the present theory. The theory predicts that behavior and symptoms in ADHD result from the interplay between individual predispositions and the surroundings. The exact ADHD symptoms at a particular time in life will vary and be influenced by factors having positive or negative effects on symptom development. Altered or deficient learning and motor functions will produce special needs for optimal parenting and societal styles. Medication will to some degree normalize the underlying dopamine dysfunction and reduce the special needs of these children. The theory describes how individual predispositions interact with these conditions to produce behavioral, emotional, and cognitive effects that can turn into relatively stable behavioral patterns.

Keywords: catecholamine; clumsiness; dopamine; hyperkinesis; hyperkinetic disorder; impulsivity; monoamine; neuromodulator; overactivity; pollutants; reinforcement; reward; verbally governed behavior; soft signs; variability

1. Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) (American Psychiatric Association 1994) is a seemingly heterogeneous group of behavioral disorders affecting between 2% and 12% of grade-school children (American Academy of Pediatrics 2000; Swanson et al. 1998; Taylor 1998; Taylor et al. 1998). The disorder usually, but not always, manifests itself before the child is 7 years old (Applegate et al. 1997). Of children diagnosed with ADHD, 50% to 70% will have problems related to social adjustment and functioning and/or psychiatric problems as adolescents and young adults (Cantwell 1996). Of these, 20% to 30% will continue to suffer from ADHD during late adolescence and adulthood (Muglia et al. 2000), whereas the full ADHD syndrome is found in only 4% of the adult population (Mannuzza et al. 1998). However, the persistence of ADHD into adolescence and young adulthood varies according to who is being interviewed and the criteria used to define the disorder (Barkley 2002). The finding by Mannuzza and coworkers (Mannuzza et al. 1998) is based exclusively on self-report and is probably an underestimation. In addition, remission rates can be defined as either syndromic (less than full syndrome), symptomatic (less than subthreshold diagno-
In childhood, the disorder is more common in boys than in girls. In the general population, ~9% of males and ~3% of females are found to have behaviors consistent with ADHD (American Academy of Pediatrics 2000). During adolescence and young adulthood relatively more females are affected (Biederman et al. 1994). There might be slight geographic variations in the percentage of children diagnosed as ADHD (Alarcon et al. 1999; Meyer 1998; Taylor 1998). Some of this variation could be caused by different referral practices and different diagnostic criteria (Swanson et al. 1998).

There have been multiple changes in diagnostic criteria for ADHD over the past two decades. Research in this period has sought to identify more homogeneous subtypes. The emphasis has shifted from a unidimensional conceptualization to a model consisting of two factors: hyperactivity/impulsiveness and inattention (for a review of the history, see Taylor et al. 1998). The latter model is based on exploratory and confirmatory factor analyses (cf. Willett et al. 2000). Thus, overactivity, impulsiveness, and inattention are presently regarded as the main clinical symptoms of ADHD (American Psychiatric Association 1994).

The ADHD diagnosis has three subtypes based on two behavioral dimensions: the ADHD predominantly inattentive subtype that is more typical among girls than boys (Taylor et al. 1998), the ADHD predominantly hyperactive/impulsive subtype that is more typical among boys than girls with a diagnosis of ADHD (Taylor et al. 1998), and the combined subtype. The inattention dimension includes difficulty in sustaining attention, distractibility, lack of persistence, and disorganization. The hyperactivity/impulsiveness dimension includes excessive motor activity and impulsive responding (Lahey et al. 1998). Admittedly, the symptoms are not that well defined, and requirements vary somewhat between the ICD and DSM taxonomies (Swanson et al. 1998; Taylor 1998). According to DSM-IV criteria, it is possible to have “ADHD” without being inattentive. Inattentiveness is, however, a necessary requirement for a hyperkinetic disorder according to ICD-10 criteria (Taylor 1998).

Disruptive behavioral disorders and internalizing disorders are the most common comorbid disorders in ADHD. The disruptive behavioral disorders, Oppositional Defiant Disorder (ODD) and Conduct Disorder (CD), coexist with ADHD in ~35% of children. Internalizing disorders like anxiety and mood/depressive disorders coexist with ADHD in ~25% and ~18%, respectively (American Psychiatric Association 1994; American Academy of Pediatrics 2000). Finally, learning disabilities (e.g., reading disorder, dyslexia, dyscalculia, problems with writing) are common (~25%), especially in boys with ADHD (Biederman et al. 2002c; Seidman et al. 2001). Estimates of comorbid learning disabilities range from ~7% to ~92%, depending on the definitions used (DuPaul & Stoner 1994).

### 1.1. A dynamic developmental theory of ADHD

#### 1.1.1. Behavioral foundations of ADHD symptoms

The search for a pivotal behavioral deficit in behaviorally defined ADHD and its corresponding neurobiological correlates has proven particularly challenging. A comprehensive neuropsychological model of ADHD has yet to be proposed, although models of other psychopathologies have been suggested previously, for example, by Gray (Gray 1982; Gray et al. 1991).

We will offer a novel behavioral theory of ADHD that to
A large extent is inspired by behavioral analysis (see Catania’s precommentary accompanying this article). In parallel we will suggest how this theory may be related to neurobiological factors. There is increasing agreement that dysregulation of frontostriatal circuits may underlie many of the behavioral symptoms of ADHD (Biederman & Faraone 2002; Castellanos 1997; Castellanos & Tannock 2002; Grace 2001; 2002; Johansen et al. 2002; Sagvolden & Sergeant 1998; Solanto et al. 2001b). We will explore behavioral predictions from the point of view of interactions between dysregulated frontostriatal circuits and hypofunctioning dopamine systems. We realize that many other changes besides hypofunctioning dopamine systems necessarily will be present in ADHD, including upregulation of parts of these systems. We suggest that explanations and predictions derived from hypofunctional dopamine system branches should be explored to the fullest before aiming research at other neurotransmitter systems. By ignoring other possible changes, we hope to be more concrete in the theoretical issues involved. It might also facilitate the design of future studies.

In a dynamic developmental theory we will argue that there might be two main behavioral processes causing ADHD: altered reinforcement of novel behavior and deficient extinction of previously reinforced behavior. These processes may primarily be associated with a hypofunctioning mesolimbic dopamine system (Johansen et al. 2002) and will probably interact with effects of other hypofunctioning dopamine systems: a hypofunctioning mesocortical dopamine system associated with deficient attention and poor behavioral organization and a hypofunctioning nigrostriatal dopamineergic system impairing motor functions and causing poor nondeclarative habit learning (Fig. 1). The stunted dopamine responses might be caused by a combination of insufficient glutamate input from the prefrontal cortex to dopamine neurons and faulty regulation of dopamine release (see sects. 2.1 and 2.4 below).

The main behavioral selection mechanisms, reinforcement and extinction, are associated with dopamine neuron activity, which at a neurobiological level may have the function of constantly reprogramming neuronal connections by strengthening (reinforcing or potentiating) connections associated with reinforced (usually adaptive) behavior, while at the same time weakening (extinguishing or depressing) other neuronal connections associated with nonreinforced (usually maladaptive) behavior. Reinforcement operates within a limited time window from the occurrence of the behavior to the perception of the consequences of this behavior (for more details see sects. 2.5.1 and 3.3).

We argue that the time available for associating behavior with its consequences will be shorter in ADHD than in normal children if dopamine systems are hypofunctioning. A narrower time window in ADHD will restrict the stimuli controlling their behavior and therefore explain some of the attention problems seen in ADHD. Such a narrower time window will also preferentially select short sequences of behavior giving rise to motor impulsiveness. In addition, we suggest that hypofunctioning dopamine systems lead to a deficient behavioral extinction process. This will cause excessive behavior, usually labeled hyperactivity, and increased behavioral variability, frequently interpreted as failure to inhibit responses. We argue that response disinhibition is at best misleading and usually a misinterpretation.

The dynamic developmental theory disentangles aspects of various deficient executive functions in ADHD into impulsiveness caused by inefficient reinforcement, deficient extinction of previously acquired behavior, and impaired motor control. The concept of impulsiveness has both a motor and a cognitive component. Motor impulsiveness is presently defined as bursts of responses with short inter-response times (IRTs). This behavior has been shown to emerge in children with ADHD (Sagvolden et al. 1998) as well as in the best-validated animal model of ADHD (Sagvolden 2000). Cognitive impulsiveness implies that private events like thoughts and plans are dealt with for short sequences of time with rapid shifts, resulting in problems with generating and following plans, problems with organizing own behavior, forgetfulness, and inefficient use of time. Although some aspects of cognitive impulsiveness may fit the notion of response “disinhibition” (Barkley 1997b), we will argue that these aspects may be explained as being caused by slower acquisition of long sequences of behavior and deficient extinction of previously reinforced behavior.

ADHD behaviors such as increased reaction times and speed variability (Oosterlaan & Sergeant 1995b; Rubia et al. 1998) have been described as evidence of impaired executive functions by some authors (Kooijmans et al. 2000) and as response “disinhibition” by others (Sonuga-Barke 2002; Pliszka et al. 2000). These behaviors will be explained as more fundamental, simpler motor problems: impaired timing of starting and stopping of responses; impaired acquisition, retrieval, and relearning of programs for sequential motor tasks; and deficient nondeclarative habit learning and memory.

1.1.2. ADHD in a developmental perspective. Behavior and symptoms in ADHD result from the interplay between individual predispositions and the surroundings. Thus, the dynamic developmental theory predicts that the exact ADHD symptoms at a particular time in life will vary and be influenced by factors having positive or negative effects on symptom development (Fig. 2).

The theory describes how individual variations in dopamine functioning may affect learning processes and motor
functions, thereby producing ADHD behavior: attentional problems, hyperactivity, and impulsiveness. The theory also predicts increased behavioral variability. Altered, or deficient learning and motor functions will produce special needs for optimal parenting and societal styles. Medication will to some degree normalize the underlying dopamine dysfunction and reduce the special needs of these children. The theory describes how individual predispositions interact with these conditions and produce behavioral, emotional, and cognitive effects that can turn into relatively stable behavioral patterns.

1.2. Symptoms of ADHD

Inattention, hyperactivity, and impulsiveness are regarded as the main clinical symptoms. These symptoms are frequently explained as caused by faulty executive functions and/or deficient behavioral inhibition (see sect. 1.2.3).

1.2.1. Deficient sustained attention. Attention, in the widest sense, refers to the relationship between behavior and the environment. One is attending to a stimulus, or stimulus property, when variation of that stimulus or stimulus property changes behavior (Catania 1998). Attention is modified by a multitude of psychological factors like sensory and motivational processes. In various forms, inattention is found in most psychiatric disorders except mania (Taylor 1994), and it could well be that some non-ADHD disorders masquerade as ADHD (Mannuzza et al. 1993). Inattentive behavior is poorly operationalized, and judgments about inattentiveness are inferred from observed behavior. This means that so-called inattentive behavior may be produced by other deficits leading to poor test scores on measures of attention; that is, behavioral changes may be misinterpreted as inattention.

Functional mapping of brain electrical activity indicates multilevel deficits in sensory processing in children with ADHD (Pliszka et al. 2000). Thus, a combination of cognitive and sensory-processing deficits may be the underlying bases of inattentive behavior observed in ADHD. It is, however, beyond the scope of the present article to review the neuropsychology of attention (for a review, see Posner & Petersen 1990).

1.2.2. Hyperactivity. An excessive level of activity is typically seen in ADHD as restlessness, fidgeting, and a general increase in gross body movements (Porrino et al. 1983; Taylor 1998; Teicher et al. 1996). Ratings of hyperactivity (and of impulsiveness) involve an element of overstepping...
impulsive or explicit social rules and are judged according to situational appropriateness (Taylor 1998). Although children with ADHD move twice as frequently and cover a fourfold wider area, the ADHD movement pattern is less complex and more linear (side to side) compared to normal controls (Teicher et al. 1996). Overactivity is seen in some situations such as the classroom but might not be present in others such as play (Porrino et al. 1983). It seems that the ADHD overactivity is absent in novel situations (Sagvolden et al. 1998; Sleator & Ullman 1981). Clinical evaluation of hyperactivity statistically often overlaps with impulsiveness (Taylor 1998).

1.2.3. Impulsiveness and executive functions. ADHD impulsiveness has often been explained as being caused by faulty executive functions (EFs). In DSM-IV, impulsiveness is operationalized as blunting out answers before questions have been completed, having difficulty waiting one's turn when this is appropriate, and frequent interruption and intrusion on activities of other people. In general terms, impulsiveness means acting without reflecting and failure to plan ahead. In the literature, however, impulsiveness is a heterogeneous concept, including terms such as over-rapid responsiveness, sensation seeking, risk taking, novelty seeking, excessive attraction toward immediate reward, boldness, adventuresomeness, accident-proneness, boredom susceptibility, unreliability, and disorderliness. Measures of impulsiveness necessarily become heterogeneous, ranging from motor and cognitive measures to more complex behaviors.

Executive functions denote psychological processes involved in the organization and planning of behavior (Denckla 1996; Tannock 1998). Building on more fundamental cognitive processes, executive functions consist of an assembly of higher-order cognitive functions, and the term is used interchangeably with concepts like self-control. Impulsive behavior has been suggested to be a result of executive dysfunction caused by behavioral disinhibition (Barkley 1997b). However, the concept of behavioral inhibition is an ambiguous term, both regarded as one of the executive functions as well as referring to one of the fundamental processes underlying executive functions. In addition, the concepts of inhibition/disinhibition have multiple meanings and operationalizations (Sergeant et al. 1999). Also, empirical findings on disinhibition as a characteristic of ADHD are inconclusive (Scheres et al. 2001).

The concept of inhibition has a variety of meanings and a long history (MacLeod et al. 2003). According to Webster's New Universal Unabridged Dictionary, the Latin verb *inhibere* means to hold back, restrain, or curb. This dictionary lists two main meanings: (1) to prohibit, forbid; and (2) to suppress, withhold, or check. MacLeod lists two main meanings in medicine and behavioral science, on a neuronal and on a behavioral level. He questions the evidence for the cognitive concept of inhibition, but not the neurobiological concept. The main problem is that one cannot derive the concept of inhibition directly from the concept of inhibition at the neural level (MacLeod et al. 2003).

In what is now known as neuroscience, the phenomenon of “inhibition” had its breakthrough in a monograph showing that stimulation of the cut vagus nerve causes temporary cessation of the heart beat (Weber & Weber 1846). The concept of “inhibition of responses,” and its relation to limbic areas of the brain, has a long history starting with the seminal electrophysiological works of the neurophysiologist Birger R. Kaada (1951). Kaada showed, along with many other observations, that electrical stimulation of the subcallosal-septal area produced inhibition of respiration, spinal reflexes, and cortically induced movements. Stimulation of the cingulate cortex produced facilitation of these reflexes. These results were generalized from reflexes to more complex behavior by Robert A. McCleary (1966) showing a double dissociation following lesions of the following areas: In passive avoidance where the subject has to withhold responding to avoid the aversive stimulus, lesions of the septal nuclei produced deficits; but there was no change following cingulate lesions. In active avoidance, where the subject is required to perform an active response to avoid the aversive stimulus, septal lesions improved performance, but there were impairments following cingulate lesions.

Excitation and inhibition are fundamental synaptic processes that may explain reflexes involving a few synapses (cf. Kaada 1951), but as any textbook in neurobiology will point out, even simple spinal cord reflexes are highly intricate and multidimensional. Kaada’s and McCleary’s results were generalized into a theory of psychopathology by Jeffrey A. Gray (1982), which later was developed into theories of ADHD by Herbert C. Quay (1988), Russell A. Barkley (1997b), and others.

Mainly based on Barkley’s and Quay’s theories, response inhibition is now used extensively as an explanation of ADHD symptoms, perhaps without realizing that the neuronal activities behind integrated behavior are the results of extremely complex sequences of excitations and inhibitions, probably involving large numbers of synapses in highly complex neuronal networks, making exact predictions from a synaptic level to behavior exceedingly difficult. Further, the response unit that is supposed to be inhibited is hard to define empirically (Catania 1998). Few studies have addressed empirically whether functional behavioral units are only active responses or include “passive responses,” for example, recordable inactivity in a sequence of active responses. Iversen has argued that units should be based on functional analyses, not a priori assumptions regarding behavioral structure. The proper unit is what emerges when a reinforcement contingency is applied (Iversen 1991). Further, functional response units are unlikely to remain constant even within the same individual, for example, under the influence of drugs when response chains sometimes appear to be truncated (Lyon & Robbins 1975).

As long as the unit of behavior (in this case the inhibited response) is not identified, the nature of response inhibition and disinhibition remains enigmatic (Johansen et al. 2002; Sagvolden & Sergeant 1998). Hence, it is unclear whether response inhibition is as basic a mechanism as is often suggested, for example, by Barkley (1997b). The use of “response inhibition” as an explanation of ADHD symptoms may be another example of an overly simplistic idea that influences research primarily because of its appeal.

1.2.4. One or two disorders? The various attention problems associated with the ADHD subtypes are quite different from each other (Barkley 1997b; Johansen et al. 2002; Taylor 1998). Children with the ADHD inattentive subtype are often non-hyperactive, rather dreamy, and inert children. Their attention problems are non-specific and related to deficient sensory processes, poorly focused attention,
and less accurate information processing. Such problems lead to reading disorder and learning disability and may be associated with reduced IQ. Usually, such attention problems are associated with a family history of learning problems, sluggish cognitive processes, and school failure (Taylor 1998; Willcutt et al. 2000). Children with ADHD predominantly hyperactive/impulsive subtype do not have general attention problems in the same sense. Their attention problems are more specifically related to distractibility and reduced persistence and are present after correcting for IQ (Taylor et al. 1991). Furthermore, this subtype is associated with memory retrieval problems, disruptive behavior, and peer rejection.

The subtypes may have very different developmental courses both in terms of outcome and comorbidity (Willcutt et al. 2000). Although individuals with ADHD predominantly inattentive subtype may be more socially withdrawn, experience greater academic problems, and develop comorbid anxiety or other mood disorders, early hyperactive/impulsive behavior is associated with externalizing problems like aggression, oppositional behavior, adolescent delinquency, and substance abuse (Barkley 1997b).

There is little or no data on medical treatment of ADHD inattentive subtype (NIH Consens Statement 1998), although there seems to be a common clinical notion that methylphenidate also helps these children. However, response to central stimulant medication is not specific to ADHD and cannot be used as a diagnostic criterion: both methylphenidate and d-amphetamine have been shown to have similar effects in boys with ADHD and healthy boys (Conners 2002; Rapoport & Inoff-Germain 2002; Rapoport et al. 1978; 1980); d-amphetamine has been shown to decrease impulsive choice in healthy volunteers (de Wit et al. 2002) and methylphenidate to increase the amount of self-control choices in non-ADHD criminals and former substance abusers with or without conduct disorder (Pietras et al. 2003).

In conclusion, symptoms and developmental course indicate that the present ADHD diagnosis consists of two separate disorders probably with separate etiology: (1) Attention Deficit Disorder predominantly inattentive type without impulsiveness and hyperactivity (ADD in the text following) and (2) the Hyperactive/Impulsive Disorder with hyperactivity, impulsiveness, and problems with sustaining attention developing into ADHD combined type (ADHD in the text following). We suggest that the latter disorder might be named Reinforcement/Extinction Disorder (RED) according to the proposed underlying dysfunctions (cf. Sagvolden & Archer 1989).

We acknowledge the likelihood of ADHD subtypes and also recognize the importance of other neurobiological factors. However, most explanatory models of ADHD address mainly the predominantly hyperactive/impulsive or the combined subtype (Castellanos & Tannock 2002; Tannock 1998). This is also the case with the dynamic developmental theory of ADHD.

2. Etiology

Abnormal dopamine function has been the focus of attention in the search for the neurobiological basis of ADHD because of the assumed dopamine agonistic action of the stimulant drugs (Biederman et al. 2002a; Castellanos 1997; Castellanos et al. 2002; Johansen et al. 2002; Volkow et al. 1998) that for several decades have provided the primary pharmacological treatment for ADHD (Bradley 1937; Conners 2002; Rapoport & Inoff-Germain 2002; Solanto et al. 2001b).

2.1. Neurobiological bases of ADHD

Dopamine effects on prefrontal functioning are complicated (for reviews of dopamine neuroanatomy and physiology, see Grace 2002; Haber et al. 2000; Missale et al. 1998; Schultz 2002). Dopamine exerts a strong regulatory effect on prefrontal cortical pyramidal neuronal activity. These neurons exhibit bistable membrane potentials alternating between a hyperpolarized, non-firing state and a depolarized, action-potential-firing state. The effects of dopamine stimulation on these prefrontal cells depend on this state (Grace 2002). The glutamatergic output from these neurons projects to the nucleus accumbens and the ventral tegmental area (VTA) and exerts a strong regulation of the activity in these areas.

We suggest that dopamine ought to be thought of as a neuromodulator rather than as a neurotransmitter (Siegelbaum et al. 2000). Its effects are relatively long-lasting ones acting on metabotropic receptors coupled to G proteins (Missale et al. 1998). The dopamine actions may best be described not in terms of inhibition or excitation, but rather as gating of inputs and modulation of states of neuronal elements (Grace 2002). Dopamine has potent regulatory control over interactions between neighboring neurons in target areas of the brain (Grace 2002). At the systems level, dopamine exerts a focusing effect whereby only the strongest signals will pass through the striatum to the pallidum (Schultz 2002). On a behavioral level, “the arrival of the dopaminergic input to the striatum is best seen as providing a temporal window permitting change, rather than as providing a direction to that change” (Gray et al. 1991, p. 17).

Dopamine is the predominant catecholamine neuromodulator in the mammalian brain (Missale et al. 1998). There are at least five distinct G protein-coupled dopamine receptor subtypes all with seven transmembrane domains (Missale et al. 1998): Two D1-like receptor subtypes (DRD1 and DRD5) are primarily situated postsynaptically and are coupled to the stimulatory G protein Gs by a short third intracellular loop activating adenylyl cyclase and thereby stimulating cAMP formation (Fig. 3). The D1-like receptors increase intracellular calcium via various mechanisms. Furthermore, there are three D2-like receptor subtypes (DRD2, DRD3, and DRD4) that are coupled to the inhibitory G protein Gi by a long third intracellular loop common to receptors inhibiting adenylyl cyclase and thereby cAMP formation. The D2-like receptors are found both pre- and postsynaptically. Postsynaptically, these receptors activate K+ channels and reduce calcium influx into the cell via various mechanisms.

The pharmacological profiles of the D-1-like and D-2-like receptors are different (for a review, see Missale et al. 1998). However, the pharmacological differences within each dopamine receptor subfamily are relatively small and in general related to differences in affinity of various agonists and antagonists.

Dopamine receptors are also found outside of the central nervous system, even in places where there are no dopamine-releasing varicosities, such as the cardiovascular...
system where it is involved in controlling microcirculation (Krimer et al. 1998). For example, the dopamine receptor 4 (DRD4) is found on the heart (Missale et al. 1998). It seems that the function of dopamine receptors within the cardiovascular system control synergistically operating systems reducing or increasing blood pressure. Defective renal dopamine production and/or dopamine receptor function have been reported in human primary hypertension as well as in genetic models of animal hypertension, suggesting that dopaminergic abnormalities are not a secondary effect of hypertension (Amenta et al. 2001). Dopamine receptors are also found in the renin-angiotensin-aldosterone system (Missale et al. 1998) involved in stress and blood pressure control. This might both suggest an association between ADHD, type-A personality, and hypertension (Whalen & Henker 1986), and explain why the spontaneously hypertensive rat strain turns out to be a good animal model of ADHD (Sagvolden 2000). Finally, dopamine controls sodium chloride concentrations in the kidneys (Amenta et al. 2001). It might be that reduced dopamine functions change thirst and micturition in children with ADHD.

2.2. Genetic bases of ADHD

Mental disorders like ADHD are extremely challenging to genetic researchers because they do not stem from errors in single genes, but from polymorphisms that create subtle differences in human behavior and are likely to interact with the environment to create symptoms and functional impairment. In addition, both genes and environment appear to be complexly and interactively involved in the development of mental disorders, perhaps with multiple components of each. Furthermore, a mental disorder such as ADHD probably represents the pathological end of a continuum that includes normal functions.

The genetic basis of ADHD might be rather complicated. No single gene stands out as an obvious candidate. This reflects a polygenic and multi-determinant etiology of ADHD. Evidence from twin, adoption, and family studies has found heritability to be ~50% in ADHD. However, high heritability does not imply neurobiological determinism; the behavioral result will still heavily depend on interactions with the environment (Biederman et al. 2002a; Taylor et al. 1998). Dopamine genes have been the initial candidates for investigation (Solanto et al. 2001b). Several studies have concentrated on possible links between genes coding for dopamine receptors and ADHD.

Dopaminergic neurons are complicated structures with intricate interactions with other neurons and glial cells (Fig. 3). Even the simplest of behavioral reflexes is controlled by many neurons involving several neuronal signal substances and a multitude of receptors. The various neuromodulators that have been implicated in ADHD are very tightly linked neuroanatomically, such that functional changes in one undoubtedly will affect the functioning of the others (de Vilers et al. 1995). ADHD is most likely a polygenetic disorder (Taylor 1998) where the polygenetic contribution to the disorder interacts with environmental factors in producing the behavioral expression (Taylor et al. 1998). The most frequently found genes linked to ADHD are almost all associated with neuromodulatory functions.

The high heritability of ADHD is likely to be due to multiple genes with small effect size rather than a few genes of major effect. A lot of scientific interest has focused on the human DRD4 gene mapped to chromosome 11p15.5. DRD4 is highly expressed in the frontal cortex, the amygdala, the hippocampus, the hypothalamus, and the mesencephalon, and to a lesser extent in the globus pallidus and the substantia nigra pars reticulata. Finally, the DRD4 is found on the heart and in the retina (Missale et al. 1998). Unlike most G-protein coupled receptors that have no introns interrupting the coding sequence, the DRD4 gene has three such introns.

The human DRD4 gene exhibits extensive polymorphic...
variations within the coding sequence. There are several insertions in the functionally significant third intracellular loop. A 48-base pair sequence in the third intracellular loop exists either as a single repeat of the sequence or as multiple repeats. The most common is the 4-repeat form, followed by the 7- and 2-repeat forms (Missale et al. 1998). The 7-repeat allele in exon 3 of the DRD4 gene may be associated with a subsensitive postsynaptic receptor (Missale et al. 1998). Several candidate gene studies have identified an association between a 7-repeat variant in exon 3 of the DRD4 (or a neighboring locus) and ADHD (as well as novelty seeking and Tourette’s syndrome) (Barr et al. 2000a; Holmes et al. 2002; Manor et al. 2002), although other studies have failed to replicate this association (Castellanos et al. 1998; Fisher et al. 2002; Mill et al. 2002; Smith et al. 2003). A recent meta-analysis indicates a small but real association between the DRD4 7-repeat and ADHD (Faraone et al. 2001).

Although the presence of the DRD4 7-repeat allele may be associated with a modestly increased risk for ADHD, it is not a necessary condition because about half of the children with ADHD do not have a 7-repeat allele. Nor is it a sufficient condition because ~20% of the unaffected controls have a 7-repeat allele (Swanson et al. 2000b). Surprisingly, in this study the ADHD subgroup defined by the presence of the 7-repeat allele showed normal response speed and variability in neuropsychological tests designed to probe attention networks with neuroanatomical foci in D4-rich brain regions, whereas the subgroup of children with ADHD without the 7-repeat variant of the DRD4 showed the expected abnormality of slow and variable responses. Dopamine receptors are distributed differently in population groups across the world, and DRD4 7-repeat allele might be associated with novelty seeking, perseverance, and migration (Ding et al. 2002).

Other dopamine receptor genes have also been investigated. The dopamine receptor 1 (DRD1) is the most widespread dopamine receptor in the brain. DRD1 gene polymorphism does not seem to be associated with ADHD (Kuntsi & Stevenson 2000). The dopamine receptor 2 (DRD2) is mainly expressed in the neostriatum and in the olfactory tubercle. The DRD2 gene is associated with ADHD in some (Conings et al. 1996), but not all studies (Kuntsi & Stevenson 2000; Todd & Lobos 2002). There might be an association with substance abuse (Blum et al. 1995). The dopamine receptor 3 (DRD3) does not seem to have a role in ADHD (Barr et al. 2000b). The dopamine receptor 5 (DRD5) is found in the hippocampus (where the DRD5 is highly present compared to the DRD1), the dentate gyrus, the entorhinal cortex, the lateral mammillary nucleus, the diagonal band of Broca, the prefrontal and premotor cortices, the lateral thalamus, and the neostriatum. DRD5 levels are relatively low compared to those of DRD1. In general, the dopamine receptors are found on inhibitory GABA neurons, but the DRD5 are also situated on large cholinergic interneurons (Missale et al. 1998). There might be an association between ADHD and a polymorphism near the DRD5 gene (Tahir et al. 2000) in certain ADHD families (Fisher et al. 2002; Hawi et al. 2003).

The plasma membrane dopamine transporter (DAT1) provides major regulation of synaptic and extrasynaptic levels of dopamine and is a principal target of psychostimulant drugs (Grace 2002; Missale et al. 1998; Volkow et al. 1998). The DAT1 gene has 15 exons, several introns, and several polymorphisms. The 10-repeat allele of the DAT1 gene may be associated with increased reuptake of dopamine (Swanson et al. 2000h). Allelic variations of the DAT1 gene have been linked to ADHD in some (Chen et al. 2003; Comings et al. 1996; Hawi et al. 2003; Kuntsi & Stevenson 2000) but not in all studies (Fisher et al. 2002; Muglia et al. 2002; Palmer et al. 1999). Finally, catechol-O-methyl transferase (COMT), an enzyme-metabolizing catecholamine, may be involved in ADHD gender differences in Han Chinese (Qian et al. 2003).

The dopamine systems are linked to the noradrenergic (NE) neuromodulator system originating in the locus coeruleus. Plasma norepinephrine concentrations may be significantly increased in ADHD children with reading disorder and other cognitive disabilities compared to ADHD children without learning disabilities (Halperin et al. 1997). ADHD, especially when associated with learning disabilities and poor grade-school academic performance, has been shown to be associated with the dopamine-beta-hydroxylase (DBH) enzyme converting dopamine into norepinephrine (Comings et al. 1996; Hawi et al. 2003; Smith et al. 2003) and noradrenergic genes: the adrenergic alpha2A receptor (ADRA2A), adrenergic alpha2C receptor (ADRA2C), and DBH genes (Comings et al. 1999). Although the dopamine transporter DAT1 may be involved in ADHD, the gene for the norepinephrine transporter (NET1) does not seem to be a susceptibility factor in ADHD (Barr et al. 2002).

The dopamine systems are also closely linked anatomically to the serotonergic (5-HT) neuromodulator systems originating in the brainstem raphe nuclei. Reduced central serotonergic activity has been implicated in poor impulse regulation and aggressive behavior. There is evidence for an involvement of 5-HT transporter polymorphism in ADHD (Cadoret et al. 2003; Fisher et al. 2002; Kent et al. 2002). A linkage between polymorphisms in the serotonin HTR2A receptor gene and ADHD has been shown (Quist et al. 2000), but not in all studies (Levitan et al. 2002; Zoroglu et al. 2003). The 5-HT1B receptor, however, may be involved in ADHD (Quist et al. 2003). It could be that norepinephrine and serotonin imbalances contribute to a dopaminergic imbalance, which underlies the possible complex interplay among the neurotransmitter systems in the etiology of ADHD.

In conclusion, it might not be one critical gene associated with ADHD. Instead, ADHD could be the result of one of several combinations of genes producing postsynaptic changes of a magnitude exceeding the capacity of normal neuronal or behavioral compensatory mechanisms. This may explain why the same gene allele has not been found to be critical in all studies. Another possibility is that environmental factors (e.g., density of reinforcers, or number and intensity of environmental stimuli) contribute to normalization of synaptic function despite an unfavorable genetic constitution. In addition, it is conceivable that ADHD consists of subgroups that can be differentiated according to the genetic make-up.

### 2.3. Nongenetic factors in the etiology of ADHD

As reviewed above, dopamine dysfunction seems to play a pivotal role in the neurobiology of ADHD. Reductions in dopaminergic functioning can result from genetic as well as nongenetic factors (Fig. 1). For example, dopamine agonist
drugs such as cocaine, crack, and amphetamines produce down-regulation of dopamine synthesis (Scalfi et al. 1996). The down-regulation and ADHD-like symptoms persist until dopamine functions normalize.

Drug addicts and children exposed to drugs of abuse prenatally exhibit ADHD-like behavior (Mick et al. 2002; Vogel 1997). Development of ADHD symptoms is a dynamic process of adaptation to defective neurotransmission in the developing brain. It is important to understand the current status of the nervous system of children with ADHD to gain insight into the pathogenesis of ADHD. Excitatory inputs to the ventral tegmental area (VTA) dopamine neurons and to the nucleus accumbens are critical for the development of sensitization and addiction to drugs of abuse (Bonci et al. 2003; Ryu et al. 2002; Saad et al. 2003; Thomas et al. 2001; Vanacore et al. 2002; Wolf 1998). Sensitization involves incremental adaptations to these drugs.

Chronic in vivo administration of cocaine increases dopamine release in the nucleus accumbens and elicits a long-lasting depression of synaptic strength at synapses made by prefrontal cortical afferents onto medium spiny neurons in the shell subdivision of the nucleus accumbens, a change that is required for the maintenance of behavioral sensitization and addiction (Thomas et al. 2001). As a result of the cocaine-induced decrease in synaptic strength of cortical afferent connections, the magnitude of long-term depression (LTD, see sect. 2.5.2) is reduced in the nucleus accumbens shell (Thomas et al. 2001), thereby impairing extinction.

We suggest that inappropriate overactivity of mesolimbic VTA dopamine neurons at an early stage of development of ADHD could similarly increase excitatory synaptic transmission in the VTA dopamine neurons. This could perhaps result in depolarization block of VTA dopamine neurons and hypoactivity of the mesolimbic dopaminergic system.

Worldwide, more than 3,000 chemicals are produced in high volumes (over 500 tons per year). Few of these have been adequately tested for their effects on the developing brain. It is documented that some environmental toxins cause a wide variety of problems, including impairments in attention, memory, learning, social behavior, and IQ (Stein et al. 2002). Some environmental pollutants cause dopamine dysfunction. Epidemiological studies have linked insecticide, herbicide, and fungicide exposure to Parkinson's disease. The concentrations and types of these chemicals vary between countries and regions within a country. Pyrethroid insecticides reduce striatal dopamine function (Pittman et al. 2003) and induce anxiety-like behavior in rats (Righi & Palermo-Neto 2003). The insecticide rotenone causes the death of dopaminergic neurons in vitro and in vivo by mitochondrial chain complex I inhibition, and it is widely used to model Parkinson's disease in animals (Beal 2003; Imam 2003; Vanacore et al. 2002). The herbicide paraquat and the fungicide maneb enhance sensitivity of the aging nigrostriatal dopamine pathway resulting in irreversible and progressive neurotoxicity in mice (Thiruchelvam et al. 2003). It is not known whether these chemicals are able to induce ADHD-like symptoms.

Polychlorinated biphenyls (PCBs) constitute a group of halogenated aromatic hydrocarbons that is lipophilic and, consequently, bioaccumulating (Holene et al. 1998). The lipophilic nature of PCBs makes organs like the brain particularly vulnerable. Intake of these pollutants causes developmental abnormalities in human babies including low birth weight, disruptive behavior, and overactivity (see See-gal 1996 for references). A series of studies of effects of PCB exposure on behavior and brain chemistry (Holene et al. 1995; 1998) showed that normal male rats exposed to subtoxic doses of the PCB congener 153 through mother's milk when pups were hyperactive and impulsive after they had grown up. Their behavior was closely similar to that shown by the spontaneously hypertensive rat, the best validated animal model of ADHD (Sagvolden 2000). Similar behavioral changes are shown by rats either consuming food adulterated with the commercial PCB mixture Aroclor 1248 or PCB-contaminated St. Lawrence River carp (Berger et al. 2001). Although the various PCBs work via different routes, the most likely mode of action of di-ortho-substituted PCB congeners like PCB 153 producing hyperactivity and motor impulsiveness is via monoaminergic pathways. Dopamine and serotonin levels are reduced (Chu et al. 1996) probably by a combination of an inhibition of dopamine synthesis and deficient vesicular storage or release (Chishiti et al. 1996).

There is an increasing amount of evidence from prospective studies suggesting a strong linkage between maternal smoking during pregnancy and the development of ADHD, conduct disorder, learning difficulties, and later substance abuse in the offspring (Weissman et al. 1999). Also, fetal exposure to alcohol is associated with adolescent behavioral and learning problems (Olson et al. 1997; Weinberg 1997). The direct impact of and neurological mechanisms involved in such exposure on the development of ADHD in the child is not yet established; they are considered to be rather general (but highly increased) risk factors for later behavioral, social, and learning problems.

Finally, children with ADHD are about three times more likely to have been born with low birth weight (LBW) than non-ADHD children. Although birth weight is highly heritable, the increased incidence of LBW among children with ADHD might have other, nongenetic causes. Children with LBW, however, make up a relatively small proportion of children with ADHD after attending to potential confounders such as prenatal exposure to alcohol and cigarettes, parental ADHD, social class, and comorbid disruptive behavior disorders in parents and offspring (Mick et al. 2002). A recent prospective study following children from birth to mid-adolescence found that small for gestational age status had only modest independent impact on learning, cognition, and attention in adolescence (O'Keefe et al. 2003).

2.4. Important neuronal loops

The brain serves behavior by increasing or decreasing activity in neural networks that connect neurons in different anatomical regions of the brain that communicate with each other. The functioning of midbrain dopaminergic neurons and their projection areas, particularly the prefrontal cortex and striatum, has been implicated in ADHD (Castellanos 1997; Castellanos & Tannock 2002; Grace 2001; Sagvolden & Sergeant 1999; Solanto et al. 2001b; Teicher et al. 2000). Dopamine and the other neuromodulators exert distinct regulatory actions on the transfer of information through neural circuits that connect, among other structures, frontal cortical areas with the striatum (the nucleus accumbens septi, the caudate nucleus, and the putamen), the pallidum, the thalamus, the substantia nigra, and the ventral tegmental area (Alexander et al. 1986).
In general, neurons in the frontal cortical areas send excitatory glutamatergic projections to the striatum including the nucleus accumbens (ventral striatum). These structures send inhibitory GABAergic projections to the pallidum and substantia nigra that inhibit thalamic nuclei through GABAergic connections. Finally, the thalamus completes the circuit by sending excitatory glutamatergic projections to cortical neurons. The figure illustrates that there are several circuits like this probably with distinct functions. On each level, the functioning may be modulated by the actions of stimulatory G proteins, $G_s$, and inhibitory G proteins, $G_i$, associated with various neuromodulators. (Adapted from Alexander et al. 1986.)

Figure 4. Neurons in the frontal cortical areas send excitatory glutamatergic projections to the striatum. These structures send inhibitory GABAergic projections to the pallidum (and substantia nigra) that inhibit thalamic nuclei through GABAergic connections. Finally, the thalamus completes the circuit by sending excitatory glutamatergic projections to cortical neurons. The figure illustrates that there are several circuits like this probably with distinct functions. On each level, the functioning may be modulated by the actions of stimulatory G proteins, $G_s$, and inhibitory G proteins, $G_i$, associated with various neuromodulators. (Adapted from Alexander et al. 1986.)

In general, neurons in the frontal cortical areas send excitatory glutamatergic projections to the generally silent, medium-spiny neurons of the striatum including the nucleus accumbens (ventral striatum). These structures send inhibitory GABAergic projections to the normally active neurons of the pallidum and the substantia nigra that inhibit thalamic nuclei through GABAergic connections. Finally, the thalamus completes the circuit by sending excitatory glutamatergic projections to cortical neurons (Fig. 4). Figure 4 is a simplification, however, as it omits interactions between the various loops. For example, dopamine release in the nucleus accumbens shell influences dopamine release in the core which in turn influences dopamine release in the caudate-putamen, and so on, in an ascending spiral (see Haber et al. 2000).

There are three circuits like this that may have distinct functions: first, a prefrontal loop, apparently involved in functions like planning of future behavior, short-term memory, and directing attention (cf. Posner & Peterson 1990); second, a limbic loop, involved in reinforcement and extinction of behavior (cf. Schultz 2002, Waelti et al. 2001); and third, a motor loop, seemingly involved in timing the starting and stopping of responses in the acquisition, retrieval, and relearning of programs for sequential motor tasks (Jog et al. 1999) and nondeclarative habit learning (discriminative stimulus–response–reinforcement relations: Knowlton et al. 1996).

The nucleus accumbens consists mainly of medium-spiny neurons that are surrounded by a “cloud” of glutamate and dopamine from 10,000 to 15,000 inputs from excitatory glutamatergic neurons and 5,000 to 6,000 inputs from dopaminergic neurons overlapping the glutamatergic inputs (Grace 2001). The main inputs are from the prefrontal cortex, the hippocampus subiculum (information about the behavioral context), the amygdala (affective information), and the ventral tegmental area (Grace 2001; Gray et al. 1991). The medium-spiny accumbens neurons exist in a bistable state, either hyperpolarized and non-firing, or at a depolarized plateau where action potentials are generated (Grace 2001). The hippocampus subiculum controls this bistable state and is, therefore, able to gate information from the prefrontal cortex to pallidum, thalamus, and back to neocortex (Grace 2001).

As reviewed by Grace (2001), the phasic and tonic dopamine components are functionally distinct and normally tightly regulated. Schultz suggests that the phasic component may be subdivided into two subcomponents: a fast (100 to 300 ms) component signaling erroneous “reward prediction,” and an intermediate subcomponent (lasting from seconds to minutes) involved in reinforcement, sex, movement, punishment, and stress (Schultz 2002). The phasic component releases dopamine as a brief pulse in association with an action potential, or spike. Released dopamine is rapidly removed from the synaptic cleft by the plasma membrane dopamine transporter (DAT1). The tonic dopamine level controls the phasic dopamine release via synthesis- and release-modulating autoreceptors on the dopamine terminals. Normally, the tonic extracellular dopamine...
pool is too low in concentration to stimulate postsynaptic dopamine receptors. However, the concentration is sufficient to provide a tonic down-modulation of action-potential-dependent dopamine release by stimulating (presynaptic) release- and synthesis-modulating dopamine autoreceptors on dopamine terminals. This causes a decrease in the action-potential-dependent dopamine release in the synaptic cleft (Grace 2001; 2002).

The tonic dopamine level is controlled by two sources: low concentrations of dopamine that has escaped from the synaptic cleft and glutamate released from (mainly prefrontal) cortical afferents in close proximity to the dopamine terminal. This glutamate stimulates close-by presynaptic heteroreceptors on the dopamine terminal to release dopamine from an intraterminal pool of vesicles directly into the extrasynaptic space. Normally, a low tonic dopamine level will lead to elevated action-potential-driven phasic dopamine responses (Grace 2002). An underdeveloped, immature, or hypoactive prefrontal cortex will reduce this glutamatergic input, resulting in abnormally low tonic dopamine levels in ADHD (Grace 2001; Solanto et al. 2001b). We suggest that ADHD is associated with a dysregulation of tonic/phasic dopamine control, causing stunned phasic dopamine responses (Russell et al. 1995) despite low tonic dopamine levels. This might be caused by several factors, for example, genetic mechanisms uncoupling this normally tight regulation.

There are several indications of anatomic and functional changes in the frontal lobes of ADHD (Castellanos 2001). In a series of studies, Castellanos and collaborators have been using automated methods measuring initial volumes and prospective age-related changes of total cerebral, cerebellum, gray and white matter for the four major lobes, and caudate nucleus of the brain in patients and controls. Patients with ADHD had significantly smaller brain volumes in all regions. This global difference was reflected in smaller total cerebral volumes and cerebellar volumes. Also, previously unmedicated children with ADHD demonstrated significantly smaller total cerebral and cerebellar volumes. Unmedicated children with ADHD also exhibited strikingly smaller total white matter volumes compared with controls and with medicated children with ADHD. Volumetric abnormalities persisted with age in total and regional cerebral measures and in the cerebellum. Caudate nucleus volumes were initially abnormal for patients with ADHD, but diagnostic differences disappeared as caudate volumes decreased for patients and controls during adolescence. Results were comparable for male and female patients on all measures. Frontal and temporal gray matter, caudate, and cerebellar volumes correlated significantly with parent- and clinician-rated severity measures within the ADHD sample (Castellanos et al. 2002). Because dopamine is involved in controlling cerebral circulation (Krimer et al. 1998), circulatory changes due to hypofunctioning dopamine systems may be one reason why brain-imaging studies have shown relatively global functional and structural differences between subjects with ADHD and controls.

2.5. Roles of dopamine in neuronal processes involved in reinforcement and extinction

The dopaminergic system has several branches: the mesolimbic and mesocortical branches originating in the ventral tegmental area, projecting to the nucleus accumbens septi, the olfactory tubercle (the mesolimbic branch), and to the prefrontal cortex (the mesocortical branch); and the nigrostriatal branch originating in the substantia nigra and projecting mainly to the striatum (Fig. 1). Imbalances in dopamine transmission in these branches will inevitably lead to imbalances in other neurotransmitter systems, producing specific behavioral effects related to the different systems and depending on situational fluctuations.

2.5.1. Reinforcement. Reinforcers are required both in acquisition and in maintenance of behavior. Reinforcement describes either a procedure (delivering a reinforcer) or a process (“strengthening” the likelihood that the reinforced response[s] will be repeated later in the same or a similar situation). A stimulus is a positive reinforcer if its presentation increases the probability of future occurrence of the response that produced it. The reinforcement contingencies are the conditions under which a response produces a reinforcer (Catania 1998).

The concept of reinforcer is strictly behavioral and makes no reference to subjective or cognitive states. The alternative concept of reward is more cognitive and connotes several subjective states like “pleasure” as well as “reinforcer” and “incentive” (Robbins & Everitt 1996). Therefore, there is not a perfect overlap between reinforcer and reward. We prefer the more descriptive and less ambiguously defined concept of reinforcer rather than reward.

A large body of evidence shows the importance of increased activity of the mesolimbic dopamine system, particularly in the nucleus accumbens, during reinforcement (Di Chiara & Imperato 1988; Robbins & Everitt 1996; Schultz 1998; 2002). This does not imply that dopamine activity is only, or always, involved in reinforcement. In general, dopamine is released in the nucleus accumbens, but not necessarily in the dorsal striatum, when novel associations between stimuli, or stimuli and responses can be formed (Datla et al. 2002). These stimuli may be reinforcers, but also seemingly neutral stimuli without apparent motivational or incentive value, even stressors or aversive stimuli. Only small increases in accumbal dopamine levels are produced by such stimuli when presented alone and out of context, or even by consuming a palatable reinforcer (Datla et al. 2002). Accumbal dopamine release is also seen when associations between two stimuli without apparent motivational value are formed (Young et al. 1998).

Dopamine neurons normally fire at a low tonic rate. Following a reinforcer, there is a phasic burst of activity of intermediate duration (Schultz 1998; 2002; Waelti et al. 2001) (Fig. 5). Reinforcement-induced burst firing of dopaminergic neurons produces a global dopamine signal that advances as a parallel wave of activity from the midbrain to the (ventral) striatum and the frontal cortex (Schultz 1998; 2002). Synaptically released dopamine diffuses out of the synaptic cleft and gives rise to transient peaks of extracellular dopamine concentrations (Schultz 1998). Thus, in the present theoretical framework, the burst of dopamine neuron activity seems to be linked to stimuli that function behaviorally as reinforcers. These reinforcers may either be primary or secondary (conditioned), scheduled or unscheduled (unpredictable, “free”) (Datla et al. 2002; Schultz 2002). The phasic burst activity following a reinforcer seems to occur whenever the delivery of this reinforcer deviates from the organism’s acquired behavioral relationships, for example, reinforcer delivery during ac-
The acquisition of novel behavior, delivery at an unusual time, or when the reinforcer has a higher-than-usual reinforcing value. The phasic dopamine activity level is gradually transferred to the earliest stimulus predicting future reinforcers. This stimulus is functioning behaviorally as a discriminative stimulus with secondary reinforcer properties (Schultz 2002). Apparently, there is no change in the tonic dopamine activity when stable-state behavior is established, and reinforcer deliveries are according to acquired stimulus–response–reinforcer relations (Schultz 1998; 2002; Waelti et al. 2001) (Fig. 5).

The duration of the time window available for coincidence detection with NMDA receptor stimulation is obviously critical for AMPA receptor mobilization and for subsequent LTP. NMDA-receptor-induced excitation necessary for LTP is enhanced by DRD1 receptor activation and attenuated by DRD2 activation (Cepeda et al. 1993; Kerr & Wickens 2001; Pedarzani & Storm 1995). Thus, DRD1-receptor activation may synergistically increase the excitatory actions of glutamate at NMDA receptors by increasing the opening time of NMDA receptors and therefore the time window available for coincidence detection. NMDA receptors are necessary for LTP in the hippocampus (Malenka et al. 1999), the corticostriatal synapses (Calabresi et al. 1997) and the nucleus accumbens (Kelley et al. 1997). Within the striatum, LTP (and LTD) only occurs in the presence of dopaminergic input (Grace 2002). Phasic application of dopamine in the prefrontal cortex facilitates LTP (Blond et al. 2002).

The dopamine-induced enhancement of synaptic transmission by accelerating LTP of the synapses in these areas is in accordance with a three-factor Hebbian learning rule. Synaptic transmission undergoes plastic changes when presynaptic (glutamatergic) input, postsynaptic activation, and the dopamine signal occur simultaneously at the same neuron. Thus, the homogeneous dopamine signal associated with reinforcement will selectively reinforce the weights of synapses that are active around the time of behavioral reinforcement (Wickens et al. 1996). At a systems level, dopamine exerts a focusing effect whereby only coincident inputs are reinforced and subject to LTP, whereas unsynchronized activity has no such LTP effect. Dopamine probably exerts its reinforcing effects by acting on D1-like receptors (DRD1 and DRD5; cf. Schultz 2002), stimulating adenylyl cyclase to produce cAMP that is essential for activation of PKA (cAMP-dependent protein kinase A). The resultant phosphorylation of CREB (cAMP response element binding protein), activation of gene transcription, mRNA, protein synthesis, and structural changes are required for memory consolidation (Bailey et al. 2000).

Using the hippocampal-slice preparation, it has been shown that the time frames of synaptic plasticity of burst activity in hippocampal CA1 pyramidal cells are closely similar to that of the equivalent behavioral phenomena. The spontaneous bursting of individual CA1 pyramidal neurons may be reinforced with activity-contingent injections of dopamine and cocaine, whereas CA3-bursting responses may be reinforced with contingently applied dynorphin A (Stein & Belluzzi 1989; Stein et al. 1983). Burst-contingent injections of the excitatory neurotransmitter glutamate failed to reinforce CA1 bursting. It is likely that dopamine acts as a “neurochemical reinforcer” through D1-like receptors in cellular models (Schultz 2002).
As a failure to inhibit responses. (Adapted from Waelti et al. 2001).

crease in the dopamine activity (D) is stunted in ADHD due to a
tivity changes. It is predicted that, in particular, the phasic de-
ing dopamine systems in ADHD result in stunted dopamine ac-
when a predicted reinforcer is not delivered (D). Hypofunction-
when a “free” reinforcer is delivered (C), whereas a decrease occurs
s in reinforcement schedules. An increased activity takes place when an unpredicted
when a predicted reinforcer is not delivered (D). Hypofunction-
ing dopamine systems in ADHD result in stunted dopamine ac-
activities. There is no depression of dopaminergic neuronal activ-
when a predicted reinforcer is not delivered (D). Hypofunction-
ing dopamine systems in ADHD result in stunted dopamine ac-
When a reinforcer is delivered according to an established rein-

Figure 6. Dopamine neurons normally fire at a low tonic rate. When a reinforcer is delivered according to an established reinforcement schedule, there is no change in dopamine activity either when the reinforcer is delivered (A), or when an omission of a reinforcer is signaled or predicted (B). Phasic changes are observed whenever there are unpredicted deviations in reinforcement schedules. An increased activity takes place when an unpredicted (“free”) reinforcer is delivered (C), whereas a decrease occurs when a predicted reinforcer is not delivered (D). Hypofunction-
ing dopamine systems in ADHD result in stunted dopamine ac-
activity. On a synaptic level, it might be predicted that reduced availability of dopamine will start an LTD process. The open time of the NMDA-receptor associated ion channels will be reduced and less calcium will be allowed to enter the cell. The reduced intracellular calcium levels will activate protein phosphatases removing phosphate groups from proteins, and start removing the AMPA receptors from the active zone of the synapse (Luscher et al. 2000). The associated LTP will no longer be maintained and will therefore gradually be reduced.

3. A dynamic developmental theory of ADHD

ADHD is currently defined as a cognitive/behavioral disorder with no biological marker. We will consequently offer a dynamic behavioral theory. To break the potential intrinsic circularity involved in explaining behavior by behavioral principles, we will suggest how this theory may be related to some of the presently less well-established genetic and neurobiological correlates to ADHD reviewed earlier.

The dynamic developmental theory of ADHD focuses on dopamine hypofunction because the majority of findings from a variety of research fields seem to converge on dopamine in the etiology of ADHD (Biederman & Faraone 2002; Castellanos & Tannock 2002; Grace 2001; Johansen et al. 2002; Sagvolden & Sergeant 1998; Sagvolden et al. 1998; Solanto et al. 2001b). The importance of other neuromodulators must not be underrated, however, and the present model may be applicable mainly to a subgroup of ADHD linked to dopamine hypofunction. However, in the genetics section we concluded that ADHD should be analyzed on a systems level, not on a single-gene, or a single-synapse level. It might well be that behavioral processes like reinforcement and extinction constitute the most elementary level at which it is possible to identify factors that are universal in a disorder like ADHD. It is likely that the development and severity of symptoms are linked to degree of dysfunction in the various dopaminergic systems.

The neuromodulator dopamine will regulate the processing of the information the brain receives via neurotransmitters like glutamate (Deutch & Roth 1998). Genetic links to ADHD do not represent mutations but polymorphisms that create subtle differences between normal and ADHD behavior. The theory offers an explanation of why ADHD is not a pathology that represents a separate entity with behavior qualitatively distinct from normal behavior but is a case where functions of the central nervous system occasionally exceed the limits of normal variation and adaptation.

2.5.2. Extinction. Procedurally and behaviorally, extinction is defined in relation to reinforcement. Discontinuation of reinforcer deliveries (actually, discontinuation of a reinforcement schedule) is termed an extinction procedure. This procedure starts an extinction process. The process has traditionally been understood as part of the process generated by reinforcement: Responding is maintained as long as reinforcers are delivered contingent on the responses and stops, or it is reduced to the level prior to reinforcement (the operant level), when this contingency is discontinued (Catania 1998). Thus, for operant behavior, extinction is the other side of reinforcement. Operant extinction may be the demonstration that the effects of reinforcement are temporary (cf. Catania’s precommentary in this issue of BBS.)

Neurobiologically, however, reinforcement and extinction may be separate processes associated with different aspects of dopamine neuronal activity. Depression of dopamine activity seems to occur when, in our terms, previously established stimulus—response—reinforcer relations are discontinued, (cf. Schultz 2002; Waelti et al. 2001). Extinction (and reinforcers with lower than previously experienced reinforcer value) is accompanied by a short-lasting (100 to 300 msec) phasic decrease in the tonic level of dopamine activity. There is no depression of dopaminergic neuronal activity, however, when an omission of a reinforcer is signaled (Fig 6) (Waelti et al. 2001). Thus, the extinction process depends on phasic depression of the tonic level of dopamine activity. On a synaptic level, it might be predicted that reduced availability of dopamine will start an LTD process. The open time of the NMDA-receptor associated ion channels will be reduced and less calcium will be allowed to enter the cell. The reduced intracellular calcium levels will activate protein phosphatases removing phosphate groups from proteins, and start removing the AMPA receptors from the active zone of the synapse (Luscher et al. 2000). The associated LTP will no longer be maintained and will therefore gradually be reduced.

3.1. Attention deficits

Attention encompasses highly multifaceted functions that are modified by a multitude of psychological factors like sensory and motivation processes. Excellent reviews of attention processes and networks are found elsewhere (Posner & Petersen 1990) and are outside the scope of the present review. It has been established that both dopamine and norepinephrine are important neuromodulators in attention processes (e.g., Arnsten 2001). The catecholamines may contribute to different aspects of attention processes. As our focus is on dopamine, we will suggest that both the
prefrontal loop and the limbic loop (Fig. 4) are involved in different aspects of attention.

The prefrontal loop is mainly involved in directing attention and selecting the behavior needed to achieve a given goal in a given situation (cf. Posner & Petersen 1990). It is suggested that a dysfunctions mesocortical dopamine branch will cause various attention deficiencies, such as inefficient orienting responses and abnormal control of eye saccades (Mostofsky et al. 2001), as well as poorer attention toward target (Kojima & Goldman-Rakic 1982). These problems will, in a developmental perspective, result in difficulties with controlling behavior and directing actions toward long-term accomplishments.

The limbic loop is mainly involved in reinforcement and extinction processes, the main components in the establishment of stimulus control and verbally governed (“rule-governed”) behavior (verbal instructions that regulate the behavior of the listener) (Catania 1998). Stimulus control is considered to be a prerequisite for the establishment of verbally governed behavior. A dysfunctioning mesolimbic dopamine branch will contribute to problems establishing these functions. A lack of stimulus control will be manifested in deficient sustained attention. Problems in establishing verbally governed behavior will result in difficulties with making and following plans. Thus, in our theory, the multifaceted attention problems of children with ADHD may be due to at least two different neurobiological systems related to dopamine dysfunction.

3.3. Clumsiness

Children with pervasive ADHD are more likely to show developmental delays in language (mainly expressive) and in motor functions, and to have an onset of symptoms in the first two years of life (Blondis 1999; Gillberg & Rasmussen 1982; Kadesjo & Gillberg 1999; Polataloko et al. 1995; Willoughby & Polataloko 1995; Taylor 1998; Teicher et al. 1996). A dysfunctioning nigrostriatal dopamine branch (Fig. 1) will probably cause several “extrapyramidal” symptoms (neurological “soft signs”) associated with ADHD in the form of clumsiness, that is, poor motor control, longer and more variable reaction times (Oosterlaan & Sergeant 1997). A dysfunctioning nigrostriatal dopamine branch (Fig. 1) will probably cause several “extrapyramidal” symptoms (neurological “soft signs”) associated with ADHD in the form of clumsiness, that is, poor motor control, longer and more variable reaction times (Oosterlaan & Sergeant 1997). Poor response timing, poor handwriting, poor correlation of the activity of different body parts, and so on. Also deficient nondeclarative habit learning and memory (Knowlton et al. 1996) might result from a dysfunction of the nigrostriatal dopamine branch. Thus, findings previously attributed to response disinhibition may rather be due to impaired motor control and to deficient nondeclarative habit learning and memory associated with dopamine dysfunction in the neostriatum (Sagvolden & Archer 1989; Sagvolden & Sergeant 1998).

3.3. Reinforcement and extinction

It is likely that a two-factor explanation of ADHD (delay aversion and response disinhibition) is better than a one-factor explanation (Solanto et al. 2001a; Sonuga-Barke 2002). The present dynamic developmental theory suggests that the delay aversion is associated with a dysfunctioning mesolimbic dopamine branch producing a shorter delay-of-reinforcement gradient (sect. 4.1). The response disinhibition may partly be rooted in an extinction deficit and partly be caused by a dysfunctioning nigrostriatal dopamine branch causing impaired modulation of motor functions in terms of poor timing of starting and stopping of responses, deficient acquisition, retrieval, and relearning of programs for sequential motor tasks.

As reviewed earlier, release in association with reinforcement is one of several dopamine functions. We will argue that this is a particularly important function for understanding ADHD. Further, reinforcement is associated with a phasic increase of dopamine activity (Schultz 1998; 2002; Waelti et al. 2001). Dopamine depletion of nucleus accumbens biases animals from instrumental responding for a normally highly preferred food to consumption of freely available, but normally less preferred food (Salamone et al. 1994). This behavior appears similar to ADHD children’s aversion to delayed reinforcers and preference for immediate reinforcers even when those have a lower value than reinforcers that are available after a delay (Sonuga-Barke 2002; Sonuga-Barke et al. 1992) and may indicate reduced accumbal dopamine functioning associated with reinforcers also in ADHD.

The postulation of a hypofunctioning dopamine system leads to several interesting predictions about reinforcement and extinction processes in ADHD if one assumes that the same phasic extracellular concentration of dopamine is required in the brains of children with ADHD as in normal children for reinforcement and extinction to occur. Compared to normal children, a reduced tonic dopamine level in children with ADHD will require an increased phasic release of dopamine to produce the postsynaptic changes required for normal reinforcement to take place. Similarly, normal tonic, but reduced phasic dopamine release associated with a reinforcer, will also result in less efficient reinforcement in ADHD. In both cases, an elevation of the reinforcer value is required to normalize the reinforcement process. These arguments are in accordance with the clinical observation that children with ADHD have a “motivation” problem: Stronger and more salient reinforcers are needed to control their behavior. They are also less sensitive to changes in reinforcement contingencies (Kollins et al. 1997). Assuming these underlying principles, it is unnecessary to predict general facilitating or inhibitory deficits associated with ADHD. Synapses that are active at the same time repeatedly, whether excitatory or inhibitory, are probably active because they participate in the same function.

As discussed earlier, extinction is associated with a phasic depression of tonic dopamine neuronal activity. We predict that abnormally low tonic dopamine activity associated with ADHD may cause failure of extinction, in particular of previously reinforced behavior, due to a “floor” effect (Fig. 6). Similar arguments have been forwarded by Wolfram Schultz: “Hypodopaminergic function will lead to a deficient prediction error and result in slower and less efficient learning” (Schultz 2002, p. 256). Hence, a hypofunctioning mesolimbic dopamine branch in ADHD may alter both reinforcement and extinction processes and thereby be the neurobiological basis of the altered reinforcement processes repeatedly suggested as one factor in ADHD symptomatology (Douglas 1983; Douglas & Parry 1994; Johansen et al. 2002; Sagvolden & Archer 1989; Sagvolden et al. 1995a; Sonuga-Barke 2002; Sonuga-Barke et al. 1992; Wender 1971). This suggestion is supported by several studies showing that the behavior of children with ADHD is differently affected by reinforcement contingencies (Douglas & Parry 1994; Sagvolden et al. 1998; Sonuga-Barke et al. 1992).
The theory predicts that symptoms like deficient attention processes and impaired motor functions may be caused by hypofunctioning dopaminergic loops. These symptoms will be modified by the altered reinforcement processes and deficient extinction, and develop dynamically as the child grows older, interacting with within-family factors and societal demands (see sect. 4).

Reinforcers act on responses that have already taken place by increasing the probability of future responding (Catania 1971; Catania et al. 1998). Thus, reinforcement is the selection mechanism in the evolution of behavior ontogenetically. Reinforcers may vary along several dimensions like density (frequency), the temporal response–reinforcer relationship (contiguity, delay of reinforcement), predictability, and value (attractiveness). The reinforcing effect is largest when the reinforcer is delivered immediately after the occurrence of the response and wanes as a function of the delay in reinforcer delivery. This relation between the effect of the reinforcer and the time interval between response and reinforcer is commonly known as the “delay-of-reinforcement gradient,” or simply as the “delay gradient” (Catania et al. 1988; Sagvolden et al. 1998) and may be expressed as a hyperbolic decay function of time (Johnson & Bickel 2002).

We have argued that a main component of the altered reinforcement process is a shorter and steeper delay-of-reinforcement gradient in ADHD (Fig. 7, left), implying that mainly responses in close proximity to the delivery of the reinforcer will be effective (Johansen et al. 2002; Sagvolden & Archer 1989; Sagvolden et al. 1989; 1998). In a novel situation, there will be a stream of spontaneously emitted, random responses of various kinds. If one of these, for example, $R_c$, is reinforced, the reinforcer will be less effective in ADHD than in normal children. This means that it is less likely that the child with ADHD will repeat the response than a normal child (though—as the stimulus is functioning as a reinforcer—the child with ADHD is more likely to repeat the response than he or she was before the reinforcer was delivered). Further, a reinforcer acts not only on the response that produced it, but, to a lesser extent, also on responses emitted earlier (Catania 1971). Thus, the response $R_c$ will be reinforced but to a lesser extent in a child with ADHD than in a normal child. The $R_{D}$ response will normally be reinforced, but it is outside the reach of the reinforcer when the delay gradient is short and steep.

The establishment of novel behavioral relations by reinforcement is essentially a matter of neuronal detection of coincident response–reinforcement or stimulus–response–reinforcement relations. Consequently, despite apparent differences in time scales, we suggest that the delay-of-reinforcement gradient neurobiologically is associated with the time window available for coincidence detection and thereby for mobilization of “silent” AMPA receptors necessary for LTP to take place (Malenka et al. 1999). Dopamine stimulation, as well as stimulation by other monoamines and estrogen, may increase the opening time of NMDA receptors and therefore the time window available for coincidence detection (Pedarzani & Storm 1995; Stein et al. 1993). Consequently, reduced dopamine function associated with ADHD produces shorter than normal time windows for coincidence detection resulting in a shorter delay gradient.

A multitude of processes contribute to reduction of responding: neurobiological factors associated with the extinction procedure, lack of maintenance of acquired responses, and acquisition of incompatible responses (Fig. 7, right). The dynamic developmental theory proposes that extinction is less efficient in ADHD than in normal children. This means that the normal elimination, in particular of previously established but no longer reinforced responses, will take place to a lesser extent in ADHD than in normal children. This view is consistent with studies finding excessive responding during extinction of previously reinforced responses in children with ADHD (Sagvolden et al. 1998). In a novel situation, there will be a stream of spontaneously emitted, random responses of various kinds. If one of these, for example, $R_c$, is reinforced, the reinforcer will be less effective in ADHD than in normal children. This means that it is less likely that the child with ADHD will repeat the response than a normal child (though—as the stimulus is functioning as a reinforcer—the child with ADHD is more likely to repeat the response than he or she was before the reinforcer was delivered). Further, a reinforcer acts not only on the response that produced it, but, to a lesser extent, also on responses emitted earlier (Catania 1971). Thus, the response $R_c$ will be reinforced but to a lesser extent in a child with ADHD than in a normal child. The $R_{D}$ response will normally be reinforced, but it is outside the reach of the reinforcer when the delay gradient is short and steep.

The establishment of novel behavioral relations by reinforcement is essentially a matter of neuronal detection of coincident response–reinforcement or stimulus–response–reinforcement relations. Consequently, despite apparent differences in time scales, we suggest that the delay-of-reinforcement gradient neurobiologically is associated with the time window available for coincidence detection and thereby for mobilization of “silent” AMPA receptors necessary for LTP to take place (Malenka et al. 1999). Dopamine stimulation, as well as stimulation by other monoamines and estrogen, may increase the opening time of NMDA receptors and therefore the time window available for coincidence detection (Pedarzani & Storm 1995; Stein et al. 1993). Consequently, reduced dopamine function associated with ADHD produces shorter than normal time windows for coincidence detection resulting in a shorter delay gradient.

A multitude of processes contribute to reduction of responding: neurobiological factors associated with the extinction procedure, lack of maintenance of acquired responses, and acquisition of incompatible responses (Fig. 7, right). The dynamic developmental theory proposes that extinction is less efficient in ADHD than in normal children. This means that the normal elimination, in particular of previously established but no longer reinforced responses, will take place to a lesser extent in ADHD than in normal children. This view is consistent with studies finding excessive responding during extinction of previously reinforced responses in children with ADHD (Sagvolden et al. 1998). In a novel situation, there will be a stream of spontaneously emitted, random responses of various kinds. If one of these, for example, $R_c$, is reinforced, the reinforcer will be less effective in ADHD than in normal children. This means that it is less likely that the child with ADHD will repeat the response than a normal child (though—as the stimulus is functioning as a reinforcer—the child with ADHD is more likely to repeat the response than he or she was before the reinforcer was delivered). Further, a reinforcer acts not only on the response that produced it, but, to a lesser extent, also on responses emitted earlier (Catania 1971). Thus, the response $R_c$ will be reinforced but to a lesser extent in a child with ADHD than in a normal child. The $R_{D}$ response will normally be reinforced, but it is outside the reach of the reinforcer when the delay gradient is short and steep.

A multitude of processes contribute to reduction of responding: neurobiological factors associated with the extinction procedure, lack of maintenance of acquired responses, and acquisition of incompatible responses (Fig. 7, right). The dynamic developmental theory proposes that extinction is less efficient in ADHD than in normal children. This means that the normal elimination, in particular of previously established but no longer reinforced responses, will take place to a lesser extent in ADHD than in normal children. This view is consistent with studies finding excessive responding during extinction of previously reinforced responses in children with ADHD (Sagvolden et al. 1998). In a novel situation, there will be a stream of spontaneously emitted, random responses of various kinds. If one of these, for example, $R_c$, is reinforced, the reinforcer will be less effective in ADHD than in normal children. This means that it is less likely that the child with ADHD will repeat the response than a normal child (though—as the stimulus is functioning as a reinforcer—the child with ADHD is more likely to repeat the response than he or she was before the reinforcer was delivered). Further, a reinforcer acts not only on the response that produced it, but, to a lesser extent, also on responses emitted earlier (Catania 1971). Thus, the response $R_c$ will be reinforced but to a lesser extent in a child with ADHD than in a normal child. The $R_{D}$ response will normally be reinforced, but it is outside the reach of the reinforcer when the delay gradient is short and steep.

The establishment of novel behavioral relations by reinforcement is essentially a matter of neuronal detection of coincident response–reinforcement or stimulus–response–reinforcement relations. Consequently, despite apparent differences in time scales, we suggest that the delay-of-reinforcement gradient neurobiologically is associated with the time window available for coincidence detection and thereby for mobilization of “silent” AMPA receptors necessary for LTP to take place (Malenka et al. 1999). Dopamine stimulation, as well as stimulation by other monoamines and estrogen, may increase the opening time of NMDA receptors and therefore the time window available for coincidence detection (Pedarzani & Storm 1995; Stein et al. 1993). Consequently, reduced dopamine function associated with ADHD produces shorter than normal time windows for coincidence detection resulting in a shorter delay gradient.

A multitude of processes contribute to reduction of responding: neurobiological factors associated with the extinction procedure, lack of maintenance of acquired responses, and acquisition of incompatible responses (Fig. 7, right). The dynamic developmental theory proposes that extinction is less efficient in ADHD than in normal children. This means that the normal elimination, in particular of previously established but no longer reinforced responses, will take place to a lesser extent in ADHD than in normal children. This view is consistent with studies finding excessive responding during extinction of previously reinforced responses in children with ADHD (Sagvolden et al. 1998).
al. 1998) as well as in an animal model of ADHD (Sagvolden 2000). It is also consistent with studies showing that children with ADHD are not hyperactive in novel situations (Sagvolden et al. 1998; Sleator et al. 1981). We suggest that a failure to inhibit responding (Barley 1997b) in most cases is the result of a faulty extinction process.

3.4. Overactivity

Introducing a reinforcer may lead to induction (response generalization), which is a general increase in responding. Responses may be defined either as belonging to a descriptive or nominal class (the responses that are reinforced), or a functional class (all the responses generated by reinforcement). During differentiation, responding gradually becomes more restricted to the nominal class producing the reinforcer (Catania 1998), that is, the more frequent than other responses (Fig. 7). At a neurobiological level, both the phasic increase in dopamine release associated with reinforcement and the phasic decrease in dopamine neuronal activity associated with extinction may be necessary for efficient differentiation of responses. In ADHD, the establishment of functional response classes and differentiation may be inefficient due to the less effective extinction of behavior. On a behavioral level, responses in general will be induced resulting in an increased frequency of all responses in the functional class without the normal differentiation into the nominal response class.

The dynamic developmental theory predicts that the failing extinction process in ADHD will result in an increased number of responses, as well as an increased behavioral variability (see sect. 3.5), despite a reduced effect of each reinforcer. Altered reinforcement processes characterized by a shorter delay gradient in ADHD will not by itself generate the gradually developing overactivity. It is hypothesized that the ADHD overactivity is acquired and maintained by a combination of scheduled and unscheduled reinforcers and failing extinction, increasing the frequency of acquired responses without the pruning of ineffective and inadequate responses (Fig. 7). The deficient extinction process will lead to an accumulation of responses that may be seen as excess motor activity where no reinforcer can be identified (cf. Porrino et al. 1983; Sagvolden et al. 1998a; Teicher et al. 1996). An increased number of responses with short inter-response times (motor impulsiveness) in ADHD is also contributing to the overactivity (see sect. 3.6).

3.5. Increased behavioral variability

Clinically, ADHD behavior varies according to situational and task characteristics (American Psychiatric Association 1994). Experimentally, it has been shown to be more variable than normal (Kinsbourne 1990; Rubia et al. 1998; Scheres et al. 2001; Teicher et al. 1996). Variability acts as an operant that may be modified by reinforcers (Mook et al. 1993; Saldana & Neuringer 1998). Just as variability in the form of spontaneous mutations is necessary for evolution to take place, so is variability of spontaneously emitted behavior necessary for the emergence and shaping of new behavior (Catania 2000). According to the dynamic developmental theory, a combination of a general induction of responding and inefficient response differentiation due to a deficient extinction process in ADHD will result in an increased number of slightly different responses in the functional class and hence increased behavioral variability (Fig. 7). This means that normal children's responding increasingly will be within the nominal response class (i.e., the class of responses that generates reinforcement) (Catania 1998) and inefficient responses (responses that do not generate reinforcement) will be extinguished. However, the behavior of children with ADHD will continue to include responses outside the nominal class. In addition, a response accidentally occurring just before the delivery of a reinforcer may quickly be part of the behavioral repertoire of a child with ADHD (cf. superstitious behavior, Skinner 1948) and not be extinguished in spite of lack of subsequent reinforcement.

An efficient reinforcer may select short sequences of behavior that function well under one set of circumstances, like during learning of new material when the situation is motivating. But, as situations change, the behavior of a child with ADHD will not change accordingly, and the learned behavior will not adapt to changes in the reinforcement contingencies (e.g., Kollins et al. 1997). Therefore, as the child may seem to function well under one set of conditions, the lack of adaptability of behavior to slight changes in the environment will be characterized as dysfunctional by the surroundings. The dynamic developmental theory of ADHD may thus explain the common observation that ADHD behavior is quite variable.

3.6. Impulsiveness

Impulsiveness is often exemplified by the choice of a small or less attractive reinforcer that is available immediately, in preference to a larger but delayed reinforcer. Selective lesions of the nucleus accumbens core induce persistent impulsive choice in rats. In contrast, damage to two afferents of the nucleus accumbens, the anterior cingulate cortex and the medial prefrontal cortex, does not increase impulsiveness (Carli et al. 1985). Thus, dysfunction of the nucleus accumbens core, and therefore reinforcement functions, may be a key element in the neuropathology of impulsiveness.

Not only single responses (e.g., R_C, see Fig. 8, top), but also the relationships between responses (e.g., inter-response times, IRTs, see Fig. 8, bottom) are conditioned and maintained by reinforcers (Catania 1971; Catania et al. 1998; Sagvolden et al. 1998). In contrast to the normal delay gradient, only short IRTs are reinforced and maintained by a short delay gradient. Only the normal gradient is long enough to reinforce the long IRT involved in the sequence R_D – R_C (Fig. 8, bottom). This reinforcement process explains why motor impulsiveness, responses emitted with short IRTs, is not present in a novel situation but develops gradually as more reinforcers modify the behavior (Sagvolden et al. 1998a). In addition, because the normally occurring medium and long IRTs necessarily will reduce the overall behavioral output, the selective reinforcement of short IRTs due to a short delay gradient probably explains a substantial part of the ADHD overactivity.

The importance of reinforcement in impulsive behavior is supported by the fact that children with ADHD are not always impulsive as they temporarily do manage to plan ahead, organize themselves, and remember important things, if this behavior is maintained by potent and frequent reinforcers (Douglas 1999). Further, impulsiveness is not unique to ADHD. All children are impulsive as infants and young toddlers (Fig. 9).
Behavior is gradually brought under discriminative control, including the establishment of verbally governed behavior, as a function of training (Barkley 1997b; Catania 1998). Verbally governed behavior, or the control over behavior by verbal stimuli, will be gradually established as the child enters the verbal community (i.e., learns to understand speech and to speak). Verbal stimuli controlling future nonverbal behavior includes instructions, directions, demands, requests, urges, and written or spoken rules or norms for conduct in specified situations (see Skinner 1957). Verbal stimuli may also be defined as contingency-specifying stimuli, as they often describe the situation in which a specific behavior is warranted (e.g., “at the second intersection, turn left”) (Catania 1998). Very briefly, the establishment of verbally governed behavior goes on continuously from (but probably even before) the child learns to name objects and to use object names to get what they want. Simultaneously, the parents introduce instructions (e.g., “Look at me”) and immediately reinforce compliant behavior (e.g., “Good!”). A further step is through play and interaction with other children and parents where overt verbal directions for actions play a central part (“I go to your house and you open the door”). Gradually, more sophisticated, covert verbal self-talk directs more behavior over longer time periods. This account is in line with Vygotsky’s theory of the development of private speech (Vygotsky 1978) and internalization of speech (Winsler & Naglieri 2003).

The development of longer sequences of behavior and establishment of verbally governed behavior will be hampered by a short delay-of-reinforcement gradient. Therefore, both in normal and in children with ADHD, impulsiveness will be reduced as they grow older, but this process is stunted in children with ADHD. ADHD impulsiveness will consequently be manifested differently at different ages. Motor impulsiveness (bursts of responses) is predominant in infants and young toddlers, while cognitive impulsiveness (poor verbal control of behavior) is more prevalent in older children and adolescents. Clinically, this will mean that diagnosing ADHD at very early ages will be difficult partly because impulsiveness is typical of all young children’s behavior. Thus, ADHD impulsiveness may be understood as a maturational lag with later achievement of language milestones, simpler expressive language, impaired sensory-motor coordination, poor handwriting, and reading ability that are all behind that which is expected for this child’s chronological age (Saugstad 1994a; 1994b; Taylor et al. 1998).

### 3.7. Impaired sustained attention

Attention denotes the control over behavior by some stimulus features and not by others (Catania 1998). Sustained attention means that this stimulus controls behavior over time.

The establishment of the relation between a discriminative stimulus, behavior, and the reinforcement contingency (i.e., the three-term contingency) is a prerequisite for stimulus control. Stimulus properties and reinforcer timing are the two essential factors. First, important stimulus properties include the contribution of new and significant information about reinforcement. If behavior is already controlled by one stimulus, the behavioral effects of adding a new stimulus is “blocked” that is, behavior will not be controlled by this new stimulus (Catania 1998). Neurophysiologically, blocking is seen in the lack of a phasic dopamine response if the added stimulus is later presented alone.
Second, the introduction of a reinforcer must be contingent on the behavioral changes following stimulus changes for the behavior to be related both to the stimulus and the reinforcer. The three-term contingency will not be established if the onset of the discriminative stimulus is outside the reach of the reinforcer. The potency of a stimulus as a conditioned reinforcer depends on the time between its onset and the subsequent delivery of a reinforcer in its presence, according to the same delay gradients that operate for the relation between responses and subsequent reinforcers (Figs. 7 and 8). Ordinarily, the delay gradient decreases slowly enough that stimuli become effective even when many seconds pass between their onset and the reinforcer (Fig. 10, top). But if the gradient is short, the reinforcer must follow quickly after stimulus onset. If not, the stimulus does not become a potent reinforcer, and the individual will not attend to it when it appears (Fig. 10, middle). Thus, the sustained attention deficit is derived from the same source as hyperactivity. The dynamic developmental theory predicts that the delay gradient will bridge longer time intervals in normal children (Fig. 10, top) than in children with ADHD (Fig. 10, middle).

The deficit in sustaining attention may be remedied by increasing the reinforcer frequency and thereby increasing the probability of bringing the stimulus onset within reach of the delay-of-reinforcement gradient (Fig. 10, bottom). A short delay gradient implies that poorly sustained attention in ADHD is seen whenever the frequency of reinforcers is too low for stimulus control to be properly established in children with ADHD, but high enough for such control of normal children’s behavior. Further, it is predicted that normal children will also show lack of sustained attention if the frequency of reinforcers is very low. In addition, there will be individual differences both among children with ADHD and among normal children, because stimulus control by distanced stimuli is a result both of individual dopamine levels and individual learning histories.

As Catania describes in more detail in the precommentary that follows this article, according to our theory, individual differences between delay gradients in children with ADHD will give rise to differences in symptoms. Consider a child whose gradient quickly drops to zero. Then, responses must be very close to the reinforcer to be captured by it, and only single responses will be strengthened (Fig. 8, top). If chains of responses are not strengthened, there will be little motor impulsiveness (Fig. 8, bottom). In this instance, there should be a profound sustained attention deficit, because only brief stimuli quickly followed by reinforcers will acquire any conditional reinforcing effectiveness.

Now consider a child with ADHD with a somewhat longer delay gradient. This child is likely to show a lot of responses with short IRTs (impulsiveness). Poor stimulus control (attention deficit) is still likely to be a problem. Therefore, in this instance, we can expect to see both impulsiveness and sustained attention deficit.

If we were able to lengthen the delay gradient, for example, by medication (see sect. 3.8), the longer time period would mean that sustained attention deficit would be less of a problem (discriminative stimuli in the presence of which reinforcers become available soon enough would acquire conditional reinforcing properties of their own, and therefore they would be attended to when they appeared), but the IRTs that can be captured by the reinforcer would still be shorter than some of those captured by a normal gradient. In this case, the individual would probably show mild impulsiveness and mild attention deficit, with the former dominant.

3.8. Effects of stimulants

A review of the effects of stimulant drugs is outside the scope of this target article. In brief, psychostimulants such as methylphenidate and amphetamines have for several decades provided the primary pharmacological treatment for ADHD (Bradley 1937; Solanto et al. 2001b; Vitiello et al. 2001). Methylphenidate probably acts by blocking
However, electrically stimulated release of dopamine synaptic dopamine receptors (Seeman & Madras 2002). Gerded dopamine release, resulting in less activation of post-mine activates DRD2 autoreceptors that inhibit dopamine dopamine concentrations, because the increased dopa-malized by low doses of psychostimulants. Low doses of psychostimulants primarily affect the tonic dopamine level that is increased several-fold (Seeman & Madras 2002). Consequently, the therapeutic effect of psychostimulants may be mediated by an increase mainly in the tonic level of dopamine activity, thereby improving reinforcement and extinction on a behavioral level. However, the exact mechanisms of action of stimulant drugs are not known (Solanto et al. 2001b) and may differ across brain regions (Porrino & Lucignani 1987; Russell et al. 1998).

The mechanism by which psychostimulants alleviate ADHD symptoms may be by increasing tonic extracellular dopamine concentrations, because the increased dopamine activates DRD2 autoreceptors that inhibit dopamine release and reduce the amplitude of action-potential-triggered dopamine release, resulting in less activation of post-synaptic dopamine receptors (Seeman & Madras 2002). However, electrically stimulated release of dopamine in vivo is not in fact reduced; it is increased by low, clinically relevant doses of d-amphetamine (Parker & Cubeddu 1986; Seeman & Madras 2002; Suaud et al. 1989), which would support the dopamine hypofunction hypothesis of ADHD. In addition, extrasynaptic dopamine may be required to act at more distant DRD4 heteroreceptors to inhibit glutamate release from corticostriatal afferents (Berger et al. 2001; Tarazi et al. 1998). We suggest that inappropriate overactivity of mesolimbic VTA dopamine neurons at an early stage of development of ADHD could activate DRD5 receptors on dendrites of VTA dopamine neurons and increase expression of functional NMDA receptors in VTA dopamine neurons. Increased NMDA function could give rise to compensatory changes that would result in depolarization block of VTA dopamine neurons and hypoactivity of the mesolimbic dopaminergic system.

4. ADHD in a developmental perspective

In a developmental perspective, one has to consider the child's behavioral characteristics, the neurobiological development during the child's life, and the interplay between these two factors and the environment (Karniloff-Smith 1998). This interplay is not unidirectional and will have many different outcomes, as the capacity for learning and change is life-long. Herein lies also the possibility that a caregiver may adjust the environment to the child's needs for optimal development of adaptive skills. These skills may, of course, develop into a behavioral style with which the world is met, determining the long-term consequences of the initial interplay between the child and the environment.

At a neurobiological level, all neurotransmitter and neuromodulator systems undergo growth spurts and pruning several times during ontogenesis (Andersen et al. 1997; Saugstad 1994a; 1994b). The growth spurts and pruning will be associated with synaptic supersensitivity and therefore associated with enhanced vulnerability to negative as well as positive environmental (parental, familial, and societal) influences. Considering the neurobiological bases of acquisition and maintenance of behavior at such critical stages in the individual's neurodevelopmental history, the environment may influence symptom development in either a negative or positive direction.

ADHD in a neurodevelopmental perspective is a vast topic. So far we have considered the dynamic interplay among neurobiological processes, environmental events, and behavior. The following subsections will discuss some aspects of the dynamic developmental of the behavior of the child with ADHD on a macro level, taking into account behavioral and environmental properties and principles. We will limit the discussion to the most important predictions for within-child factors (sect. 4.1), then proceed to consider these factors in a family (sect. 4.2) and a societal perspective (sect. 4.3), pointing out important relations and how they can lead to different short- and long-term consequences. The discussion is summarized in Figure 2.

4.1. Within-child factors

There is substantial evidence for a neurobiological predisposition in ADHD. Increasing amounts of genetic, neurobiological, and neuropsychological data support the biological underpinning of the disorder (Wilens et al. 2002a; 2002b). In addition, ADHD is often chronic with prominent symptoms and impairment spanning into adulthood. ADHD is often associated with co-occurring anxiety, mood, and disruptive disorders, as well as substance abuse (Wilens et al. 2002a; 2002b). The neurobiological predisposition can be viewed as a risk factor or vulnerability for maladjustment. In the dynamic neurodevelopmental theory of ADHD, the vulnerability consists, in particular, of inefficient reinforcement and extinction processes.

A short and steep delay-of-reinforcement gradient implies that reinforcement should be immediate to be effective. As discussed earlier, the short delay gradient and impaired extinction may cause impulsiveness and hyperactivity and hamper the establishment of stimulus control and verbally governed behavior. Disrupted discriminative control of behavior will result in developmental delays in several areas of daily life. A young child with ADHD will have problems with learning the relationships between situational or instructional demands and the child's own behavior, and will thus receive little reinforcement for compliant behavior. As the child grows older, he or she will have problems with anticipating the proper behavior for a given situ-
A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD)

Sagvolden et al.

BEHAVIORAL AND BRAIN SCIENCES (2005) 28:3

416

Early-onset CD almost invariably occurs in combination with ADHD (Pliszka 1999). It is not yet clear whether the combined ADHD-CD case is a separate disorder, or a more severe case of ADHD. What seems to be the case is that late-onset CD (with or without ADHD) probably is a product of psychosocial influence, whereas early-onset CD (which never occurs without ADHD) is genetically based. The interactions are not simple. The probability of developing CD from early oppositional behavior seems to be mediated by high levels of socioeconomic disadvantage and negative family climate, whereas this probability is almost absent given low levels of these risk factors (McGee & Williams 1999). Parent–child conflict appears to act as a common vulnerability that increases risk for multiple childhood disorders. Furthermore, the association between parent–child conflict and childhood disorders is mediated via common genetic and environmental factors. These findings support the idea that the comorbidity among these disorders partially reflects core psychopathological processes in the family environment that link putatively separate psychiatric disorders (Burt et al. 2003).

Most children with ADHD and comorbid CD also meet criteria for ODD, which usually precedes CD onset by several years. However, there seem to be two subtypes of ODD associated with ADHD: one that is prodromal to CD and another that is subsyndromal to CD but not likely to progress into CD in later years (Biederman et al. 1996; Jensen et al. 2001). The possibility of a combination of disruptive behavior being reinforced by its short-term consequences and deficient extinction in ADHD is severe. In ADHD with co-occurring disruptive behavior, the short-term consequences of lying, stealing, threatening, and the like, can reinforce and maintain the disruptive behavior. In the dynamic developmental theory, we predict that a subgroup of the disruptive behavior disorders is caused by the core deficits involved in ADHD and hence is secondary to ADHD (same etiology). Thus, this behavior may also be controlled by a short delay-of-reinforcement gradient. Law-breaking behavior is often associated with lack of proper verbally governed behavior, sensation seeking, and substance use (Rasmussen et al. 2001). The extinction problem in ADHD will add to the negative effect, as the deviant behavior is not easily reduced by punishment or lack of reinforcement. The societal actions like punishment and prison will be ineffective and in the long term possibly lead to an elevated prevalence of criminal offense in persons with ADHD symptoms (Crowley et al. 1998).

High heritability or a neurobiological basis does not imply determinism. The “positive” or successful adult with ADHD might have had insightful teachers and parents who understood the importance of immediacy of reinforcers and computer-assisted instruction. As adults, children with ADHD may very well end up with a type-A-like personality (Whalen & Henker 1986), directing activity toward work, being creative and relatively well adapted, although they might be easily stressed and develop hypertension.

The dynamic developmental theory explains why the severity of the behavioral problems of individuals with ADHD varies tremendously, not just between persons, but also within individual persons, as they encounter changing situations with differing contingencies operating. The variability is enhanced by the nature of the long-term neurodevelopmental changes caused by dopamine influences where the time scale is not milliseconds, but rather seconds and minutes (Byrne 1998). This fact explains why people with ADHD can stay focused when high densities of reinforcement or potent reinforcers are operating, for example, when playing video games or performing hazardous acts. Then the reinforcers may release enough dopamine and computer-assisted instruction. As adults, children with ADHD may very well end up with a type-A-like personality (Whalen & Henker 1986), directing activity toward work, being creative and relatively well adapted, although they might be easily stressed and develop hypertension.
central nervous system within normal functional range without medication. Increased release of dopamine might be a part of a sensation-seeking behavior associated with ADHD (Blum et al. 1995; Petry 2001). Substances of abuse also increase dopamine levels (Di Chiara & Imperato 1988), which might be an important aspect of self-medication too often leading to substance abuse associated with ADHD (Biederman et al. 1999).

4.2. Family interactions and parenting style
The dynamics of family interaction are influenced both by behavioral characteristics of the child and the parenting style of the child's primary caretakers. A child with ADHD affects the family interaction in ways other than normally developing children. Research indicates that the presence of ADHD in a child is associated with disturbances in family and marital functioning, disrupted parent–child relationships, reduced parenting self-efficacy, and increased levels of parental stress (DuPaul et al. 2001; Johnston & Mash 2001).

Genetics, not family environments, produce ADHD (Rey et al. 2000). However, negative emotional family environments predispose unfavorable behavioral development in a child with ADHD (Hinshaw et al. 2000) and increase the risk of later ODD and CD (Biederman et al. 2001; Taylor 1999), particularly in boys (Biederman et al. 2002a). Biederman and colleagues showed that a number of risk factors like low social class, maternal psychopathology, and family conflict were associated with a greater risk for ADHD and other comorbidity in a “dose-dependent” fashion, irrespective of gender, parental ADHD, and maternal smoking during pregnancy (Biederman et al. 2002a). A possible developmental trajectory is outlined and evidenced in coercion theory for the development of antisocial behavior in children (e.g., Patterson 2002). This theory explains how coercive behavior develops through reinforcement processes: The child's nagging is reinforced when the parent gives in, and the parent's behavior of giving in is reinforced by the removal of nagging, that is, negative control. According to the dynamic developmental theory, this behavior, once established, is harder to extinguish in the child with ADHD than in other children.

Having a child with ADHD requires exceptional parenting skills. Preschool children with ADHD are at an early age (typically 3 to 5 years) rated by their parents as showing more noncompliant and inappropriate behavior, and they are significantly more aggressive, more demanding of parental time, less socially skilled, and less adaptable to change in routine, as compared with parent ratings of normally developing children (DeWolfe et al. 2000; DuPaul et al. 2001). To secure an optimal upbringing, caregivers have to adapt to the needs of the child with ADHD by taking into account the implications of the underlying deficits and adjust expectations and demands to the child's functional age (Barkley 1998). Thus, in addition to coping with ongoing challenging behavior, the altered reinforcement and extinction processes require parents to behave in a consistent and organized way toward their child. This includes reinforcing adaptive behavior by frequent and immediate reinforcers and at the same time not allowing maladaptive behavior to develop. However, 15%–20% of the mothers and 20%–30% of the fathers may also have ADHD themselves. Furthermore, parents of children with ADHD often show conduct problems and antisocial behavior (~25%), alcoholism (14%–25%), histrionic or affective disorder (10%–27%), or learning disabilities (Barkley 1998). Thus, parents with any of these problems will have even greater difficulty than other parents in coping with the special needs of their children with ADHD.

Parental ADHD is associated with a disruptive family environment, which increases the risk of a negative outcome in the child with ADHD (Biederman et al. 2002b). Weiss and coworkers (Weiss et al. 2000) have suggested several ways that adult ADHD may influence parenting skills: reduced patience with and responsiveness to the child; difficulty maintaining attention during supervision; difficulty remembering or keeping appointments with day care or school; difficulty with instrumental and organizational tasks like remembering birthday parties, activities, or play dates; problems with disengaging emotionally in their child's temper tantrum and instead contribute to escalation; and difficulties with organizing both domestic duties and care for the child. Fathers with ADHD use less effective discipline toward their child with ADHD than fathers without ADHD (Arnold et al. 1997). In these circumstances the parent will not be able to create a predictable environment for the child, where certain behaviors consistently are followed by certain consequences.

Supporting this, maternal ADHD has been shown to be the sole factor accounting for lack of change in child ADHD after intensive parent training while the presence of ADHD symptoms in the child was significantly and long-lastingly (15 weeks) reduced when mothers scored low on ADHD symptoms (Sonuga-Barke et al. 2002). The long-term consequences of an upbringing characterized by inconsistency, impulsiveness, and disorganization are grave compared to a well-structured environment. A corollary of this reasoning is that the situation may improve if the ADHD parent was allowed adequate medication (Fig. 2) in addition to attending parent-training programs.

In the framework of our theory, a normal parent will have a long delay-of-reinforcement gradient and good stimulus control in the sense that she or he can verbalize the rules applicable in a certain situation and behave accordingly (Fig. 11, upper left). Combined with an understanding of the need for frequent and immediate reinforcers, the dynamic developmental theory predicts that establishment of stimulus control is possible (Fig. 11, lower left). When the parent also has ADHD, it is likely that there is deficient stimulus control, and she or he may have poor verbally governed behavior (Barkley 1997b) (Fig. 11, upper right). In this case, establishment of adequate stimulus control in the child with ADHD will be unlikely (Fig. 11, lower right).

4.3. Societal style
From time to time, professionals and lay people suggest that ADHD is a product of the Western way of life where events happen quickly, contingencies change incessantly, and reinforcers never have to be postponed. Such allegations are contradicted by research showing that ADHD is found in all kinds of cultures around the world (e.g., Meyer et al. 2004). This is not to say that societies do not create influential contingencies for its inhabitants. In the dynamic developmental theory, societal style is predicted to influ-
ence the behavior of people by the prevailing "culture" of, for example, child upbringing and in the way disorders and disabilities are defined. In Western cultures, children are allowed to behave in certain ways when they are young ("Let him keep on, he is just a child!") but when the child gets older, unwanted behaviors are supposed to extinguish (by parenting practices like rule learning, lack of reinforcers, punishment, and ignoring). A child with ADHD in a Western culture will have acquired quite a lot of the behavior described as unwanted when young, but combined with the ADHD extinction deficit, getting rid of it will be difficult. Other cultures with a stricter child upbringing than is common in Western countries may see less maladaptive behavior and lower prevalence of ADHD because disruptive behavior is not accepted even in very young children (Meyer et al. 2004).

The severity of the behavioral problems of children with ADHD varies. Approximately 50% have significant problems in social relationships with other children (Bagwell et al. 2001). Not only the parents but also the society in general interact with the child and shape its behavior. The society requires that its inhabitants develop adequate self-control, learn to use time efficiently, learn to foresee consequences of their behavior to socialize, obtain an education, and get a job. All these requirements are very difficult for people with ADHD. Behavioral training programs may generate optimal environments with frequent and immediate reinforcers as well as short and clear instructions. For example, in the multimodal treatment study of ADHD (MTA) the children receiving either only intensive behavioral treatment or the combination of medication and behavioral treatment started the treatment period with an eight-week summer school program (Pelham et al. 2000). Here, all children continuously received reinforcers for proper, prosocial behavior; rules of conduct were explicit and frequently repeated; and violations to the rules resulted in predictable consequences. Behavior was evaluated by parents, and there were no differences between the children that received medication in addition to the intensive behavioral treatment and the children that only had the behavioral treatment package; and they all showed significant improvement over a range of behaviors (Pelham et al. 2000).

The problem is that optimal contingencies only exist during the training session or under certain circumstances. Outside these, inconsistent and unpredictable contingencies are the rule. The school may, however, help a child with ADHD to adjust to the school requirements by creating an optimal learning environment (Hoffman & DuPaul 2000). Such an environment should include structure, clear instructions, and frequent reinforcers to establish stimulus control and verbally governed behavior. Programs like "Positive Behavior Intervention and Support" (e.g., Wolf 1998) specifically seek to optimize these contingencies by increasing reinforcer density and clarify rules for preventing and treating conduct disorders on a school-wide basis. The effect of such programs on the behavior of children

![Figure 11. An abnormally steep and short delay gradient will result in poor stimulus control when reinforcers are infrequent both in children and adults with ADHD, but not when the density of reinforcement is high enough for the three-term contingency to work. Accordingly, the dynamic developmental theory predicts that it is possible to establish stimulus control in ADHD by presenting the reinforcers frequently. A normal parent will have a long delay-of-reinforcement gradient and good stimulus control in the sense that she or he can verbalize the rules applicable in a certain situation (upper left). Combined with frequent and immediate reinforcers, establishment of stimulus control and verbally governed behavior in a child with ADHD may be possible (lower left). In the case when the parent also has ADHD, there is poor stimulus control, and she or he may have poor verbally governed behavior (upper right). Under such circumstances, there will be poor stimulus control, and a child with ADHD will probably not establish verbally governed behavior (lower right).](image-url)
with ADHD has yet to be established empirically, but according to our theory programs built on the principles listed previously should improve their level of functioning.

5. Conclusions

The dynamic developmental theory for the ADHD predominantly hyperactive/impulsive and combined subtypes is based on the hypothesis that altered dopaminergic function plays a pivotal role by failing to modulate non-dopaminergic (primarily glutamate and GABA) signal transmission appropriately. Genetic links to ADHD do not represent mutations, but polymorphisms.

1. The theory offers an explanation of why ADHD is not a pathology that represents a separate entity with behavior qualitatively distinct from normal behavior, but is a case where the function of the central nervous system occasionally exceeds the limits of normal variation and adaptation.

2. A dysfunctioning mesolimbic dopamine branch will produce altered reinforcement of behavior and deficient extinction of previously reinforced behavior. This will, on a behavioral level, give rise to delay aversion, development of hyperactivity in novel situations, impulsiveness, deficient sustained attention, increased behavioral variability, and failure to “inhibit” responses (“disinhibition”). It might be that the disorder in the future should be named RED (reinforcement/extinction disorder).

3. A dysfunctioning mesocortical dopamine branch will cause attention response deficiencies (deficient orienting responses, impaired saccadic eye movements, and poorer attention responses toward a target) and poor behavioral planning (poor executive functions).

4. A dysfunctioning nigrostriatal dopamine branch will cause impaired modulation of motor functions (poor timing of starting and stopping of responses, deficient acquisition, retrieval, and relearning of programs for sequential motor tasks) and deficient nondeclarative habit learning and memory. These impairments will give rise to apparent developmental delay, clumsiness, neurological “soft signs,” and “failure to inhibit” responses when quick reactions are required.

5. The theory predicts that symptoms will in part be produced by deficient regulation of attention and impaired motor functions. These symptoms will develop as a result of the altered reinforcement processes and deficient extinction and be dynamically modified as the child grows older interacting with societal and within-family styles.

ACKNOWLEDGMENTS
This work was supported by grants from The Research Council of Norway grant no. 136108/310 (T.S.), Dr. med. Letten F. Sangstads Fund (E.B.J.), The National Council for Mental Health – Norway (H.A.), The University of Cape Town, and The South African Medical Research Council (V.A.R.). We also thank A. Charles Catania for valuable discussions during the early phase of preparing this article.

Precommentary

Jeffrey Gray was the BBS Editor for this treatment. When he accepted Sagvolden et al. for publication, he invited Catania, who was one of the reviewers of the original submission, to prepare a precommentary. Commentators were then invited to respond to the Sagvolden et al. article, to the Catania precommentary, or to both.

Attention-deficit/hyperactivity disorder (ADHD): Delay-of-reinforcement gradients and other behavioral mechanisms
A. Charles Catania
Department of Psychology, University of Maryland, Baltimore County (UMBC), MD 21250. catania@umbc.edu
http://www.umbc.edu/psych/personal/catania/cataniash.html

Abstract: Sagvolden, Johansen, Aase, and Russell (Sagvolden et al.) examine attention-deficit/hyperactivity disorder (ADHD) at levels of analysis ranging from neurotransmitters to behavior. At the behavioral level they attribute aspects of ADHD to anomalies of delay-of-reinforcement gradients. With a normal gradient, responses followed after a long delay by a reinforcer may share in the effects of that reinforcer; with a diminished or steepened gradient they may fail to do so. Steepened gradients differentially select rapidly emitted responses (hyperactivity), and they limit the effectiveness with which extended stimuli become conditioned reinforcers, so that observing behavior is less well maintained (attention deficit). Impulsiveness also follows from steepened gradients, which increase the effectiveness of smaller, more immediate consequences relative to larger, more delayed ones. Individuals who vary in the degree to which their delay gradients are steepened will show different balances between hyperactivity and attention deficit. Given the range of ADHD phenomena addressed, it may be unnecessary to appeal to additional behavioral processes such as extinction deficit. Extinction deficit is more likely a derivative of attention deficit, in that failure to attend to stimuli differentially correlated with extinction should slow its progress. The account suggests how relatively small differences in delay gradients early in development might engender behavioral interactions leading to very large differences later on. The steepened gradients presumably originate in properties of neurotransmitter function, but behavioral interventions that use consistently short delays of reinforcement to build higher-order behavioral units as a scaffolding to support complex cognitive and social skills may nonetheless be feasible.

Keywords: ADHD, attention deficit, delay gradient, exponential decay, extinction deficit, hyperactivity, impulsiveness, intervention, observing responses, self-control

Sagvolden, Johansen, Aase, and Russell (Sagvolden et al.) provide an interpretation of attention-deficit/hyperactivity disorder (ADHD) at levels of analysis that range from neurotransmitters to behavior. In the long run, the success of their account will depend on the adequacy with which fine details of dopamine systems are linked via grosser cellular and neuroanatomical levels to their eventual molar behavioral products. To the extent that evolutionary contingencies have selected nervous systems on the basis of the behavior that they engender, we must understand the properties of that behavior if we are to understand how the brain serves it (Catania 2000). My main objective here is to elucidate aspects of Sagvolden et al.’s account that bear on the possible roles of delay-of-reinforcement gradients and other behavioral phenomena in producing ADHD.

The ubiquity of delayed reinforcement. Much important behavior, called operant behavior, occurs because of its consequences, that is, its effects on the environment. Some important consequences are those that afford opportunities for new behavior, as when something one does allows eating or drinking or playing, or as when one’s shift of attention leads to new things seen or felt or heard. Responses that produce particular consequences are said to be members of operant classes. Some consequential effects are im-
meditate, and others are delayed, and their immediacy determines the potency with which they change or maintain behavior. In other words, the extent to which consequences such as reinforcers operate to alter the future likelihood of responses in the class that produced them depends, along with many other variables, on the delays between the responses and their consequences.

Delay of reinforcement is a ubiquitous effect even if reinforcers are delivered very promptly upon responses, because other responses typically precede the one that actually produces the reinforcer (Dews 1962). “The reinforced response is followed by the reinforcing stimuli; the preceding unreinforced responses are also followed by the reinforcing stimuli, though not quite so promptly. Indeed, the whole pattern of . . . responding is followed by the reinforcing stimuli and so, in a sense, is reinforced” (Dews 1966, p. 578). It was once regarded as paradoxical that schedules of intermittent reinforcement produced more behavior than the reinforcement of every response. But if only every tenth response produces a reinforcer, ten responses, not just the last one, share in the effects of that reinforcer. The earlier responses make a smaller contribution than the later ones by virtue of the longer delays that separate them from the reinforcer, but the sum of all ten contributions is necessarily greater than that from the tenth response alone.

One way of thinking about how reinforcers work is to assume that responses weighted according to a decay function by the delays that separate them from a reinforcer contribute to a reserve of potential behavior, and that subsequent responding depends on the magnitude of that reserve, which is then depleted when responding occurs without reinforcement (e.g., Catania 2001; 2005). Skinner (1938) proposed a reserve that received contributions only from the response that just preceded the reinforcer, but retracted the proposal when it became clear that it could not accommodate data from schedules of reinforcement (Skinner 1940). The retraction might have been unnecessary if the contributions of responses preceding the one that produced the reinforcer had been recognized (Catania 1971).

Furthermore, delays may affect behavior in other ways. The onset of a stimulus that sets the occasion for responding may be followed by a reinforced response after a shorter or a longer delay. If reinforcers are delivered in its presence, the stimulus will become a conditional reinforcer, but its potency will depend on the delay (Dinsmoor 1983; 1995). One simple but exceedingly important response that is maintained by such a stimulus is that of attending to it. A stimulus in the presence of which an opportunity for reinforcement is likely to arise very soon is more likely to be observed or looked at or attended to than one in the presence of which that opportunity is still some time away.

**Experimental assessments of delay gradients.** Figure 1 provides examples of two delay gradients obtained with pigeons. The first shows rates of responding as a function of the time between one response and the later reinforcement of a different response; the second shows rates of responding maintained by a response-produced stimulus as a function of the time between the onset of that stimulus and the subsequent delivery of a reinforcer in its presence. In both cases the data have been fit by exponential decay functions. Candidates for the delay gradient have included exponential, hyperbolic, and logarithmic functions, but the appropriateness of one or the other depends on both procedural and statistical considerations. For example, integrals of hyperbolic functions approach logarithmic functions, so the former are better fits to data from procedures that assess one point on the gradient at a time, whereas the latter are better fits to data from procedures that assess rates of responding over long time periods and therefore across a range of delays. Furthermore, variance in the decay parameters of exponential functions may generate hyperbolic functions when data are averaged (Killeen 1994; 2001).

The first experiment illustrated in Figure 1 involved random-interval reinforcement of a sequence of pecks on two keys by a pigeon. For example, if reinforcement was contingent on exactly four left pecks followed by exactly four right pecks, left pecks would always be separated from the reinforcer by the time taken to emit the right pecks, and that time could be manipulated by varying the required number of right pecks. The data for Pigeon 73 in Figure 1 were obtained by varying the required number of pecks on the right key (R), while the number required on the left key (L) was held constant (cf. Catania 1971). Similar data can be generated with procedures that alter the time it takes for the pigeon to emit its right-key pecks; such procedures demonstrate that time rather than the intervening number of responses is the appropriate dimension along which to measure the effects of delayed reinforcers (cf. Catania 1991).

---

![Figure 1](Catania). Pigeon 73: Rate of left-key pecks as a function of the delay between the last left-key peck (*) and a reinforcer produced by a right-key peck (†). Pigeon 47: Rate of key-A pecks as a function of the delay between the key-A peck that turned on the key-B stimulus (*) and the later production of a reinforcer by a key-B peck in the presence of that stimulus (†). Procedures are shown schematically below each graph.
The second experiment involved an observing-response procedure (Kelleher et al. 1962). During successive presentations of yellow on the right key (B), contingencies irregularly alternated between a fixed-interval schedule of reinforcement and an equal duration of extinction. These presentations were preceded by brief presentations of the left or observing-response key (A), lit white. If a white-key (observing) peck occurred during a brief window of time before the onset of the right-key stimulus, the right key lit green if the current contingency was fixed-interval reinforcement, and the right key lit red if it was extinction. Procedures that allow observing pecks to produce only green if fixed interval, or only red if extinction, show that observing pecks are maintained because green under these circumstances functions as a conditional reinforcer. Essentially, pigeons peck the observing key in order to get a look at green on the right key. But, as shown in Figure 1, the rate of left-key pecking decreases as a function of the duration of the fixed interval. The potency of green as a conditional reinforcer that maintains the observing response depends on the delay from the onset of green to the later delivery of a reinforcer. A substantial body of evidence demonstrates that organisms work to observe discriminative stimuli correlated with the delivery of reinforcers; they do not work to observe discriminative stimuli that are equally informative but are instead correlated with extinction or aversive events (Dinsmoor 1983; 1985).

Both delay gradients in Figure 1 extend over many seconds. They are the facts about behavior that must be taken into account by hypotheses about mechanism. The gradients may be expected to vary as a function of a variety of parameters, and their properties are presumably influenced by such factors as whether response sequences are homogeneous or heterogeneous, and whether the responses that make up those sequences are relatively simple units or are instead integrated higher-order, and perhaps temporally extended, ones (Catania 1995; 1998). In any case, the durations of the delays considered here differ by orders of magnitude from those of synaptic events or even of cascading neuronal processes involving large numbers of cells.

**Implications of anomalous delay gradients.** Now we are ready to examine the implications for ADHD. As argued by Sagvolden et al., the two major components of the ADHD syndrome, hyperactivity and attention deficit, can each be interpreted as consequences of a delay-of-reinforcement gradient that is more limited in its temporal range than the ordinary delay gradient. Figure 2 illustrates the rationale by comparing one hypothetical exponential decay gradient with another that declines more steeply over time. Each gradient is assumed to end when it reaches the previous reinforcer, based on data showing that the retroactive effects of reinforcers do not extend back past the previous reinforcer to still earlier responses (Catania et al. 1988), though this blocking might be attenuated in situations where reinforcers vary in kind or magnitude.

If gradient 1 operates for the reinforced behavior of a given organism at a given time, then the five responses in A as well as the five in B will share in the effects of the reinforcer, though the summed effects in B will clearly be greater than those in A. Similarly, it will support the stimuli in both C and D as conditional reinforcers, but the effectiveness as a conditional reinforcer of the stimulus in C will clearly be weaker than that in D. With gradient 2, however, the early responses in A and the stimulus with early onset in C will be outside the range of effectiveness of the reinforcer, because at those longer delays the gradient is at near-zero levels. This gradient will differentially strengthen relatively rapid sequences of responses, and only stimuli with relatively short delays from onset to reinforcer will be sufficiently effective as conditional reinforcers to sustain observing behavior. The outcome will be rapid responding accompanied by deficits in observing behavior or, in other words, hyperactivity plus attention deficit. The differential strengthening of relatively rapid responding takes time, so a delay function like that of gradient 2 may engender hyperactivity, but the hyperactivity may take a while to develop and may develop separately in different environments.

The case for steepened delay gradients as a mechanism underlying ADHD is strengthened by comparisons of the behavior of Wistar Kyoto (WKY) and spontaneously hyperactive (SHR) rats (though the latter abbreviation was originally based on the hypertension of those rats, which was discovered first, rather than on their hyperactivity). Sagvolden et al. present the argument for SHR rats as a nonhuman model for ADHD in some detail (and see also Sagvolden 2000; Sagvolden et al. 1993; 1988). In other research with WKY and SHR rats, reinforcers were arranged for a fixed consecutive number of responses on one lever followed by a single response on a second lever, and longer response sequences were maintained by WKY rats than by SHR rats (Evenden & Meyerson 1998). This is what we would expect if delay gradients for SHR rats were abridged or steepened relative to those of WKY rats, and it suggests that a direct comparison of delay gradients for SHR and WKY rats in experiments similar to those illustrated in Figure 1 would be of substantial interest. And if a quick way could be developed to obtain such gradients from non-ADHD and ADHD children (say, using computer games on laptop computers), such data would not only help to validate Sagvolden et al.’s SHR model but might also be of considerable diagnostic value.

To this point I have considered only gradients based on reinforcing events. It would be useful to know about the properties of delay gradients involving aversive stimuli. Aversive stimuli may reduce behavior when they are contingent upon responses in punishment procedures, or they may maintain behavior when they are postponed or cancelled by responses in avoidance procedures (Catania 1998, pp. 88–110). Steepened gradients would probably make a difference in either case. Steepened punishment gradients would reduce the effectiveness of both natural punishment contingencies (e.g., getting burned upon touching a hot stove) and artificial ones (getting scolded after teasing a sibling); this could be manifested in promiscuity to accidents as well as in disobedience. Steepened avoidance gradients would make it more difficult to maintain avoidance behavior, because such behavior makes only indirect contact with aversive events (after a successful avoidance response, nothing happens); this could be manifested in risk-taking or other varieties of carelessness.
Impulsivity. One aspect of behavior often included in diagnoses of ADHD is impulsivity or impulsiveness, where behavior with fairly immediate consequences dominates over behavior with larger but more delayed consequences. Impulsivity is sometimes described in terms of executive dysfunction, or disinhibition, or failure to withhold behavior, and it is typically regarded as the inverse of self-control (Rachlin & Green 1972). An account of impulsivity and self-control in terms of hypothetical delay gradients is illustrated in Figure 3 (cf. Rachlin 1995, p. 111, Fig. 1).

Imagine a rat given access to two levers on trials that occur every minute or so. A press on the first lever 10 seconds into the trial or later produces a small reinforcer, and a press on the second lever 30 seconds into the trial or later produces a large reinforcer. Each trial ends as soon as either reinforcer is delivered. If 10 seconds pass and the rat presses the first lever, it receives the small reinforcer but has permanently lost the large one on that trial. The only way to obtain the later large reinforcer is to refrain from pressing the first lever until the large reinforcer is available for a press on the other lever. On the left, Figure 3 shows the respective exponential decay gradients engendered by the smaller but earlier reinforcer arranged for the first response at time A, and by the larger but later reinforcer arranged for the other response at time B.

This example assumes some separate experience with the contingencies arranged for each lever. A rat in this situation for the first time might start with presses on the A lever, always producing the smaller, more immediate reinforcer, and so might never reach the time at which its press on the B lever could produce the larger but later one. The relative heights of the respective gradients can be taken as representing the relative likelihoods of the two responses during the time leading up to the earlier reinforcer. The two gradients are shown starting at different maxima reflecting the different A and B reinforcer magnitudes; if they started at equal maxima and decayed at equal rates, they could not cross at E.

In this example, the B response is more probable than the A response up until time E, but thereafter the A response becomes more probable. One way to overcome the higher probability of A (or, in other words, to show self-control rather than impulsiveness) is if a B response prior to time E becomes a commitment of some kind. For example, the B response might make the A response unavailable (perhaps via retraction of the A lever) for the remainder of the time until the B reinforcer becomes available. Under such circumstances, we might observe many instances of self-control, in the sense that B responses committing to the larger reinforcer would occur before any A responses that would produce the smaller earlier reinforcer and therefore end the sequence.

Now consider the steeper gradients on the right in Figure 3. In this instance, the gradient engendered by the smaller earlier reinforcer is everywhere higher than the other gradient in the time leading up to C, even though the D gradient starts at a relatively higher maximum. With these steepened gradients, there will be no circumstances in which the probability of the D response exceeds that of the C response, so self-control will be completely displaced by impulsivity. Impulsivity follows so directly from these kinds of gradients that it is not necessary to appeal to deficient extinction or executive dysfunction.

For impulsivity, as for hyperactivity and attention deficit, no problems are posed by issues of localization, such as Sagvolden et al.’s discussion of dopaminergic systems in mesolimbic, mesocortical, and nigrostriatal branches (e.g., the target article’s Fig. 1). Delay gradients with common decay properties could as easily operate for behavior classes intermixed within a single area as for those discretely localized in separate areas.

Individual differences in the balance between hyperactivity and attention deficit. As outlined in Sagvolden et al.’s review of ADHD, some individuals display both hyperactivity and attention deficit, but in others one or the other component dominates. These individual differences vary with gender, age, and other variables (e.g., Sagvolden & Berger 1996). They can be accommodated by assuming delay gradients that decline at different rates. Varieties of presentation of ADHD symptoms are perhaps best viewed not as separate classes but rather as lying along a continuum involving rate of decay of the delay gradient as a parameter. Two ways in which delay gradients might vary are illustrated in Figure 4.

Consider first the family of gradients on the left, in which the highest gradient (a) represents a normal or non-ADHD gradient. Let us start with the steepest gradient, furthest from the normal gradient. For the individual whose gradient drops asymptotically...
to near zero within a second or so, responses must be very close to the reinforcer to be captured by it. The time period is so short that only single responses can typically be strengthened. If sequences of responses cannot be strengthened, there will be no hyperactivity. But this gradient will generate profound attention deficit, because only brief stimuli quickly followed by reinforcers will acquire any conditional reinforcing effectiveness. (We might also expect such other problems as severe impulsiveness and poor acquisition of coordinated sequential behavior.)

Next consider a gradient that drops asymptotically to near zero only after a delay of a couple of seconds or so. Attention deficit is still likely to be a problem, but in this case sequences of rapid responses will sometimes be fully captured within the effective temporal extent of the gradient. They will come to dominate over slower sequences of responses, so in this instance we can expect to see both attention deficit and hyperactivity.

Finally, consider a gradient that drops asymptotically to near zero only after several seconds and therefore is closer to the normal gradient (a). The longer time period means that attention deficit will be less of a problem, because stimuli will acquire conditional reinforcing properties, though perhaps with slightly diminished potency. But faster response sequences will still be differentially strengthened relative to more leisurely ones. In this case hyperactivity will dominate and any attention deficit that becomes evident is likely to be mild.

We could play out the details further (e.g., by extending the argument to impulsivity), but the point is that a single parameter determining the rate of decay of the delay gradient might be sufficient to determine both the absolute and the relative severity of the attention and hyperactivity components of ADHD. If a compromised dopamine neurotransmitter mechanism is implicated in ADHD, as proposed by Sagvolden et al., graded behavioral outcomes should be expected from variations in the degree of compromise. The account is of special interest because it promises to subsume a range of individual differences under a single mechanism.

But this is only one way in which the parameters of delay gradients might vary. Another possibility is illustrated in the right graph of Figure 4. In that case, the normal or non-ADHD gradient (b) is the one that crosses the y-axis at the lowest point. The others decline more steeply, like those in the left graph. Here the area under each curve is equal to a constant. Such functions might be appropriate, for example, if variations in the rate of decay depend on how quickly a fixed quantity of some neurotransmitter is depleted. Such depletion can occur either slowly or rapidly, as in the family of curves on the left, but the steeper the rate of decay, the higher the maximum would have to be to hold the area constant. Differential selection of response sequences and maintenance of attention would still vary with the rate-of-decay parameter, but these curves have some additional implications.

One argument in favor of the equal-area functions on the right over the exclusively decremental functions on the left is suggested by the impulsivity examples in Figure 3. An account of impulsivity in terms of exponential gradients will not work unless the gradients generated by different reinforcer magnitudes start at different maxima. Furthermore, if the effects are everywhere decrements, as on the left in Figure 4, then the only source of higher rates of responding would be the differential selection of rapid sequences; with extreme decrements, little if any responding could be supported by reinforcers. This might be an appropriate model for other behavior pathologies, but it seems not to capture the defining features of ADHD.

The equal-area functions in Figure 4, however, are consistent with a model in which a reserve of potential behavior is replenished by responses weighted according to the delays that separate them from a reinforcer and in which subsequent responding depends on the magnitude of that reserve. In this case, hyperactivity follows not only from the differential strengthening of more rapid sequences but also from the direct strengthening of responses that are very quickly followed by reinforcers. With equal-area functions, greater strengthening occurs with steeper functions, but with steeper and steeper functions, the temporal window within which responding will be strengthened progressively narrows.

Sagvolden et al. argue that children with ADHD are less sensitive to changes in reinforcement contingencies and require stronger and more salient reinforcers. This might seem consistent with the decremental (left) gradients of Figure 4, but problems that appear to be motivational might instead be problems of contingencies. Apparent insensitivity to reinforcement contingencies can come about not only because of weak reinforcers but also because of strong reinforcers presented after a delay. Furthermore, the latter problem will be more likely with steeper delay gradients.

**Extinction deficit.** I have so far emphasized delay gradients. But along with their presentation in terms of delay gradients, Sagvolden et al. have also offered extinction deficit as an alternative mechanism contributing to the complex of symptoms that define ADHD. We have already seen that delay gradients on their own adequately account for many features of ADHD, but there are other reasons besides parsimony to question the role of extinction deficits.

Extinction demonstrates that the effects of reinforcement are temporary, and Sagvolden et al. correctly point out that the variables that produce increments in responding when reinforcement.
begins may be different from those that produce decrements af-ter it ends. It is therefore appropriate to consider different mech-anisms for reinforcement and for extinction. But extinction deficit, the absence of the response decrements that typically occur dur-ing extinction, has no relevant temporal parameters and therefore is not applicable to situations that can be interpreted in terms of differential delays (that is another reason why the direct determina-tion of delay gradients with WKY and SHR rats might be especially valuable).

One problem with assessing extinction effects is the metric used to assess the progress of extinction. For example, if extinction for SHR rats begins with higher baseline rates of responding than for WKY rats, should comparisons be based on relative declines in responding or on the absolute levels reached at certain times? Procedures that changed baseline rates of responding for one or the other group in an attempt to match baseline rates would have to deal somehow with the differential effects of the contingencies that such matched baselines would require.

Another and perhaps even less tractable problem with assessing extinction deficit, however, is that extinction is rarely studied in isolation. In Jokhansen and Sagvolden (2004), for example, extinc-tion was studied in successive sessions that each began with a fixed period of reinforcement. Thus, the procedure involved the acquisition of a discrimination between the early and the late por-tions of each session. If attention deficit affects orientation toward visual cues, it presumably also affects attention without evident motor components, such as attention to temporal cues. (I here treat attention as a variety of behavior, but one defined by the en-vironmental contingencies it can enter into rather than by a par-ticular topography.) Thus, even if SHR rats responded more in extinc-tion than WKY rats, the difference could be attributed as readily to differences in attention to temporal stimuli as to an extinc-tion deficit.

Failure to attend to temporal cues rather than extinction deficit might also account for continued responding early in the individ-ual segments of fixed-interval (FI) schedules of reinforcement. A similar confusing exists in procedures that compare reinforce-ment versus extinction contingencies arranged in the presence of different visual or auditory stimuli, where what might seem like extinction deficit might depend instead on a failure to attend to relevant stimuli. Thus, it seems reasonable to consider the possi-bility that extinction deficit is not a separate source of some of the properties of ADHD, but rather is a derivative of the kinds of anomalies of delay gradients that we have already considered.

I have had little to say here about other factors that might con-tribute to ADHD, such as executive functions, verbal governance, and other higher-order processes. But given differences in delay gradients similar to those already considered, it is plausible that complex skills such as the hierarchical structuring of verbal gov-erned behavior and the monitoring of one’s own behavior would develop differently in a child with, than in a child without, ADHD.

**ADHD and development.** As we know from the analysis of non-linear systems, very small differences in initial conditions can re-sult in exceedingly large long-term differences (Gleick 1987). For example, even if the only problem with autism was aversion with regard to both eye contact and touch, many of the everyday con-tingencies that build social interaction would be missed, such as not noticing when a parent smiles at something one has done. These interactions provide the scaffolding on which more com-plex social behavior depends, including verbal behavior, so the effects will be seen in all of the other behavior that depends on them. This is presumably why early intervention matters so much.

One significant feature of Sagvolden et al.’s account is the par-allel case they have presented for ADHD. It should be no surprise that different early histories with ADHD, especially in combina-tion with the variations in delay gradients that we have enter-tained, could lead to vastly different spectra of behavioral compen-ses and difficulties. Might small path dependencies lead sometimes to oppositional defiant disorder and sometimes to con-duct disorder and sometimes to neither? Even the dominance of motor versus cognitive components might depend on differences in historical paths, and perhaps we should also entertain the pos-sibility that such behavioral trajectories can drive certain features of brain organization rather than be driven by them. As suggested by Sagvolden et al., analyses in terms of the ebb and flow of com-plex interactions of behavior with contingencies involving parents, peers, teachers, and others are a daunting but unavoidable chal-lenge.

Perhaps there are also circumstances in which features of ADHD are advantageous. With experimental contingencies that favor varied over stereotyped response sequences, for example, comparisons of the behavior of WKY and SHR rats have shown that SHR rats learn to vary rather than repeat sequences more readily than WKY rats (Mook et al. 1993). Variable behavior pro-vides the raw material upon which the selection of behavior by contingencies operates within individual lifetimes, so this behav-ioral capacity may have been selected by evolutionary contingen-cies (cf. Neuringer 2002). We may argue from our anthropocen-tric view that an organism with more extended delay gradients will be more capable of taking into account events that are more remote in time, but such capabilities surely must be balanced against the importance of its sensitivity to the immediate consequences produced by its behavior.

**Interventions and implications.** If delay gradients are impli-cated in ADHD, their properties presumably originate in the properties of neurotransmitter function, but this does not imply that pharmacological interventions are the only recourse. Behav-ioral interventions that use consistently short delays of reinforce-ment to build higher-order behavioral units as a scaffolding to sup-port complex cognitive and social skills may nonetheless be feasible. For example, the shaping of behavior with prompt con-sequences both correlated with and intermixed with longer-term ones might provide the prerequisites for building conditional rein-forcers that maintain longer periods of attention and that bridge increasingly extended delays. The decremental (and detrimental) effects of delays might be attenuated with the creation of higher-order temporal units, especially if they also involve mediation by verbal behavior. Computer games may be particularly useful tools, because their rapid responsiveness, which sometimes so easily cap-tures the behavior of children with ADHD, allows both for the precise control of contingencies relating skilled behavior to its consequences and for the structured embedding of minimal behav-ioral units into higher-order coordinated units. Behavior is the interaction of an organism with its environment, so such inter-ventions might teach us things not only about how brain structure drives behavior but also about how behavior drives brain struc-ture.

It may be worth noting that this account has mostly dealt with behavior in its own terms. Although the interpretation of ADHD in terms of delay gradients is theoretical, delay gradients them-selves are not theory but rather are measurable properties of behav-iour. At least in part because of the limitations of my expertise, this commentary has only occasionally made contact with other levels of analysis. One of the great strengths of Sagvolden et al.’s contribution is its articulation among the several levels, and I look forward to the buttressing and the widening of the bridges that they have begun to build among those levels. The following quo-tation is particularly apt:

> Valid facts about behavior are not invalidated by discoveries concern-ing the nervous system, nor are facts about the nervous system invali-dated by facts about behavior. Both sets of facts are part of the same en-terprise, and I have always looked forward to the time when neurology would fill in the temporal and spatial gaps which are inevitable in a behav-ioral analysis. (Skinner 1984b, p. 543)

**ACKNOWLEDGMENTS**

Eliot Shinnoff collaborated in the research that generated the data used il lustriate Figure 1. Rouhen Rostamian provided helpful insights into the properties of exponential decay functions.
Unitary or multiple pathways: The trap of radical behaviorism

Tobias Banaschewski, Sunke Himpel, and Aribert Rothenberger

Child and Adolescent Psychiatry, University of Göttingen, D-37075 Göttingen, Germany. tbanasc@gwdg.de shimpel@gwdg.de arrothen@gwdg.de http://www.gwdg.de/~ukyk

Abstract: Early and automatic neuropsychological processes may be influenced by altered dopaminergic functions but cannot be fully explained by these or by altered reinforcement and extinction processes. The reinforcement-extinction-model is excellent for understanding certain causal pathways of attention-deficit/hyperactivity disorder (ADHD), but it can hardly explain the heterogeneous developmental trajectories of ADHD fully. It should be integrated into a multiple pathways model.

Sagvolden, Johansen, Aase, and Russell (Sagvolden et al.) have conducted an outstanding review of recent behavioral and neurobiological results concerning attention-deficit/hyperactivity disorder (ADHD) and introduced a novel theoretical approach of ADHD. Their dynamic developmental theory proposes that altered reinforcement and extinction processes, mainly associated with a hypofunctioning mesolimbic dopaminergic system, affect learning processes and thereby produce ADHD symptomatology. Catania tries to explain ADHD more parsimoniously without recourse to an extinction deficit. We approach the issue of whether the proposed model(s) can explain symptomatology and correlates of ADHD (e.g., attentional dysfunctions and basic information processing alterations) and fully and completely describe the causal chains occurring in development, because a causal model for ADHD should meet these requirements (Coghill et al. 2005).

Sagvolden et al. ground their approach on alterations of the dopaminergic system, mainly the mesolimbic branch’s involvement in reinforcement and extinction. However, monoamines also have trophic functions in the developing brain influencing the development of other neurotransmitter systems and brain structures (Rice & Barone 2000): Retinal dopaminergic receptors are involved in colour perception; cerebellar dopaminergic neurons, presumably in time perception; and the tubero-infundibular system, in stress regulation. Thus, a variety of hypofunctioning dopaminergic subsystems may exist, leading to heterogeneous functional consequences. By renaming ADHD as reinforcement/extinction disorder (RED), Sagvolden et al. would reduce – in agreement with Catania – the various dopaminergic dysfunctions to a hypofunctioning mesolimbic system involved in reinforcement and extinction processes. By ignoring other possible changes (beside the hypofunctional dopamine system branches), Sagvolden et al. hope to be more concrete in the theoretical issues involved, but the converse may follow: Central concepts like attention which a causal model of ADHD would describe as heterogeneous (both cross-sectionally and during development), may remain undetermined.

Models postulating single unitary underlying mechanisms, such as the deficient inhibitory control approach (Barkley 1997b) or Sagvolden et al.’s model, implicitly assume that all children with ADHD share the same causal etiology. Sagvolden et al. and Catania seek to explain how differences between delay gradients in children with ADHD may cause differences in their ADHD symptomatology. However, the high degree of heterogeneity may indicate the existence of pathophysiological heterogeneity, that is, multiple and – to some extent independent – developmental pathways from etiological factors to brain dysfunctions and behavioural symptoms (Biederman et al. 2005). It seems that ADHD is heterogeneous not only clinically (Biederman et al. 1992), but also genetically (Willcutt et al. 2000) and neurophysiologically (Banaschewski et al. 2003a; 2003b). Sonuga-Barke (2002) has emphasised a motivational pathway (delay aversion) and a pathway of disinhibitory change, which seem to be independently associated with the diagnosis of ADHD (Solanto et al. 2001a). In addition, some studies suggest that the comorbidity of ADHD with ODD/CD may constitute a biologically distinct subtype (Banaschewski et al. 2003a; Faraone et al. 1998). Furthermore, any simple lesion model of ADHD is unlikely to fully explain the disorder because disturbances of the frontocortico-striato-thalamo-cortical circuits can also arise from dysfunctions of posterior cortical regions, the cerebellum or ascending arousal systems, which have also been implicated in ADHD (Castellanos et al. 2002; Sowell et al. 2003).

Further, fundamental early and automatic neuropsychological processes such as attentional orienting, processing speed, time processing, or motor response organization (Brandeis et al. 2002; Sergeant et al. 1999; Smith et al. 2002) may explain – at least partly – ADHD symptoms without being related to reinforcement and extinction. The same holds true for transcranial magnetic stimulation (TMS) studies, which suggest basic inhibitory deficits of intracortical interneurons involved in motor control (Moll et al. 2000a). The latter may explain hyperkinetic behavior like fidgeting largely consisting of undirected, non-responding movements. Convergent evidence of impaired processes of sensory-motor integration and motor control has been found (Banaschewski et al. 2003b; Rubia et al. 2003; Yordanova et al. 2001).

Also, attention problems may be influenced by altered dopaminergic functions but cannot be fully explained by alterations of the dopaminergic system. Animal studies (Bymaster et al. 2002; de Villiers et al. 1995), studies of event-related potentials (ERP) (Banaschewski et al. 2003a; Brandeis et al. 2002), and pharmacological data (for a review, see Banaschewski et al. 2004) suggest the involvement of dysregulated noradrenergic networks, which are associated with automatic attentional processing (Conoll 1998; Posner & Petersen 1990). It seems to be difficult to explain these various findings fully as learned behaviour, altered by mesolimbic dopaminergic dysfunctions.

Finally, developmental effects and compensatory processes need to be taken into consideration by any causal model of ADHD. Hence, the developmental description of the nonlinear, discontinuous, and asynchronous time courses of the development of symptom domains, brain structures, psychophysiological parameters, and neurotransmitter systems across the lifespan, particularly in adolescence and adulthood, is necessary (Barry et al. 2003; Biederman et al. 2000; Herschkowitz et al. 1997; Moll et al. 2000b). Therefore, modifications of both maturational lag and developmental deviation models of ADHD are needed to explain results from behavioural and electrophysiological studies adequately (Barry et al. 2003; Rothenberger et al. 1987; Woerner et al. 1987). Focusing on the underlying dysfunction rather than the behavioural profile, Clarke et al. (2002) proposed distinct heterogeneous developmental pathways within the ADHD population which are largely independent of the DSM-IV-TR (American Psychiatric Association 2000). Diagnostic categories. Sagvolden et al.’s model can hardly explain the developmental trajectories of ADHD symptomatology and correlates in concreteness. Of course – as Catania pronounces – relatively small differences early in development might engender behavioural interactions leading to very large differences later on and differences in historical paths are important, but according to a Sagvolden et al.’s model, symptoms should develop more or less continuously, and be related to correspondent changes in environmental stimuli and reinforcer contingencies, respectively.

In conclusion, the models of ADHD by Sagvolden et al. and Catania are excellent models to understand certain causal pathways of ADHD, but more studies are needed contrasting their reinforcement model(s) with other theoretical models within the same samples. The aim of these studies will be to examine the relationships between the various deficits, evaluate relative effect
sizes, and examine their relationships with genetic and environmental causal factors and associated neural-mediating mechanisms (Coghill et al. 2005). Presumably, the reinforcement-extinction model has to be set in line with other mono- or dual-pathway models of ADHD and integrated into a multimodal causal model (Coghill et al., in press) to avoid the risk of taking a partial account for a full causal model and thus being trapped into the kind of reductionism that is offered by radical behaviourism. Specific pathophysiological pathways for ADHD have yet to be identified. Many components of these pathways may well be shared with other conditions, and others may be unique to ADHD. Longitudinal studies of developmental trajectories are required to determine whether these neurocognitive correlates represent a primary abnormality or a secondary compensation mechanism (Banaschewski et al. 2005).

The role of context and inhibition in ADHD

Petra Björne and Christian Balkenius
Department of Cognitive Science, Lund University, Kungshuset, Lundagård, SE-222 22 Lund, Sweden. petra.bjorne@lucs.lu.se
cristian.balkenius@lucs.lu.se
http://www.lucs.lu.se/People/Petra.Bjorne/

Abstract: We have shown in a computational model that a poor memory for context could result in some of the behaviors associated with ADHD, which is well in line with the dynamic developmental theory. Given the important role of context in extinction, a weaker context due to a steeper delay-of-reinforcement gradient would result in impaired inhibition.

Sagvolden et al., as well as Catania, propose that the main cause of the difficulties seen in ADHD stems from a dopamine dysfunction rendering steeper and shorter delay-of-reinforcement gradients. Only reinforcement delivered close to a response will, according to this view, be associated with that response. While Sagvolden et al. also include a deficient extinction in ADHD, Catania argues that a shortened delay of reinforcement suffices to explain the altered behavior in ADHD.

An alternative explanation is that the difficulties seen in children with ADHD are the result of a weakened ability to code and maintain a context (Balkenius & Björne 2001). Following Fuster (1997), we assume that the lateral prefrontal system is involved in inhibiting reactions to stimuli irrelevant to a given task set or context and we propose that a weaker context will reduce this modulation in ADHD (Balkenius & Björne 2001). Such behavioral inhibition is probably unrelated to inhibition on a synaptic level. Context or working memory are supposedly harder to activate and would be more instable, and thus prone to be overridden by distracting and irrelevant stimuli. Thus, a context not maintained by external cues will not stay in memory and guide behavior and attention.

A computational model based on this assumption can reproduce the behavioral data from several experiments (Balkenius & Björne 2001), testing such features of ADHD as deficient inhibitory control (Börger & van der Meere 2000; Cepeda et al. 2000; Schachar et al. 2000), effort allocation during sustained attention (Börger & van der Meere 2000), task switching (Cepeda et al. 2000; Pineda et al. 1999), and latent inhibition (Lubow &ToShowa 1993).

Although our model makes no assumptions regarding dopaminergic dysfunctions, a shortened delay-of-reinforcement gradient could possibly play a part in explaining the weaker context proposed in our model. Due to the steeper gradients, learning of reinforcement contingencies would be inefficient and slower than normally seen. This is, however, not sufficient to explain the wide variety of experimental data obtained with children with ADHD, which is why we propose that the shortened delay-of-reinforcement gradient needs to be extended with the effects of context on learning and extinction.

It is essential to remember that extinction is not the unlearning of previously learnt associations between a response and a reinforcer or between a stimulus and a response. Rather, extinction is the learning of a new association that masks previously reinforced behavior (Bouton 1994; Hall 2002; Westbrook et al. 2002). This mask is relatively specific to the context where extinction has taken place. It is probable that the effects of the new associative links and the conditions that allow them to form differ from those rules pertaining to excitatory learning. Extinction in Pavlovian conditioning provides additional learning about the relation between a conditioned stimulus and a context, and this has its counterparts in instrumental conditioning (Bouton 1994; Westbrook et al. 2002). The information during learning might be coded differently, with a CS-US (Conditioned stimulus-Unconditioned stimulus) memory activated independently of a context, whereas a CS-no-US memory will be activated depending on a context (Bouton 1993). Thus, context is important in learning to inhibit prepotent responses.

The utility of contextual inhibition is further supported by our computational studies, which show that reinforcement learning, including this mechanism, can easily learn a wide range of tasks, such as the appropriate control of attention, task switching, the Wisconsin Card Sorting Test, and context-dependent categorization (Balkenius 2000; Balkenius & Winburr 2004).

That the excitatory associations are preserved through extinction procedures has been extensively studied in a series of experiments by Rescorla (1996; 1997). He concludes that extinction may not result in an inhibition of the link between a response and its outcome, but between some stimulus and a particular response, independent of the outcome.

An extinguished response may return on account of several factors. In spontaneous recovery, the response reappears after a time interval. It may also reappear due to a change in context. When an association is learnt in one context and extinguished in another, the response will reappear in the original or another context. Further, a response may reappear following a reminder, such that in Pavlovian conditioning a presentation of the US will reinstate the CR (conditioned response) when testing with the CS.

The exact nature of extinction is a question still to some extent unresolved. This is partly due to the fact that the exact nature of the associations formed during learning remains unknown, and hence, the inhibition of these associations might differ between different learning paradigms. However, the importance of the role of context in inhibition is fairly well established.

This role for context has implications for the analysis of the behavior of children with ADHD. As is pointed out by Sagvolden et al., children with ADHD are not always impulsive. In the light of frequent and potent reinforcers, they are better able to plan and control their behavior. We conclude that this would be due to an enhanced contextual understanding and maintenance, thus enabling the child to inhibit responses to irrelevant stimuli. Under other circumstances, the short and steep delay-of-reinforcement gradient would prevent the child from forming a stable enough context for maintenance of task set, and there is an obvious risk that stimuli that are not part of the context nonetheless come to be included.

When a child does display impulsiveness or a poor adjustment to the requirements of the surroundings, this might be due to a slower activation of the relevant context. In the case where the same task is repeated, the context remains the same, while a change of task requires an activation of a new context. This process seems to be slower in persons with ADHD. As argued by Cepeda et al. (2000) and Pineda et al. (1999), children with ADHD have an impaired ability to change cognitive set, that is, they show a predisposition to respond to a set of stimuli in a certain way, finding it difficult to inhibit old responses because of a contextual change.

Children with ADHD prefer immediate reinforcers to delayed ones even in cases where the immediate reinforcer has a smaller
value than the delayed one (Sonuga-Barke 2002; Sagvolden et al.). This has been termed delay-aversion. An alternative interpretation within a framework of a weak context would be that the cues reminding of the larger future reward are not salient enough for them to enter the context and thereby guiding behavior, inhibiting a response toward an immediately attainable reinforcement. Thus, rather than an aversion towards a delay, the early response may be seen as an inability to postpone a response with the help of a sufficiently strong context. As seen in extinction experiments, context plays a crucial role for inhibiting a response, that is, context provides information on the time and place of a response, not by excitation but by inhibition. Not yet, not now, or not here. This impaired contextually postponed response would be particularly evident if the early response has been previously reinforced and thus in need of inhibition.

We propose that, in addition to frequent and potent reinforcers, pedagogical interventions for children with ADHD need to be complemented with an explicit teaching of context relevant for a task. The context or cognitive set, if properly formed and maintained, will guide the behavior by directing and sustaining attention and inhibiting responses to stimuli irrelevant for the task. Not only will a child with ADHD be in need of frequent reinforcers, but he or she will also most probably benefit from highlighting of the stimuli relevant to the task. Thus, note that the stimuli to be highlighted are also temporal, given that time is also part of the context (Bouton 1993; 1994). Thus, questions such as How long?, In what order?, and What next? should be answered within the pedagogical setup. This will provide enough contextual cues for the child to successfully complete the task.

Furthermore, the response that does indeed produce reinforcement should be emphasized, and the relation(s) between contextually relevant stimuli, response(s) and reinforcement(s) should be made evident. The pedagogical aids used need to be stable, preferably visual, in order to help the child to update the memory of the current context (Peeters & Gillberg 1998).

If indeed the child should acquire a response that is unwanted, care should be taken to teach inhibitory rules across as many situations as possible (home, school, soccer team, etc.), because inhibitory learning does not generalize to all contexts, as argued earlier in the discussion of extinction (Bouton 1994).

In conclusion, though shortened delay-of-reinforcement gradients provide important cues to the impairments seen in ADHD, we argue that the explanation needs to be supplemented with a model of how context is formed and maintained, as well as how it guides goal-directed behavior. There is a need to further investigate possible impairments in extinguishing previously reinforced responses, as this will provide us with data on the nature of contextual inhibition in ADHD.

Frontal and executive dysfunction is a central aspect of ADHD

Ximena Carrasco, Vladimir López, and Francisco Aboitiz

Programa de Morfología, Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, Santiago, Chile. xcbarrasco@med.uchile.cl vlopez@puc.cl faboitz@puc.cl http://www.neuro.cl

Abstract: In the target article, Sagvolden and collaborators propose that attentional-deficit/hyperactivity disorder (ADHD) is the result of a general behavioral deficit which is mainly caused by a hypofunctioning mesolimbic dopaminergic system. Although we partly agree with this view, we think that it tends to overlook the dysfunction of prefrontal and frontostriatal executive functions by considering them to be a consequence of alterations in reward and extinction mechanisms. Rather, we believe that ADHD is the result of an overall cognitive and behavioral condition, associated to a generalized dopaminergic network dysfunction, and may not be easily attributable to a single basic behavioral function. Sagvolden et al. propose a novel approach to explain the etiology of ADHD on the basis of behavioral analysis, backed by specific synaptic mechanisms. In particular, they concentrate on hypofunctioning mesolimbic dopaminergic transmission which leads to altered reinforcement of behavior and deficient extinction. Basically, the authors argue that the time available for associating behavior with its consequences is shorter in ADHD than in normal children because of dopaminergic dysfunction. Effects on attentional mechanisms, on executive function, and in motor control are also considered, but these are seen as the products of interaction with other hypofunctioning dopaminergic systems (mesocortical and nigrostriatal, respectively). In fact, the concept of a failure to inhibit responses as central to ADHD is considered misleading, because hyperactivity is thought to result mainly from a deficient behavioral extinction process.

We agree with the proposal of a central role of dopaminergic transmission in ADHD, subdivided in three main domains (mesocortical, mesolimbic, and mesostriatal), which constitutes a widely accepted conceptual framework. Furthermore, we favor the idea that the ADHD-associated 7R-DRD4 genetic polymorphism appeared some 40,000 years ago as a result of evolutionary pressure, perhaps in an environment that favored migration and risk-taking behavior (Ding et al. 2002). However, the proposal of separating the inattentive type from the hyperactive-impulsive type as two distinct conditions may be in conflict with evidence suggesting a common genetic basis for both conditions (LaHoste et al. 1996).

Our main criticism relates to the consideration that most of the deficits observed in ADHD are a consequence of alterations in reward and extinction behavioral mechanisms. There are attentional features in ADHD that may not be easily explained by lack of reinforcement/extinction but suggest specific cognitive deficits. For example, not all attentional functions are impaired in ADHD. These patients show a consistent deficit in sustained attention and in some selective attention tasks, but they respond faster in all externally presented attentional tasks and make fewer mistakes in divided attention tasks (Koschack et al. 2003). These findings are easier to interpret in a cognitive framework, in which there is a different distribution of attentional resources in ADHD, with a wider spatial (peripheral) attentional framework and with a narrower time constant (decreased sustained attention) (López et al. 2004). Although a decreased sustained attention could be viewed in terms of a narrower time window to control behavior, the increased performance in peripheral attentional tasks is not easily interpreted in terms of an alteration in reinforcement/extinction mechanisms. We have previously suggested that an ADHD-like wide spatial attentional framework may correspond to a more primitive attentional system, whereas the mechanisms involved in sustained attention underwent an important development much more recently in human evolution, in relation to elaborate tool making, reading, and writing, among other human activities (López et al. 2004).

More generally, we consider that the dynamic developmental theory tends to overlook the dysfunction of the mesocortical system, which works in parallel to the mesolimbic system. In our view, there is a generalized deficit of neurotransmission in ADHD, which generates consequences both in the cognitive and in the behavioral control domains. For example, the deficits in the development of working memory in ADHD are perhaps better explained as a specific impairment of fronto-striatal cortical executive functions (Dowson et al. 2004; Durston et al. 2003; Mehta et al. 2004; Schweitzer et al. 2000), rather than as a consequence of impulsivity due to extinction failure. Furthermore, electrophysiological evidence indicates that, beside its positive behavioral effects, stimulant treatment improves frontal function and cognitive performance (López et al. 2004). Other authors have argued that cognitive impairment associated with ADHD may result from a
Delay of reinforcement gradients and attention-deficit/hyperactivity disorder (ADHD): The challenges of moving from causal theories to causal models

David R. Coghill
Division of Pathology and Neuroscience (Psychiatry), Ninewell’s Hospital and Medical School, University of Dundee, Dundee, DD1 4HN, Scotland, United Kingdom. david.coghill@tptc.scot.nhs.uk

Abstract: Notwithstanding the many strengths of the dynamic developmental theory, there remain challenges to be overcome before it can be incorporated into a true causal model of attention-deficit/hyperactivity disorder (ADHD). These include the development of reliable measures of reinforcement delay gradients, the validation of shortened reinforcement delay as an endophenotype, and the integration of this pathway with other potential pathways.

Sagvolden et al. describe the dynamic developmental theory, a well-thought-through causal theory of attention-deficit/hyperactivity disorder (ADHD), which, they propose, can account for the wide range of difficulties faced by those with this disorder. Their theory has many strengths, and its attempt to address several of the important barriers which have inhibited our ability to shift from positing causal theories to demonstrating formal causal effects is welcome. Unlike many theorists, Sagvolden et al. recognise – and indeed make central to their theory – the need to take developmental aspects seriously. Through a clear recognition of the potentially two-way interactions between the proposed pre-existing dopaminergic hypofunctioning and further biological, cognitive, and social developmental processes, clear and logical predictions are made concerning the resultant impact on and contribution towards the expression of the ADHD behavioural phenotype. They have also resisted the temptation, despite the high estimates of heritability of ADHD (around 0.8), to neglect the potential impact of non-genetic causal factors on their theory. This is important because it is often forgotten that pervasive non-genetic risks, such as environmental toxins and cultural factors, may inflate heritability estimates via genotype–environment correlation. There are, however, several other challenges facing this theory before it can make the transition from causal theory to causal model. I briefly discuss these challenges in the remaining part of this commentary.

Although Sagvolden et al.’s detailed description of the potential roles of and relationships between factors across multiple levels of analysis – from genetic and environmental causal factors through neural and cognitive mediating mechanisms to the behavioural manifestations – is a strength, it represents only a necessary first step in developing a full causal model of ADHD. The hard work starts here. The authors present strong evidence to support the assumptions made at each level of analysis, but considerable effort is still required for these findings to be constructed into empirically supported rather than theoretically promising causal chains. Only one small study (Johansen et al. 2002) is cited which measures reinforcement-delay gradients in children. The apparent lack of a reliable and practical way of measuring these gradients in children with and without ADHD is clearly limiting. Until such measures are developed and validated, it will be difficult to further investigate shortened reinforcement gradients as an endophenotype for ADHD. And even when measures are available, the association with ADHD, its heritability and family co-segregation and neural and physiological substrates will need to be defined. Until then, much of the theory and many of the proposed links within the causal chains must remain speculative. Such a situation is not unusual within ADHD research. The complexities inherent within each level of analysis have resulted in most ADHD researchers concentrating on one level of analysis, with few having accepted the challenge of working across the levels. Exceptions are beginning to emerge, utilizing electrophysiological (e.g., Brandeis et al. 2002), functional neuroimaging (e.g., Rubia et al. 1999), neuropsychopharmacological (e.g., Rhodes et al. 2004), and pharmacogenomic (e.g., Roman et al. 2004) approaches, all of which will be important future lines of research for the dynamic developmental theory of ADHD, once a measure of reinforcement gradients is available. Data from animal models of ADHD such as the spontaneous hypertensive rat, although helpful in this quest, are not a substitute for human studies. Although the measurement of delay aversion as an indirect indicator of reinforcement-delay gradients is also positive, formal links between a shortened delay gradient and delay aversion in ADHD remain untested, and there are, as yet, no published studies describing the neural substrates or genetic correlates of the various delay-aversion tasks studied in ADHD.

Causal models of ADHD will also need to account for the heterogeneity which is being increasingly recognised as a key factor in the understanding of ADHD. It is now generally agreed ADHD is likely to be the consequence of multiple causal pathways which may interact with each other in varying degrees (Coghill et al., in press). Sagvolden et al. address heterogeneity and multiple pathways in several ways. Their suggestion that hyperfunctioning in the various branches of the dopamine system will result in the different symptom patterns found in the inattentive and combined subtypes of ADHD is interesting, but it requires further study. They also describe:

1. The ways that reward and extinction deficits (or, in Catania’s opinion, reward deficits alone) may result in the full range of symptoms associated with combined subtype ADHD.

2. The developmental processes by which interactions between such deficits and differentially active environmental factors could result in within sample heterogeneity.

While helpful, this explanation is unlikely to fully account for the heterogeneity found within ADHD samples. A head-to-head comparison of delay aversion and behavioural inhibition in children with combined subtype ADHD suggested that although both were present within the sample, the two deficits were uncorrelated with each other (Solanto et al. 2001a). This suggests the presence of at least two independent pathways leading to the development of ADHD. There is growing evidence to suggest the presence of further pathways including deficits in timing (Toplak et al. 2003), and working and non-working memory (Rhodes et al. 2004). And there is emerging evidence to suggest that these too are often independent from and uncorrelated with each other (Nigg et al., in press). Further, it appears that not all individuals with ADHD (even when they present with severe combined subtype ADHD) manifest cognitive deficits factors (e.g., Coghill et al., unpublished data; Nigg et al., in press). Although this in no way diminishes the potential importance of the dynamic developmental theory to our understanding of ADHD, it is highly suggestive that ADHD may be the developmental outcome of a variety of anomalies in separable neural networks, including several beyond the frequently emphasized fronto-striatal/executive networks in the brain (Rhodes et al., in press). Thus, whilst the deficits proposed by Sagvolden et al. may be able to account for many ADHD symptoms, this does not mean that they do so in all cases. In view
Selectionism: Complex outcomes from simple processes

John W. Donahoea and José E. Burgosb

aDepartment of Psychology, University of Massachusetts at Amherst, Amherst, MA 01002; bCentro de Estudios e Investigaciones en Comportamiento, University of Guadalajara, Guadalajara, Jalisco, 45030, Mexico. jdonahoe@psych.umass.edu jburgos@cucba.udg.mx http://euryale.sbs.umass.edu/psych www.ceic.cucba.udg.mx

Abstract: Both the target article and the precommentary demonstrate that relatively simple biobehavioral processes have the cumulative effect of fostering behavioral outcomes characteristic of attention-deficit/hyperactivity disorder (ADHD). As such, the articles illustrate a central theme of Darwinian thinking — basic processes acting over time can produce complex and diverse outcomes. In this commentary, we indicate that tracing the action of processes over time can be facilitated by quantitative methods such as artificial neural networks.

The target article by Sagvolden et al. and the precommentary by Catania illustrate a common general theme: Basic neural and behavioral processes can produce diverse and complex outcomes when they act over time. This theme exemplifies a central insight of Darwinian thinking — complexity can result from repeated action of relatively simple processes (Campbell 1974; Donahoe 2002). Sagvolden et al. describe the cumulative effects of a dysfunction in fronto-striatal circuits involving the neuromodulator dopamine and its resulting impact on the ability of organisms to tolerate temporal delays between behavior and reinforcers. Catania explores further the cumulative effects of differences in the ability to tolerate delay of reinforcement and reveals additional implications for the emergence of complex behavior. Together, these authors provide an interpretation of how a seemingly minor dysfunction can lead to many of the complex characteristics of attention-deficit/hyperactivity disorder (ADHD), including a lack of sustained attention, hyperactivity, and impulsiveness.

The target article and precommentary also jointly illustrate a distinction articulated by B. F. Skinner between experimental analysis, whereby basic processes are identified through carefully controlled laboratory observations, and scientific interpretation, whereby the implications of basic processes are explored for phenomena that cannot be studied under circumstances that meet the demands of experimental analysis (Skinner 1957). Sagvolden et al. draw upon experimental analyses at the cellular level of the processes that affect synaptic efficacies, while Catania makes use of experimental analyses at the behavioral level of processes that affect delay of reinforcement. Using these processes, the authors provide plausible interpretations of their effects on ADHD.

Scientific interpretation differs from mere speculation: Interpretation makes use solely of processes and structures that have been identified in prior experimental analyses. Interpretation is the means by which explanation occurs in all historical sciences — evolutionary biology and cosmology as well as behavioral neuroscience. If the cumulative effects of basic processes are sufficient to account for a complex phenomenon, then the interpretation is tentatively accepted as the explanation of that phenomenon. As an example from physical science, if the cumulative effect of gravity on a swirling cloud of primordial dust particles can account for the formation of planetary systems, then the interpretation is accepted as the explanation of planetary formation even though nascent planetary systems have not been subjected to experimental analysis — and likely never will be.

In early Darwinian interpretations of evolution, ordinary language was used to explore the implications of the process of natural selection. However, natural selection provided an interpretation of evolution that even other scientists did not find compelling until ordinary-language interpretations were supplemented by the more formal methods of population genetics (Fisher 1930; Haldane 1931/1966; Wright 1939) and, later, computer simulations (e.g., Maynard-Smith 1982). A similar transition in the nature of scientific interpretation is occurring in biobehavioral research with the advent of such techniques as artificial neural networks. Artificial neural networks that qualify as scientific interpretations must be informed and constrained by experimental analyses of neuroscience and behavior. In this respect, such networks differ fundamentally from superficially similar methods in normative psychology. In normative psychology, the characteristics of neural networks are inferred from the behavioral observations that they seek to explain and are not informed by independent experimental analyses. Behavior analysis is uniquely positioned to interface with neuroscience because, if behavioral processes are to be supplemented by infrabehavioral processes, those processes must be the fruits of independent experimental analysis at the neural level and not mere inferences from behavior (Donahoe 2002).

One approach to artificial neural networks in behavioral neuroscience is selection networks (Donahoe 1997; Donahoe & Palmer 1994; Donahoe et al. 1993). Selection networks consist of two inter-related neural subsystems: a motor subsystem that simulates the effects of the neuromodulator dopamine on synaptic efficacies of neurons in the frontal lobes and a sensory subsystem that simulates the effects of the hippocampus on synaptic efficacies of neurons in the parietal and temporal lobes. Only the motor subsystem is considered in the present interpretation of delay of reinforcement.

Simulated increases in synaptic efficacies in the motor subsystem occur when pre- and postsynaptic units are recently coactive and their coactivity is accompanied by activation of units in the simulated ventral tegmental area (VTA). Coactivity of units in the motor subsystem results from the action of glutamate on postsynaptic striatal connections that play a central role in the account of delay of reinforcement provided by Sagvolden et al.

Figure 1 (Donahoe & Burgos). A minimal neural architecture of a selection network. Environmental events stimulate sensory units that probabilistically activate units in the sensory-association subsystem. These units, in turn, probabilistically activate motor-association and output units in the motor subsystem. Activated output units simulate the emission of the operant response (R) and the elicitation of the reinforcer-evoked (i.e., unconditioned) response (UR). The circled MA-to-VTA connections simulate the fronto-striatal connections that play a central role in the account of delay of reinforcement provided by Sagvolden et al.
naptic AMPA and NMDA receptors. The pathways that simulate the release of dopamine from neurons in the VTA are represented by the grey regions in the motor subnetwork of Figure 1. The motor subnetwork includes motor-association (ma) units and output units. VTA units are activated by environmental events that stimulate the input unit for the reinforcing stimulus (SR). The neural processes and structures simulated by selection networks are informed by the research findings of the sort summarized by Sagvolden et al. (see Donahoe 1997; Donahoe & Palmer 1994; Frey 1997). Previous simulation research has shown that selection networks can simulate important aspects of a wide range of behavioral phenomena including acquisition, extinction, discrimination, timing, and revaluation (Donahoe & Burgos 1999; 2000). Of interest here are the effects of the pathways from ma units to VTA units, for these pathways are critical to conditioning with delay of reinforcement.

As Catania noted, delay of reinforcement is inevitable at the cellular level. The intracellular events that are essential for changes in synaptic efficacy endure for only a few hundred milliseconds within the dendritic compartments. This interval is much less than the irreducible delay between the occurrence of an operant (e.g., lever pressing) and the delivery of a reinforcer (e.g., a food pellet). In recognition of this constraint, before operant conditioning is instituted in the laboratory some distinctive stimuli (e.g., a "click" produced by the operation of a pellet feeder) is repeatedly paired with a reinforcer. Then, when operant conditioning begins, the distinctive stimulus is presented immediately after the occurrence of the operant. The neural solution to the problem of delay of reinforcement is accomplished in selection networks by means of connections from ma units to VTA units. These simulate the fronto-striatal pathways identified by Sagvolden et al. as critical for interpreting ADHD. As conditioning proceeds, VTA units become activated not only by the reinforcing stimulus (SR) but also by ma units that are activated by the effects of environmental stimuli on input units (e.g., S1).

Figure 2 displays the results of the simulation of operant conditioning with two different delays of reinforcement—a short delay of three time-steps and a longer delay of six time-steps. During each time-step, the simulated synaptic efficacies were changed according to the learning rule. Changes were a function of the magnitudes of the activations of the pre- and postsynaptic units and of the VTA unit that simulated dopaminergic release during that time-step (see Donahoe et al. 1993 for details.) The input unit for the reinforcer (SR) was stimulated at the final time-step when the activation level of the operant output unit (R) exceeded zero. The upper panels show simulations when the ma-to-VTA connections were intact. Under these conditions, the VTA could exert its reinforcing effect on synapses in the motor subnetwork when the VTA was activated either by the environmental reinforcer (SR) or by MA units (via the circled connections) throughout all time-steps. As can be seen, acquisition occurred with both delays of reinforcement, albeit somewhat earlier and more rapidly with the shorter delay. The lower panels show acquisition with the same two delays but with no connections from ma to VTA units; that is, without the conditioned reinforcement supplied by the activity of ma units. With the short delay of three-time-steps, acquisition continued to occur although somewhat more slowly and more variably. In contrast, acquisition failed to occur with the longer delay of six time-steps when the ma-to-VTA connections were absent. (Other simulations with as many as 3,000 training trials did not respond above the levels shown in Fig. 2.) Thus, the simulations provide a neural-network interpretation that supports the critical role assigned to fronto-striatal pathways by Sagvolden et al. Neural-network simulations of behavioral phenomena are at an early stage of development and are incomplete in many respects. Nevertheless, they already show promise of providing compelling, biobehaviorally informed interpretations of complex behavior.

A comprehensive and developmental theory of ADHD is tantalizing, but premature

Canan Karatekin
Institute of Child Development, University of Minnesota, Minneapolis, MN 55455. karat004@umn.edu
http://education.umn.edu/ICD/KaratekinLab/Default.html

Abstract: In this commentary, I argue that the theory presented by Sagvolden et al. can be much stronger if its scope is limited, if its developmental aspects are refined, if it can be made to generate testable predictions, and if it can be supported with more data from humans.

The theory presented by Sagvolden et al., based in part on work by Catania, merits attention for several reasons. The field of attention-deficit/hyperactivity disorder (ADHD) sorely needs solid theories from which testable predictions can be derived. The theory promises to be an overarching developmental theory. It tackles an important problem in ADHD. Finally, Sagvolden et al. discourage the use of behavioral intervention techniques (defined as frequent and immediate reinforcement) as impractical and advocate the use of medications for both children and their parents.

Although tantalizing, an overarching and developmental theory of ADHD is premature. In this commentary, I focus on several weaknesses of Sagvolden et al.’s theory, but end with the hope that it will generate more research.

The theory is too comprehensive. The theory is based on the finding that the dopamine (DA) system is implicated in ADHD, and on the premises that DA dysfunction in certain fronto-striatal circuits causes ADHD and that the most important behavioral functions of these circuits are reinforcement of novel behavior and extinction of previously reinforced behavior. Sagvolden et al. argue that these two behavioral processes “cause” ADHD and can explain all the symptoms of ADHD, as well as individual and developmental differences in the nature and severity of symptoms, and can be used to link molecular to societal levels of explanation. However, children diagnosed with ADHD are very heterogeneous, and it is easy to bring up examples of findings that are not easily explained by the theory (patterns of comorbidity across the
lifespans and findings of restless behavior even during sleep, to name a couple). Therefore, the authors are setting themselves up for failure by proposing a single-neurotransmitter, two-process model to explain the full complexity of the disorder. Because the theory tries to explain so much, Sagvolden et al. end up making too many promises (in sects. 1.1.1 and 1.1.2) that are not followed up on later, and overly general predictions such as “the exact ADHD symptoms at a particular time in life will vary and be influenced by factors having positive or negative effects on symptom development” (target article, Abstract).

The theory needs to be refined further and supported with data. There is a tension throughout the Sagvolden et al. article between trying to impose boundaries on the theory and using it to explain the whole disorder. For instance, Sagvolden et al. acknowledge the possibility of dysfunction in other neurotransmitter systems and complex interactions among these systems, yet they propose that their theory can explain all the symptoms of ADHD. This makes it difficult to test the theory. Likewise, they note that “the present model may be applicable mainly to a subgroup of ADHD linked to dopamine hypofunction” (sect. 3, para. 2). To state that a theory of DA dysfunction is limited to a hypothetrical subgroup with DA dysfunction makes it untestable. It is also not clear why Sagvolden et al. give primacy to the mesolimbic branch of the DA system and not functions in certain parts of the article and why they place equal emphasis on all three branches in other parts. Why not, for instance, propose that deficits in attention or habit learning can explain response to reinforcers?

On the one hand, the theory attempts to incorporate all levels from molecular to societal and complex constructs like creativity and self-esteem. On the other hand, it seems to view humans as passive organisms at the mercy of stimuli controlled by others. This tension is evident in the lack of clarity of the definitions of reinforcing and aversive stimuli. The behavioral aspects of the theory are based on rats and pigeons, and it is plausible to assume that a food pellet has the same value for animals kept hungry for the same number of hours. What is reinforcing becomes harder to define a priori from the participant’s point of view when it comes to humans. For example, could group differences in delay gradients be explained by differences in the value placed on the reinforcer or perceptions of aversiveness? Similarly, the theory seems to be limited to tangible reinforcers or aversive stimuli whose effects can be measured on a scale of seconds. What about social reinforcers, promise of reinforcers or aversive consequences in the long run, and internal reinforcers such as feelings of accomplishment? Do these have the same delay gradients and neurobiological substrates as tangible reinforcers, and how can their effects be teased apart in studies with humans?

Those of us who work within a cognitive framework typically use experimental and control conditions with identical reinforcement contingencies and temporal structures but different cognitive demands (e.g., anti- vs. prosaccades) and often find differential impairments in participants with ADHD compared to controls. How can the theory explain these differences in a way that is testable? Both Catania and Sagvolden et al. argue that purported deficits in attention and higher cognitive functions can be better accounted for by simpler behavioral processes. However, explanations in terms of higher cognitive processes do not stand a chance in studies of rats. To make their argument stronger, the authors need to produce data showing that their theory is also superior in explaining human data.

How is ADHD different from other disorders in which the DA system is implicated? Is there any evidence of alterations in reinforcement processes in these disorders? Are there any data on the sensitivity or specificity of these deficits to ADHD?

The theory is largely unconstrained by data from humans. No empirical data are presented that (1) link specific behavioral deficits to specific neural circuits in humans with ADHD, (2) show these deficits can explain the symptoms and the multitude of cognitive impairments; (3) show that the proposed behavioral deficits are present across a range of situations, reinforcers, samples, and ages; and (4) show that psychostimulants ameliorate these deficits. In addition, Sagvolden et al. do not always make clear distinctions among clinical observations, empirical data gathered for other purposes that support the theory, empirical data gathered specifically to test their theory, and predictions derived from the theory. For example, the statement “the theory also predicts increased behavioral variability” (sect. 1.1.2) sounds more like an observation than a testable prediction.

To be convincing, the authors ultimately need to test their theory directly against alternative explanations, and demonstrate that their theory provides a better account of the data. This has yet to be done.

The theory is not very developmental. The authors suggest that there are early abnormalities in the mesolimbic system in ADHD resulting from genetic or environmental factors, and that these abnormalities cause behavioral abnormalities early in life. However, no data are presented on early deficits in the mesolimbic system or reinforcement processes in ADHD.

There is no discussion of normal developmental changes in fronto-striatal circuits, and no studies are presented that investigated age-related changes in reinforcement processes in healthy or ADHD groups.

Regardless of the progression of deficits, Catania states that small differences early in life can have large effects later on. Sagvolden et al. predict that “the development and severity of symptoms are linked to degree of dysfunction in the various dopaminergic systems” (sect. 3, para. 2) and “behavior and symptoms in ADHD result from the interplay between individual predispositions and the surroundings” (target article, Abstract). Yet, how the specific processes they propose account for the variety of individual differences in onset, trajectory, comorbidity, or prognosis is not well specified. Descriptions of age-related changes in behavior (e.g., “impulsiveness will . . . be manifested differently at different ages”; sect. 3.6, last para.) are not predictions, and the authors do not cite any studies designed to test their specific developmental predictions.

Sagvolden et al. state that DA dysfunction “probably lasts[s] for life” (sect. 4.1, para. 4). Consequently, immediate and frequent reinforcers are viewed as crucial for treatment regardless of age, and there is no discussion of possible age-related changes in the effectiveness or mechanism of action of psychostimulants.

Sagvolden et al. also do not consider the possibility that even if there is an early brain dysfunction, it does not need to be manifested early in life if the affected circuitry has a protracted period of development. Similarly, an early dysfunction in a specific neural circuit could lead to adaptive or compensatory processes in other circuits, so deficits manifested later in life may not necessarily be linked directly to the original dysfunction.

In short, the theory seems to make the questionable assumption that the DA system and the behavioral processes it supports are fully mature early in life, and that the same mechanisms operate in the same manner from infancy through adulthood.

Nevertheless, the theory has potential. The theory could be stronger if its scope were more limited and if it could generate specific, testable predictions. Among the strengths of the theory are that its core premises are plausible and worth testing. It is also very impressive in that it can generate quantitative predictions, and that it can be tested in animals. Importantly, it has the potential to stimulate theory-driven research on the development of the fronto-striatal circuits and reinforcement processes, and prospective longitudinal studies of children with, or at risk for, ADHD. The gathering of this kind of evidence can greatly advance understanding of the nature of ADHD and facilitate the development of effective treatments for the disorder.
Gradus ad parnassum: Ascending strength gradients or descending memory traces?

Peter R. Killeen
Department of Psychology, Arizona State University, Tempe, AZ 85287-1104. killeen@asu.edu http://tops.asu.edu/dresearch/blab.html

Abstract: Decay gradients are usually drawn facing the wrong direction. Righting them emphasizes the role of stimuli that mark the response, and leads to different inferences concerning the factors controlling response–reinforcer associations. A simple model of the concatenation of stimulus traces provides some insight to the problems of impulse control relevant to ADHD.

The target article constitutes an important synthesis of behavioral and biological causal factors for ADHD. It, and the precommentary, offers the promising and provocative hypothesis that, inter alia, dopamine deficits shorten and steepen the delay of reinforceability gradients (bottom). This is a difference that makes a difference. At a delay of 20 seconds, doubling the magnitude of reinforcement might improve conditioning; but the trace of the response is so weak compared to more recent stimuli and responses that much of that increased magnitude is more likely to benefit behavior other than the target response. Contrast this with operations that change the strength of the response trace. Doubling memorability at the time of the response will double memorability 20 seconds later. Even though the absolute increment at 20 sec will be much less than at 0 sec, all of it will be vested in the target response. Conversely, in situations where memorability of the response is degraded (Bottom curve, Fig. 1), the trace, and thus the reinforcer’s ability to strengthen the response, may fall below the noise level.

The literature supports this distinction. Liberman et al. (1985) showed that the presence of a light flash after a response greatly enhanced acquisition. Williams (1999) showed that such marking was much more effective in the differential acquisition of a response than having the same marker signal onset of reinforcement – and thus act as a conditioned reinforcer. In fact, the conditioned reinforcer impeded conditioning. There are three morals to this story.

1. Marking a response when it is made can facilitate conditioning.
2. Bridging stimuli intended as conditioned reinforcers might actually shorten the reach of reinforcement rather than lengthen it, as desired for behavior therapy of ADHD.
3. Dopamine released at the time of reinforcement is more likely to strengthen consummatory rather than instrumental responding. However, the dopamine released when a response has stimulus concomitants – is marked – would strengthen instrumental conditioning. All forms of conditioning are enhanced in an aroused organism (Killeen 1975), perhaps as a result of sensitized response-dependent release of dopamine.

Popular models of self-control are also exemplified with backward gradients, and they support inferences of relevance to ADHD. Most organisms choose a larger or better reinforcer over a smaller or inferior reinforcer. When the larger reinforcer is sufficiently delayed, preference switches to the smaller, more immediate reinforcer. This might be construed as a rational choice by organisms that attribute higher value to the soon-small outcome; organisms that attribute higher value to the soon-small outcome;

Figure 1 (Killeen). Decay of reinforcement gradients (top) are more properly called delay of reinforceability gradients (bottom). If memorability of the response is strengthened by marking, or weakened by conditions such as ADHD, the ability of a reinforcer to strengthen behavior is affected accordingly.
tends over a longer and longer interval as the delay to reinforcement, $td$, increases. The density of reinforcement in the presence of the memory of the CS may be calculated by dividing the saturation level by $td$ (Killeen 2001a; 2001b). If the trace gradient is exponential with rate of decay of $\gamma$, then the strength of the conditioned reinforcer is given by either:

$$S = \frac{\int_0^t e^{-\lambda t} dt}{td} = \frac{1-e^{-\lambda t}}{\lambda t_d} \quad (1)$$

or

$$S = \frac{\int_0^t \lambda e^{-\lambda t} dt}{td} = \frac{1-e^{-\lambda t}}{td} \quad (2)$$

These two forms correspond to the two types of (reversed) traces shown in Catania’s Figure 4. Equation 1 hinges the gradient at $\lambda$ when $td = 0$: Variations in the rate of decay of the trace do not affect the strength at zero delay (see Fig. 3, open symbols). Equation 2 hinges it at $\lambda$ to maintain a constant area under the curve. When these equations are embedded in a more fully articulated model, the presence or absence of the rate constant in the denominators is absorbed by other constants. But in cases where the rate parameter $\lambda$ is itself under consideration, as in the target article, the differences are important. If individuals with ADHD have steepened gradients, Equation 1 predicts that at long delays conditioned reinforcers will be devalued by the larger value of $\lambda$; Equation 2 predicts that steepened gradients would have little differential effect at long delays, but would actually be beneficial at shorter delays due to the quicker saturation of memory (see Fig. 3, filled symbols). Individuals with ADHD have difficulty deferring gratification – difficulty in ordering their behavior with respect to delayed outcomes, despite an apparent general desire to do so – and no obvious advantage at short delays, suggesting that Equation 1 may be the correct form. Figure 4 applies Equation 1 to Catania’s data, showing that it is not easily discriminated from the inverse “hyperbolic” gradient often used to fit such data.

Having established Equation 1, it may be developed to address the self-control paradigm – that is, changes in organisms’ preference for the larger delayed reinforcer as the delay to that reinforcer increases. The proportional strength of CSs signalling different delays and amounts of reinforcement may be written as

$$P = \frac{v_1 S_1}{v_2 S_1 + v_2 S_2},$$

where $v_i$ is a constant reflecting the value of the reinforcer, and $S_i$ is the strength as inferred from Equation 1:

$$P = \frac{v_1 (1-e^{-\lambda t_1}) / t_1}{v_2 (1-e^{-\lambda t_2}) / t_2} \quad (3)$$

Because the rate constant cancels out of the denominators, the same prediction also follows from Equation 2. In the case where the delay to the small reinforcer is constant, the right addend in the denominator may be assigned a constant value, such as 1.0, giving:

$$P = \frac{v_1 (1-e^{-\lambda t_1}) / t_1}{v_2 (1-e^{-\lambda t_2}) / t_2} \quad (4)$$

Equation 4, and the more general Equation 3, provides a map to the data of self-control experiments, parsing the effects into incentive value, or valence ($v_i$) and rate of gradient decay ($\lambda$). Equa-
The SHR Lo group had about the same
the latter.
up to a reinforcer, or down from a response. A case was made for
tary issue I discussed in this commentary is whether the steps lead
tation 4 is applied to the interesting data of Adriani et al. (2003),
shown in Figure 5. These authors found large intersubject vari-
bility in the performance of SHR (spontaneously hyperactive)
rats given the choice between a small immediate reinforcer and a
delayed one. They therefore did a median split on the over-
all preference to yield the graph shown in the right panel. A sim-
ilar median split on the control animals yielded very different pro-
files. Equation 4 drew the curves through the data, yielding
estimates of the two key parameters. For the WKY (Wistar Kyoto
Rat), all of the effect was due to variation in the subjective value
of the reinforcers, the Hi group preferring the large reward twice
as much as the Lo group, with \( \lambda \) remaining constant at 0.2 s \( \sim \lambda \).
The SHR Lo group had about the same \( v \) and \( \lambda \) as found in the
WKY Lo. The SHR Hi group preferred the large reward six times
as much as the Lo group and had a much flatter delay gradient (\( \lambda
= 0.04 \)). These data give no support for steeper gradients for the
SHR strain, nor for failure of impulse control, but rather, underscore
the high variability of operating characteristics in these pop-
ulations, and the need for care when drawing inferences from
pooled data.

Conclusion. The Gradus ad Parnassum – Steps to Parnassus –
was a guide to the elements of Greek and Latin for those who
would write proper prose. The aforementioned considerations
concerning delay gradients are also elements that only find their
meaning in a larger theoretical text, such as that provided by the
target article and precommentary, and by Figure 5. The elemen-
tary issue I discussed in this commentary is whether the steps lead
up to a reinforcer, or down from a response. A case was made for
the latter.

ACKNOWLEDGMENTS
The analysis and commentary was made possible by the generous support
of the Norwegian Center for Advanced Study and by NSF IBN 0236821
and NIMH 1R01MH066860. I thank Jonathan Williams for comments on
a draft.

Figure 5 (Killeen). Preference for a large (five pellets) over a small (one pellet) reinforcer as a function of the delay to the larger. Median splits on preferences yielded different characteristics for the two strains. These are parsed by Equation 4 as differences in valance of the large reward for the two WKY groups, with both groups having the same rate of decay (\( \lambda = 0.2 \)). For the SHR strains both va-
ences and gradients (\( \lambda = 0.04, 0.2 \)) differed. (Adriani et al. 2003)

ADHD, comorbidity, synaptic gates and re-entrant circuits

Florence Levy
School of Psychiatry, University of New South Wales, Prince of Wales
Hospital, Sydney, New South Wales, 2031, Australia. f.levy@unsw.edu.au

Abstract: The “dynamic developmental” theory of attention-deficit/hyperactivity disorder (ADHD) has come full circle from Wender’s (1971) reinforcement hypothesis. By specifying the principle of time constraints on reinforcement and extinction, the present theory allows for empirical validation. However, the theory implies, but does not discuss, implications for the neurophysiology of comorbidity in ADHD. The authors’ attribution of comorbid oppositional behavior to parental and societal reinforce-
ment leaves out biological factors.

Sagvolden et al. are to be congratulated on their comprehensive “dynamic developmental” theory of attention-deficit/hyperactivity

disorder (ADHD) (predominantly hyperactive/impulsive and combined subtypes), which integrates behavioural analysis with
neurobiological factors. The authors describe the consequences of a hypo-functioning mesolimbic dopamine circuit as altered rein-
fforcement of behavior and deficient extinction of previously re-
forced behavior. In some ways the theory has come full circle from
Wender’s (1971) theory of minimal brain dysfunction, which post-
ulated a reinforcement deficit. However, the strength of the dy-
namic developmental theory is that it is based on empirical animal
studies in addition to clinical observations. Sagvolden et al. quote
the three-factor Hebbian learning rule (Hebb 1949) that synaptic
transmission is facilitated when presynaptic input, postsynaptic
activation, and the dopamine signal occur simultaneously at the
same neuron. Thus, the effect of a reinforcer is more potent when
the delay between the response and the reinforcer is short rather
than when the delay is long (delay of reinforcement gradient). In
ADHD children, a steeper “delay of reinforcement gradient” al-
lows a shorter time window for effective reinforcement contin-
gencies, accounting for the necessity for immediate reinforce-
ment to control the impulsive and hyperactive behavior of ADHD
children.

Although the elaborated theory is comprehensive and heuristic in terms of parental and societal influences, and allows for emip-
rical validation, it leaves out important aspects of the behavior of
ADHD children. The authors attribute the frequent co-occur-
rence of oppositional and conduct disorders in ADHD children to
aberrant learning and insufficient extinction of antisocial behaviours, as a result of the fundamental reinforcement/extinction deficit combined with inadequate parental and societal structuring of ADHD children’s behaviour. However, reinforcement/extinction per se may not be a sufficient explanation of why some children respond to verbal requests and/or reassurance and why others refuse to comply.

Current ADHD theories have postulated a deficit of inhibition (Barkley 1997a; Quay 1988). Sagvolden et al. challenge the more current theories of ADHD and characterise inhibition as a fundamentally vague and circular concept, which is more usefully replaced by the concept of synaptic gating (Grace 1995, Levy 2004). While the dynamic developmental theory does not directly address the issue of comorbidity in ADHD, it provides a theoretical basis for understanding the pathophysiology of the core ADHD symptoms described above, which may also help to understand comorbidity. The authors draw on the work of Grace (1995; 2000a; 2000b), who showed that accumbens neurons exist in a bistable state, with their membrane potential alternating between a hyperpolarized non-firing state and a depolarised plateau lasting several hundred milliseconds, during which spike activity is generated. This bistable accumbens state allows the operation of a synaptic gating mechanism between cortical and limbic (emotional) influences on behaviour. The nucleus accumbens receives input from a number of limbic-related cortical structures, including the prefrontal cortex, hippocampus, and amygdala (Grace 2001). In particular, the hippocampus and amygdala strongly influence the ability of the prefrontal cortex to activate accumbens cell firing, allowing an emotional override to the executive system.

Goto and O’Donnell (2002) have reported timing-dependent limbic-motor synaptic integration in the nucleus accumbens (NAcc). They found that synaptic inputs from the prefrontal cortex and limbic structures interacted differently depending on their timing. Coincident inputs were likely to enhance information transmission by reducing excitatory postsynaptic potential (EPSP) amplitude variability, whereas asynchronous inputs depend on the order of arrival. Prefrontal inputs tended to dampen limbic responses, whereas limbic inputs allowed subsequent prefrontal responses by exhibiting a linear decrease in EPSP amplitude at more depolarised membrane potentials. PFC inputs were most effective in the NAcc at depolarised membrane potentials (Up state), whereas limbic membrane inputs were effective primarily during a resting membrane potential (Down state). The authors concluded that these two simultaneous mechanisms by which input (and response) selection can take place in the NAcc, depending on the state of the neurons and timing of inputs provide a mechanism for attention and emotional or motivation factors that affect responses to stimuli, with an important role in cognitive function.

A further implication of reciprocal amygdala/hippocampal/prefrontal relationships may be found in the neuroanatomical work of Heimer (2003). Heimer has described “a new anatomical framework for neuropsychiatric disorders and drug abuse.” Improved electron microscopic methods have allowed the demonstration of a ventral cortical-striatal-pallidal system. This circuit (which includes accumbens/ventral striatum) is one of three re-entrant circuits “anterior cingulate, lateral orbital frontal, and medial orbital frontal, related to the ventral emotional-motivational striatal domain.” According to Heimer, the ventral striatal-pallidal pallidal system and extended amygdala are major components of the new anatomy of the basal forebrain. “Since the entire cerebral cortex, including the hippocampus, the olfactory cortex and major parts of the amygdala project to the basal ganglia, all major telencephalic disorders are to some extent at least disorders of the basal ganglia” (p. 1737).

For present purposes, the demonstration of ventral striatal-hippocampal-prefrontal re-entrant circuits (including accumbens) allows for the possibility of iteration of emotional reactions from amygdala through hippocampus and prefrontal cortex, allowing executive monitoring of emotional behavior. Thus, impaired synaptic gating at integrative locations such as the accumbens will interfere with the development of controlled behavior. Sagvolden et al. state that behavior is gradually brought under discriminative control, including the establishment of verbally governed behavior as a result of training. They describe verbal stimuli as contingency specifying stimuli. However, rather like Chomsky’s (1959) criticism of Skinner’s explanation of language, the operant explanation of verbally governed behavior does not explain the sometimes immediate and dramatic changes in oppositional behavior of ADHD children on stimulant medication. These changes require a biological explanation, which may relate to integration at the above re-entrant circuits and between these circuits. This does not diminish the important role of language in human development. In ADHD, language deficits may well limit the scope of re-entrant circuits in the elaboration of behaviour.

What is the purpose of a new behaviorally based dynamic developmental theory of ADHD? The perspective of the educational psychologist

Paolo Moderato and Giovambattista Presti
Department of Psychology, University of Parma, 43100 Parma, Italy.
paolo.moderato@unipr.it nannip@tiscali.it

Abstract: In Sagvolden et al.’s conceptualization of how a poor behavioral, social, and academic repertoire arises from an impaired interaction with the environment of an individual with a neurological disorder, we see a convergence between the medical diagnosis and the functional assessment on which the behavioral educational approach is based. If children with such a disorder do show delay-of-reinforcement steepened gradients, it is possible to predict their behavior under given circumstances. This could bring us to more precise diagnostic criteria and better intervention technologies.

In the advancement of science, literature reviews accomplish a fundamental role: Occasionally they try to sum up the state of the art, that is, what is known on a certain subject at a certain point in time, and try to point up new insights and suggestions to understand phenomena. The target article by Sagvolden et al. thus prompts the fundamental question: Are we in need of a new point of view on the attention-deficit/hyperactivity disorder (ADHD) syndrome?

For many years, neuroscientists, psychiatrists, and psychologists have been trying to shed light on the deficits supposedly underlying attention-deficit/hyperactivity disorder (ADHD), with recent studies mostly supporting the idea that ADHD is a result of deficits in executive control and regulation that influence emotional and cognitive processes (Barkley 1998). This approach has been so influential that, starting from the end of the 1980s, ADHD has been commonly regarded as resulting more from neurological and genetic factors than from environmental events.

However, a medical diagnosis of ADHD does not necessarily imply that all the children with it show the same degree of disability, and a functional assessment is needed to fully address their behavioral and academic repertoires, if a rehabilitative intervention is to be implemented. Educational approaches based on functional analysis of the behavioral and academic repertoire, though pursuing a parallel but not strictly related path to the neurosciences, have been reported to work well. In literature a wide array of interventions planned to modify children with ADHD behavior ranging from behavioral procedures such as token economies (e.g., Williams et al. 1989), daily report cards (e.g., Burkist et al. 1987), self-monitoring (e.g., Edwards et al. 1995), verbal praise (e.g., Williams et al. 1991) and contingency contracting (e.g., Newstrom et al. 1999), to cite just a few, have been published. These and other procedures have been found to be effective in enhancing school performance and social behaviors of children with ADHD.
The two levels of analysis (neurosciences and behavior) and interventions (pharmacological and educational) have evolved strictly separated, though aiming at the same target: understanding the disorder and providing ways to deal with it. Sagvolden et al.'s article gives new insight and useful suggestions for dealing with ADHD, being able to correlate either the behavioral repertoire or the neurological impairment, both at the level of the brain pathways and at the level the neurotransmitters.

If it is true that a neurological deficit exists, then nevertheless it translates into an impaired interaction of the individual with the surrounding environment. Sagvolden et al.'s conceptualization of how a poor behavioral, social, and academic repertoire arises from an impaired interaction with the environment of an individual with a neurological disorder is consistent with a behavior analytic vision of development (see e.g., Bijou 1966). However, it is, to our best knowledge, the very first time that a thorough and coherent picture is given at both the levels of structure and of function.

The approach that is suggested can be easily translated into better prediction and control. If these children show delay-of-reinforcement steepened gradients, it is possible to predict their behavior under given circumstances. This could bring us to more precise diagnostic criteria, as suggested also in Catania's preconomy. In the point made by Sagvolden et al., we do see a convergence between the medical and the functional assessment on which the behavioral educational approach is based.

At the beginning of the 1980s these two ways of looking at developmentally retarded individuals who showed a common pattern of behavior were definitely separated. In most cases they were two antagonistic ways of conceptualizing behavior disorders. Although in some fields of the psychiatric domain, for example, anxiety or depression, functional analysis and medical diagnosis continue to be separate, this is not the case when looking at genetically based developmental disorders like fragile X or Asperger syndrome. The structural description of the behavioral phenotype, given at the medical level, is complemented by the functional description of behavior. Though not specifically linked to a precise genetic change, the same, we think, could apply to ADHD.

Up to now there has been no concrete medical test to diagnose ADHD, which often makes the diagnosis of ADHD subjective. Vague criteria in diagnosis lead to confusion in epidemiology, so that the numbers of those diagnosed range up to 17%, as reported by 19 community-based studies in the past two decades (Scahill & Schwab-Stone 2000). The differences in epidemiological surveys are a consequence of the choice of informant, methods of sampling and data collection, and, above all, the diagnostic definition. Such a big number is not confirmed by our daily experience. One should observe three to four subjects per class of 20 to 25 pupils – which is not the case. This weakness in diagnostic precision exposes scientific procedure to easy criticism, which states that ADHD does not exist up to the point where it is necessary to publish consensus statements (Barley et al. 2002). The loose descriptive category of the DSM-IV (American Psychiatric Association 1994) or the AAP criteria (Herreras et al. 2001) can be better restricted on the behavioral level by registering the ADHD child behavioral pattern with operant procedures, distinguishing true explanation at a more parsimonious level than that of the deficits preceded and consequent conditions that are really effective for individuals with special needs. The concept of prosthetic environment has been very successful in helping people with physical impairments and disabilities, and in reducing their handicaps with regard to typically developing subjects. Less successful has been the concept of helping people with “mental” atypical development.

There are many reasons for that, including the idea of mind, the idea of freedom, and the difficulty in arranging a tailored prosthetic environment for cognitive disabilities. However, this is, in our opinion, the challenge of the future, and, fortunately, a computer-based world can afford it.

Every single behavior of our life is videotaped, computer recorded, analyzed, controlled, and so on. Does all this technology allow us to implement interventions to enhance the academic and social repertoire, making the environment conducive to the delivery of specific reinforcement?

To conclude: in our opinion, Sagvolden et al. provide us with an explanation at a more parsimonious level than that of the deficits in executive control and regulation, eventually taking into account higher levels such as attention delays and failure to outline goal-oriented instructions and rules of behavior, while maintaining a strict correlation with basic research, medical practice, and education or rehabilitation of individuals with ADHD.

The question that still remains on the ground is how to move this hypothesis from the animal lab to humans and experimentally test individuals with ADHD in comparison to those without. If Sagvolden et al.'s vision is empirically confirmed, we need to definitively refocus the theoretical approach to ADHD cited in the introductory part of this commentary. So, on the basis of all the points we have considered here, including the one in the preceding paragraph, we heartily answer our question without any doubt: Yes, we do need the new point of view.
Reinforcement gradient, response inhibition, genetic versus experiential effects, and multiple pathways to ADHD

Joel Nigg
Department of Psychology, Michigan State University, East Lansing, MI 48824-1116. nigg@msu.edu http://www.msu.edu/user/nigg/nigg.htm

Abstract: Major contributions emanating from Sagvolden et al.'s theory include elucidation of the role in attention-deficit/hyperactivity disorder (ADHD) of temporal information processing, social learning, and response extinction learning. Key issues include a need for clearer explanation of the relative role of impulsivity versus response suppression/inhibition in the dual process model, and delineation of genotype-environment correlations versus interactions in the social and experiential mechanisms posited.

Neurobiological theories of attention-deficit/hyperactivity disorder (ADHD) over the past several decades have addressed executive functioning (Barkley 1997b), arousal, vigilance, activation, attention, behavioral inhibition (faulty anxiety response; Quay 1997), and reward response/reinforcement learning (Gorenstein & Newman 1980), among other mechanisms.

Following that last tradition, a signature contribution of the Oslo group over two decades has been their specification of the role of a steepened delay-reward gradient in reinforcement learning in ADHD development. This basic insight has strong appeal for at least two major reasons in addition to its ready accommodation to the literature on catecholamine dysfunction in ADHD. First, it allows integration of observations about poor temporal processing in ADHD – something which theorists have wrestled to integrate – and thus helps resolve debates about over- versus under-responding to reward in ADHD. Second, it represents a crucially needed, and too often neglected, effort to develop a social-learning account of how ADHD might unfold during development. These should prove to be landmark contributions.

Well-developed responses to four key questions may sharpen the theory's testability. First, experiential effects could be specified in a clearer relation to behavioral genetic findings. When are the experiential effects to be understood as genotype by environment interactions, and when are they to be understood as genotype-environment correlations? If disrupted parenting occurs mainly when parents themselves have ADHD, then shared or nonshared environment effects should be noteworthy. However, high heritability is consistent with genotype-environment correlations (Scarr & McCartney 1983), which virtually assure that social learning or related processes will mediate some genetic effects. Meanwhile, the important hypothesis that environmental contaminants may influence ADHD can be reconciled with high heritability, because heritability as computed from biometric models of ADHD generally signifies heritability of liability. Liability can be potentiated by environmental triggers, perhaps via genotype-environment interactions (as a partial analogy, recall that tuberculosis showed robust heritability in twin studies a half century and more ago).

Second, a need remains for more clarity regarding response inhibition and impulsivity. After a worthy reminder about the dangers of conflating “inhibition” as a neural versus a psychological process, related distinctions falter to the argument's detriment. It is crucial to distinguish (a) impulsivity, or hasty, inaccurate responding in a slow, careful response context from (b) poor response suppression, or slow mobilization of a process by which a prepotent (prepared, cued, about-to-be-executed, or already-initiated) motor response is suppressed, cancelled, or inhibited in a rapid decision context (also called executive response inhibition; Logan & Cowan 1984; Nigg 2001). These may be related (Logan et al. 1997), or may be distinct (Sonuga-Barke 2002). Nigg (2000, 2001) reviews various meanings of “behavioral inhibition” and “cognitive inhibition” in detail, including the difference between cognitive and motor control mechanisms. Sagvolden et al.'s theory is at its best in explaining impulsive responses but does not as well account (at least via the reinforcement pathway) for strategic response suppression problems, which are well replicated (Nigg 2001) in ADHD. The latter depends upon intact inferior frontal gyrus (right prefrontal cortex) as well as striatal structures including the caudate (Aron et al. 2003; Casey et al. 2002), and therefore probably lies on the meso-cortical dopamine branch described by Sagvolden and colleagues. Their linkage, in places, of this ability with the nigrostriatal or mesolimbic circuits thus may benefit from closer reconciliation with key neuroscience data.

At times the authors seem to agree with this assessment and thus suggest a dual-process model consistent with the idea that (1) executive networks are related to response suppression – and then to the symptom domain of inattention-disorganization, whereas (2) reward networks and faulty reinforcement learning are related to the symptom domain of hyperactivity and impulsive responding (see Sonuga-Barke 2002). The theory would be strengthened by clarification as to whether it bets on this dual process account, versus placing response suppression with an altered reinforcement gradient on a pathway to hyperactivity-impulsivity.

Third, how do the psychological mechanisms described interact with one another, as well as with other systems not discussed? Reactive or motivated processes (such as reward responsivity) and strategic, effortful, or executive control processes are mutually regulating (Derryberry & Rothbart 1997) both during development and in dynamic action selection. This fact has to be considered in developmental accounts. As a second example, the theory only indirectly addresses the phenomenon of slow, inaccurate responding in fast response contexts in ADHD, which may be related to low cortical arousal defined as high-scalp EEG slow-wave activity (Barry et al. 2003). More frequent reinforcement (necessary for sustained attention and behavioral control in ADHD according to this theory) should heighten arousal (Gray 1982) and/or activation (van der Meere 2002), which might drive or mediate apparent reinforcement delay-gradient effects.

Finally, how will this theory, supported by elegant animal experiments, be tested in children? In particular, how will it be shown that during early development, children with ADHD are slow to learn reinforcement parameters, yet when they do, they then fail to extinguish the maladaptive behaviors that they attempt? Some children with ADHD may emit behaviors (which are then somehow reinforced) that other children have never attempted, or they may find reinforcing responses that other children do not. Such complexities and important predictions often will require additional theoretical replication and observational developmental studies that have yet to be undertaken. However, the most striking prediction is of failed extinction learning in ADHD. This prediction is unique to this theory, relatively easy to test, and in need of further independent replication, which could then justify more costly longitudinal evaluation.

Overall, this updated articulation of the reinforcement-gradient model helps the field integrate the growing neuroscience knowledge base about ADHD and suggests how the limbic-frontal neural loop contributes to ADHD’s hyperactive-impulsive symptom domain; especially welcome is a social-learning account of ADHD development. This theory in turn requires additional integration with what is known about executive, arousal, and other neural networks involved in ADHD so that more complete multipathway accounts of the syndrome can be developed.

ACKNOWLEDGMENT
This work was supported by National Institute of Mental Health Grants R01-MH59105 and R01-MH63146.

NOTES
1. Just as debate continues about whether strategic suppression of cognitions requires an inhibitory mechanism, one could ask whether the term inhibition is confusing or useful in the context of strategic (motor) response suppression. However, there is little doubt that the ability exists (Logan 1994), and that people with ADHD do poorly at it (Nigg 2001).
2. An animal model of this construct is difficult to achieve; it is not nec-
essarily arrived at by teaching animals to withhold all responding (though that ability may be related, and it is also deficient in children with ADHD).

ADHD theories still need to take more on board: Serotonin and pre-executive variability

Robert D. Oades and Hanna Christiansen

Abstract: Correcting the relationship between tonic and burst firing modes in dopamine neurons may help normalise stimulus-reinforcement gradients and contingent behaviour in attention-deficit/hyperactivity disorder (ADHD) children. But appropriate evaluations of stimuli for developing adaptive plans and controlling impulsivity will not occur without moderating the gain-like functions of serotonin. The “dynamic theory” correctly highlights the need to account for variability in ADHD. The differentiation of pre-executive information processing is proposed as an explanation.

At the core of the article by Sagvolden and colleagues is a set of data that throws light on an aspect of the ADHD phenomenon. But one asks if the authors are a measure too brave to generalise so broadly from data that lies outside the direct influence of DA. Perhaps the mere promotion of the likelihood of DA function in and around the synapse is helpful, but indirectly so.

Let us return to the “core of ADHD problems.” Sagvolden et al. (also Castellanos & Tannock 2002) highlight the variability of response as one of the central features of ADHD. A prediction of the “Dynamic Theory” is that variability arises from the “extra-nominal” class or unusual response pattern (for a given situation) becoming “the rule.” This is an interesting and unusual form of “perseverance,” but as such is consistent with hypo-dopaminergie function. The function that we note here is the role of increasing DA activity in initiating action or promoting the likelihood of a switch between competing actions, as proposed and demonstrated elsewhere (Oades 1985; 1997). Sagvolden et al. also expressly note that the function of burst-firing DA neurons lies in the initiation of behaviour (p. 15). Thus, to a degree, we agree on the so-called role of ascending DA activity. But we also note it is consistent with the more parsimonious idea of DA having a general role in competitive information processing rather than specific reinforcement (Oades 1999).

Important for the discussion here is that there is an alternative explanation for the variability of behaviour in ADHD. Namely, that there is an impairment in top-down control of processing incoming information. This control may be independent of DA, although it may be markedly influenced by an impaired role of noradrenaline in “tuning” different inputs. This viewpoint may account for a range of anomalous features of ADHD not incorporated by the impairment of perception and integration of reinforcement described in the Dynamic Theory.

There are examples from the control of attention. Steady state visual potential latencies in ADHD suggest a decreased efficiency in coupling between PRF networks, especially in the right hemisphere (Silberstein et al. 1998). Reduced speeds of conduction (Ucles et al. 1996) and delayed latencies (e.g., P1, N1; Johnstone et al. 2001; Karayanidis et al. 2000) in late developing regions (e.g., delayed myelination) would form a good basis for response variability and poor time perception (Rubia et al. 2003; Toplak et al. 2003). The delay could also account for the slowed orienting of attention by the right hemisphere to the left visual field and evaluation of the cue eliciting the orientation (Carter et al. 1995; McDonald et al. 1999; Oie et al. 1998). With such “inefficient coupling” it is no surprise that event-related recordings show poor differentiation of Go, No-go stimuli and errors in the stages of information processing that follow (Dimoska et al. 2003; Liotti et al., in press). Further, this hypothesis is consistent with data supporting an etiology in terms of a maturational lag (Cantwell 1995), with the delay particularly affecting those parts of the frontal lobe that develop last.

Part of our suggestion may be consistent with part of the “Dual Pathway” hypothesis (Sonuga-Barke 2003). Undoubtedly the “Dynamic Theory” with its emphasis on mesolimbic reinforcement mechanisms finds support from another part of the dual pathway. Sagvolden and colleagues make a major constructive contribution to the continuing need to try to account for all the data.
Changes in sleep-wake behavior may be more than just an epiphenomenon of ADHD

Aribert Rothenbergera and Roumen Kirouv
aChild and Adolescent Psychiatry, University of Göttingen, D-37075 Göttingen, Germany; bInstitute of Physiology, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria. arother@gwdg.de ru@bio.bas.bg
http://www.gwdg.de/~ukyk http://www.bio.bas.bg

Abstract: Sleep disturbances are common for children with attention-deficit/hyperactivity disorder (ADHD) and are of great clinical significance. Brain dopamine plays an important role for both ADHD symptoms and sleep-wake regulation. We therefore suggest that one basic aspect of integrative brain-behavior relationship such as the sleep-wake cycle may certainly be addressed in a dynamic developmental theory of ADHD.

Attention-deficit/hyperactivity disorder (ADHD) represents one of the most common, socially important, and scientifically debated child psychiatric disturbances (Buitelaar & Rothenberger 2004). Hence, we must be grateful to Sagvolden et al. for carefully analyzing a large amount of data concerning ADHD and proposing a unique functional schema to explain the complexity and heterogeneity of ADHD phenotypes. However, in a general theory of ADHD, one basic aspect of integrative brain-behavior relationship such as the sleep-wake cycle needs to be included.

Sleep disturbances are common for children with ADHD (Brown & McMullen 2001; Corkum 2001; Kirov et al. 2004; Kostanecka-Endress et al. 2000). The great impact of sleep-wake problems in ADHD is evidenced by more than 15 original and 7 review articles related to this topic that were published during the period of 2003–2004 in Medline. Sleep disturbances, including motor restlessness in sleep, not only have clinical importance but may be closely related to the appearance of ADHD symptoms during the day (Lewin & Di Pinto 2004). With this background, the question is whether and how Sagvolden et al.’s behaviorally oriented theory can explain or even predict the sleep-wake changes in ADHD.

So far, a direct answer cannot be derived from the data and explanations given in the target article. However, the authors support the hypothesis that hypofunctioning mesolimbic, mesocortical, and nigrostriatal dopamine branches play a pivotal role for the core symptoms of ADHD. Yet, dopamine has recently been suggested to be functionally involved in sleep-wake regulation. Some psychiatric disorders with dopamine alterations do manifest sleep variations and a number of dopaminergic agents can pharmacologically induce sleep-wake changes (Crochet & Salai 2003; Mignot et al. 2002; Rye & Jankovic 2002). Therefore, we would like to focus on the neurobiological mechanisms of sleep and ADHD and their clinical implications.

In particular, mesolimbic and mesocortical dopamine circuits are thought to be critical for modulating the quality, quantity, and timing of rapid eye movement (REM) sleep (Keating & Rye 2003; Reid et al. 1996). Hence, a specific sleep pattern in ADHD may be characterized mainly by changes in REM sleep. In support of this suggestion, recent studies have found changes in the amount and timing of REM sleep in children with ADHD (Crabtree et al. 2003; Golan et al. 2004; Kirov et al. 2004; O’Brien et al. 2003a; 2003b; 2003c). Furthermore, periodic leg movements in sleep (PLMS) and restless leg syndrome (RLS) can be successfully treated with dopaminergic agonists (Hening et al. 2004; Stiasny et al. 2002), and both motor disturbances are associated with ADHD (Chervin et al. 2002; Picchietti et al. 1999). Also, alterations in the nigrostriatal dopamine branch are shown to cause PLMS and RLS (Michaud et al. 2002). It appears, therefore, quite possible that the hypofunctioning mesolimbic, mesocortical, and nigrostriatal dopamine branches which Sagvolden et al. consider to be essential for ADHD symptoms, may be associated with the sleep disturbances in patients with ADHD by modulating cortico-subcortical interactions.

Changes in brain dopamine may affect the sleep-wake cycle also by modulating the balance between cortical inhibition and facilitation, which is recognized as important for the sleep-wake regulation (De Gennaro et al. 2004; Gottesmann 1999; Muzur et al. 2002). There are animal-driven data showing that dopamine has a double inhibitory influence at cortical level, either directly or by favoring gamma-amino butyric acid release from interneurons (Grobin & Deutch 1998; Pirot et al. 1992; Retaux et al. 1991; Zhou & Hablitz 1999). Hence, dopamine hypofunctioning may lead to insufficient cortical inhibition. Importantly, studies with transcranial magnetic stimulation have convincingly evidenced that ADHD children display a decreased intracortical inhibition (Buchmann et al. 2003; Moll et al. 2000a; 2001), and the dopaminergic drug methylphenidate significantly improves this deficit (Moll et al. 2000a). Taken together, these results imply that sleep disturbances in ADHD may be associated not only with modified cortico-subcortical interactions following dopamine deficit but also with a related alteration in cortical excitability.

In conclusion, the neurobiological mechanisms of ADHD psychopathology and sleep-wake regulation may have much in common. Therefore, it may be suggested that specific sleep-wake patterns may characterize ADHD phenotypes resulting from variations of an impaired cortico-subcortical interplay. This aspect may certainly be addressed in a model for ADHD. However, sleep problems in ADHD can hardly be explained by Sagvolden et al.’s dynamic developmental theory, since specific sleep-wake patterns are less likely to be determined only by reinforcement/ extinction mechanisms.

RED: ADHD under the “micro-scope” of the rat model

Katya Rubia
Department of Child Psychiatry, Institute of Psychiatry, London, SE5 8AF, United Kingdom. k.rubia@iop.kcl.ac.uk www.iop.kcl.ac.uk

Abstract: Derived from a rat model, the theory of Sagvolden et al. offers an all-explanatory model of attention-deficit/hyperactivity disorder (ADHD) anatomy, behaviour, and cognition as being caused predominantly by a hypo-dopaminergic mesolimbic (affecting the mesocortical and nigrostriatal) system, leading to abnormal reward and extinction processes. This model differs from overspecialization and reductionism, reflecting the limitations of the use of animal models to explain higher mental disorders.

The target theory is interesting and potentially useful to explain specific reward related aspects of human attention-deficit/hyperactivity disorder (ADHD) behaviour. Rodent models of human disorders can be of use if they try to explain lower level functions that are shared by humans and animals, such as specific motor or limbic dysfunctions. To explain higher complex cognitive dysfunctions via rat behaviour and anatomy is, in my opinion, problematic, considering the nearly inexistent frontal lobes and reduced number of complex neural networks in rodents. No wonder then, that no animal model is fully comparable to clinical ADHD (Davids et al. 2003). The theory of Sagvolden et al. suffers from precisely this attempt to stretch an animal explanatory model of a relatively limited aspect of dopamine-mediated reward behaviour to explain all possible features of the highly complex cognitive disorder that is ADHD. ADHD behaviour (including hyperactivity, impulsivity, and inattention), and ADHD cognition (such as inhibition deficits, delay aversion, and response variability) are partly being redefined, and reduced to dopamine mediated abnormalities in reinforcement and extinction processes. Likewise, brain abnormalities in ADHD are being reduced to three dysfunctional fronto-striatal dopamine pathways insufficiently fed by reduced dopamine in the ventral tegmental area. Uni-causal explanations of complex problems, as much as they appeal to the reductionistic human brain, have shown to rarely match reality. There is evidence to doubt that
the complex features of the psychiatric disorder ADHD can be reduced to a single defect in one neurotransmitter (dopamine) and one behavioural abnormality (reward and extinction). The authors seem to be aware of the inherent reductionism of their model, but they claim that ignoring other possible changes will facilitate the isolation and future research. I cannot see the advantage of isolating a still hypothetical (and controversial) aspect of ADHD pathology (i.e., hypofunctioning dopamine systems), by ignoring the interactions of this subcomponent with other components of the pathology (i.e., other neurotransmitters/chemicals and non fronto-striatal brain regions), and then—almost in contradiction to the initial admittance of reductionism—explain every single aspect of the pathology by this isolated sub-segmental dysfunction. If all ADHD behaviour features can be explained by hypodopaminergic fronto-striatal pathways, one wonders what the other neurotransmitters and brain regions are there for? The cautious statement that “the present model may be applicable mainly to a subgroup of ADHD linked to dopamine hypofunction” (sect. 3, para. 2) is a prime example of the logical fallacy of circular reasoning. We provide an explanatory model of ADHD, but it will only apply to those patients who meet the model.

The authors claim that “the majority of findings...seem to converge on dopamine in the etiology of ADHD” (sect. 3, para. 2) and that ADHD research has now转向ed dopamine investigation, and the few studies that have investigated the involvement of other neurotransmitters have been positive. Thus, atomoxetine, a selective noradrenaline inhibitor, has shown to be effective in ADHD symptom relief (Kratochvil et al. 2003), in line with the role of noradrenaline in attention processes and ADHD (Levy & Swanson 2001). Likewise, serotonin has been related to impulsiveness in animals and humans (Krakovski 2003; Robbins 2002) and in the mechanisms of action of stimulant drugs (Gainetdinov et al. 1999; Winstanley et al. 2003), but almost completely neglected in ADHD research (Oades 2002).

Furthermore, although a dopamine dysfunction in ADHD (alongside other neurotransmitter dysfunctions) is likely, at the present state of research it is unclear whether dopamine is hyper- or hypofunctioning (Solanto 1998), whether there is a differentiation of specific dopamine systems being under- (i.e., prefrontal) systems and others over-regulated (i.e., basal ganglia; Castellanos et al. 1997; Rohde et al. 2003), or whether there is a differentiation of hypo- and hyper-dopaminergic striatal systems in ADHD patients depending on symptom severity (Teicher et al. 2000). Other ADHD animal models have found a hyper-trophic rather than hypo-trophic mesocortical dopamine system (Vigiano et al. 2003b). The exhausted theory of hypofunctioning dopamine systems in ADHD has thus in recent years been replaced by a far more sophisticated picture of multiple and divergent monoamine dysfunction.

The authors do not make any attempts to integrate recent brain imaging findings of structural and functional abnormalities in temporal, parietal or cerebellar brain regions (Castellanos et al. 2002; Durston et al. 2003; Rubia et al., in press; Sowell et al. 2003) into their model of fronto-striatal dysfunction. The rather sweep- ing statement that global functional and structural changes in ADHD may be a result of reduced blood circulation caused by decreased dopamine, is difficult to sustain. I am not aware of a direct link between brain structure and blood circulation. In functional imaging, the effect of dopamine on neural hemodynamic coupling is controversial, apart from the fact that there is no specific effect of dopamine on blood flow over other neurotransmitters (Johnston et al. 2004). Thus, dopamine enhancement as well as reduction has not shown to have an effect on neural hemodynamic coupling (Esaki et al. 2002; Rao et al. 2000). If anything, there is recent evidence that dopamine antagonists increase fronto-striatal connectivity in healthy adults (Honev et al. 2003), whereas in ADHD, dopamine agonists decrease fronto-striatal and parietal activation (Langeleen et al. 2002; Schweitzer et al. 2003; Szobot et al. 2003), which would support a hyperdopaminergic hypothesis of ADHD.

The theory of abnormal reward and extinction processes as a global explanatory model for ADHD, as the authors acknowledge, predicts deficits in learning and memory in ADHD. Contrary to their claim, however, there is hardly any evidence in the ADHD literature for learning or memory deficits, unless, obviously, in comorbidity with learning disorder (or working memory, which is an executive function). The definition of the complex feature of impulsiveness as “responses with short inter-response times” and “the choice of smaller, immediate rewards” reflects the limitation of the rat’s model viewpoint: Although this is the only way impulsiveness can be measured in rats, in humans, impulsiveness is more complex, including heterogeneous features such as poor self-control, disinhibition, prematurity, temporal myopia, delay aversion, lack of persistence, increased boredom, sensation seeking, distractibility, inattention, and irritability (Evenden 1999; Rubia 2002). To explain all of these heterogeneous cognitive features by abnormal reinforcement processes seems an oversimplification.

The problem with animal models is that, in order to compare between species at the behaviour level, higher complex human features need to be reduced to motor and limbic components that can be observed in both, and at the anatomical level, complex human neural networks have to be decomposed into simpler motor and limbic pathways. This is exactly what Sagvolden et al. have done; their research has neural disorder of ADHD is being reduced to dopamine mediated limbic reward and extinction processes. Fortunately for us, the human brain is more sophisticated than the rat brain, and I am afraid more sophisticated theories will be needed to do justice to one of the most complex and pleomorphic disorders of psychiatry that is ADHD.

Is the hypodopaminergic hypothesis plausible as neural bases of ADHD?

Adolfo G. Sadile and Davide Vigiano

Department of Experimental Medicine, Second University, Naples, 80138 Italy. adolfo.sadile@unina2.it  davide.viggiano@unina2.it

Abstract: The “dynamic developmental theory” is based on hypofunctioning dopamine systems that follow an early overactivity phase. The theory does not consider recent experimental evidences from different attention-deficit/hyperactivity disorder (ADHD) models and the heterogeneity of the disorder. Alternatives are proposed that integrate available information gathered from clinical and experimental studies, with theoretical constructs.

The dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) postulates an early hyperfunctioning followed by hypofunctioning dopamine (DA) systems. Its peculiarity consists in an early overactivity phase of DA neurons that could be the result of different genetic and epigenetic factors. Historically, the use of psychostimulant drugs such as methylphenidate (MPH) and d-amphetamine in ADHD over decades has supported the hypofunctioning hypothesis. However, our understanding of the DA systems has increasingly improved as to feedback regulation in the mesencephalon and at target sites (frontal cortex, striatum). For instance, DA neurons control their own firing. In fact, DA D2 autoreceptors hyperpolarize DA neurons, and in turn reduce their responsiveness and firing rate (see, e.g., Bonci et al. 2003). Furthermore, the membrane transporter protein for DA (DAT) reduses DA neurotransmission by re-uptaking it into the terminal (Wightman et al. 1988).

MPH blocks DAT both in the mesencephalon and target sites, thus increasing synaptic DA (Seeman & Madras 2002). This, in turn, activates DA receptors. However, low doses of MPH mainly act on mesencephalic D2 autoreceptors, leading to inhibition of DA neuron firing (Brandón et al. 2003; Ruskin et al. 2001).

Therefore, the efficacy of low doses of MPH (Solanto 2002)
Commentary/Sagvolden et al.: A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD)

Seija Sandberg
Department of Mental Health Sciences, Royal Free and University College London Medical School, London W1W 7EY, United Kingdom.
s.sandberg@ucl.ac.uk

Abstract: Attention-deficit/hyperactivity disorder (ADHD) represents adaptation to defective neurotransmission – an adaptation seldom with benefit. The resulting behavioural style not only increases vulnerability to adverse experiences, but also creates a context in which encountering adversity is more likely. Furthermore, the fact that ADHD is a highly heritable condition increases the probability of a child with a compromised neurobiological disposition being raised by caregivers with suboptimal resources.

The target article is, to my knowledge, the first serious attempt to present a unified theory expanding from the biology of brain neurochemistry to continuously evolving interaction between a child with attention-deficit/hyperactivity disorder (ADHD) and his or her social environment. It is also noteworthy that the first author, Sagvolden, is a renowned scientist in murine research.

At the basis of the dynamic developmental theory, put forward by Sagvolden et al., is a model of dysfunctional dopamine systems in the brain. Three hypofunctioning dopamine system branches and their behavioural consequences, representing the core symptoms of ADHD, are outlined. These compromised properties result from a combination of intrinsic (genetic) and extrinsic (e.g., drugs and toxins) influences on the developing brain. The altered neurobiological disposition gives rise to two main behavioural processes causing ADHD: altered reinforcement of novel behaviour and deficient extinction of previously reinforced behaviour. According to the theory, ADHD symptoms are a product of a dynamic process of the individual's adaptation to defective neurotransmission.

The authors have construed a coherent account spanning from biochemistry, via behaviour, to a reciprocal interplay between the affected child and his/her biosocial environment. The theory predicts that ADHD behaviour results from, and is continuously modified by, the dynamic context of individual predispositions and interpersonal surroundings well into adulthood. And in the case of many adults, the individual predispositions come to form the interpersonal surroundings of another individual – their child.

The individual predispositions are primarily guided by genes. However, the interplay also starts early – going back (at least) to the intrauterine life (Grossman et al. 2003; Schneider et al. 1998). By the time the child's behaviour reaches the level of abnormality qualifying for ADHD, years of active interaction have taken place. And yet, as Sagvolden et al. note, not all children presenting with the core symptoms of ADHD get identified as maladjusted. This is because the environment has been unusually insightful and supportive in guiding the child's excessive and disorganised activity into constructive creativity. The individual ADHD symptoms at different times in a person's life vary and are influenced by factors exerting either a positive or negative effect. In other words, the environment can either protect from maladjustment, or predispose to it.

Crucial here is the caregiver's ability to adjust the environment to the child's needs for optimal development of adaptive skills. The resulting behavioural style, in turn, determines the long-term consequences of the early interactions. The theory predicts that a child with ADHD finds it hard learning how to match their behaviour to the demands of a given situation. Consequently, there will be few chances for the child to be rewarded for compliant behaviour. Instead, the resulting chaotic behavioural style will only magnify the negative interactions with carers. For optimal upbringing, the caregivers have to adapt to the child's special needs by taking into account the implications of the underlying deficits and adjust their expectations and demands accordingly. As the authors spell out, "a child with ADHD requires exceptional parenting skills" (sect. 4.5, para. 3).

The biopsychosocial context of ADHD

Seija Sandberg
Department of Mental Health Sciences, Royal Free and University College London Medical School, London W1W 7EY, United Kingdom.
s.sandberg@ucl.ac.uk

Although a dysfunction of DA systems is associated with ADHD in humans and animal models, this might be a compensatory change to other primary deficits (Rubia 2002). In fact, if the system is hyperactive, target neurons will be susceptible to neurotoxicity and neurodegeneration. Therefore, low doses of MPH reduce the phasic DA release, whereas high doses produce a “generalized stimulation” (Seeman & Madras 2002) a biphasic effect that is not predicted by a hypodopaminergic hypothesis. Nonetheless, the amelioration of the ADHD symptoms would be only symptomatic, because the primary defect has not yet been ascertained.

Notwithstanding, the molecular mechanisms by which MPH determines enduring changes in DA neurotransmission remain to be elucidated, as repeated MPH treatment in juvenile SHR exert long-term effects on membrane excitability (Brandon et al. 2003) and transduction mechanisms (Sadile 2000; Andersen et al. 2002).

In conclusion, the dynamic developmental theory appears plausible and interesting; however, it should include the above-mentioned considerations to explain different ADHD variants.
In real life, however, there are too few such resourceful parents to go around; and their availability to children suffering from the problems of ADHD is even more restricted. This is because about one in five of these parents themselves have ADHD, some with added complications of depression, personality disorder, learning disability, or substance abuse. Parents with such problems of their own will have even greater difficulty coping with their child's special needs (Lesesne et al. 2003). A child with ADHD growing up in these circumstances is at high risk for additional emotional and behaviour problems, with their likelihood further increased by low social class, parental psychopathology, and family conflict (Biederman et al. 2002b; Minde et al. 2003).

To elucidate the risk mechanisms involved, the authors juxtapose predictions from their theory with those of the coercion theory of antisocial behaviour disorder by Patterson (1982). According to Patterson, child non-compliance develops through a circular process of negative reinforcement between child and parent. Sagvolden et al. argue that such coercive child behaviour, once established, is especially hard to extinguish in children with ADHD (and in their often ADHD parents).

Because it is a highly familial disorder, ADHD also means that the same parents provide the genes and the environment. Parental ADHD, as a result of its core symptoms and/or concomitancies, is associated with disruptive family environment and suboptimal parenting practices that often are resistant to modification (Chornis et al. 2004; Sonuga-Barke et al. 2002). ADHD in fathers, for example, predicts higher levels of family disruption as a consequence of parental desertion and custodial sentences for impulsive behaviour (Minde et al. 2003). The already demanding tasks of childrearing place a parent with ADHD at considerable disadvantage: Maintaining patience and emotional responsiveness towards the child, providing attentive supervision, and organising domestic duties and childcare frequently present the parent with an unmanageable challenge. Also, extrapolating from the proposed theory, a parent with ADHD will find it hard to emotionally disengage amidst a child's temper tantrum, but will easily end up contributing to its escalation, instead.

These parenting styles bear resemblance to those observed in studies of depressed mothers. For example, a recent longitudinal study involving detailed observations of the interaction between postnatally depressed mothers and their infants revealed a striking pattern of “coercive caretaking” – a phenomenon hardly ever seen in mothers who were not depressed (Murray et al. 1996). This pattern of early interaction had long-lasting connections, predicting disruptive behaviour at least to age 8 (Morrell & Murray 2003). Thus, there is a particular reason to pay attention to ADHD in girls in whom the problems are often overlooked until teenage years, or entirely missed. Compared with boys with similar levels of ADHD, girls are at a higher risk for anxiety, depression, and poor psychosocial functioning (Rucklidge & Tannock 2001). If ignored, these problems are likely to continue into adulthood and will determine the future style of parenting – of children probably sharing the mother’s ADHD genes.

It seems fit to conclude by agreeing with Sagvolden et al. in that “ADHD . . . is a case where functions of the central nervous system occasionally exceed the limits of normal variation and adaptation” (sect. 3, para. 3) – and add environmental accommodation.

The target article by Sagvolden et al. is a veritable blockbuster linking comparative psychology with clinical developmental neuropsychopathology. The neuroscience of attention-deficit/hyperactivity disorder (ADHD) has only recently begun to emerge as a major contributor to our understanding of the aetiology and development of this disorder. The target article is both timely and informative and sets a research agenda for neuroscientists in the field of ADHD. The variety of issues that have been treated and to varying degrees integrated in the dynamic developmental theory (DDT) is exceptional.

Sagvolden and colleagues argue that in DDT there are two main behavioural processes causing ADHD: altered reinforcement of novel behaviour and deficient extinction of previously reinforced behaviour. Further, the authors argue that the time available for associating behaviour with its consequences is shorter in ADHD than in normal children, on account of the delay gradient being steeper and shorter in children with ADHD than in normal children.

First, I briefly address the relationship between state factors such as sleep and diurnal rhythm and the independence or interaction of both reinforcement and inhibition. Second, I argue that the DDT does not recognize in its current form how both state and inhibition contribute to explaining ADHD. Third, I draw attention to the fact that a comprehensive model of ADHD must account not only for correct responding but also the effect of detecting an error upon the following trial. Fourth, I refer to an omission in the DDT, namely, the role of motor factors in accounting for ADHD behaviour.

Convergence. Clinical neuropsychologists have been for some time interested in the relation between performance and reinforcement in ADHD children (Douglas & Parry 1994). Few areas of neuropsychopathology have been blessed with a richly researched animal model of the disorder of interest, and it is, therefore, timely that prior to the awaited DSM-V (Diagnostic and Statistical Manual, 5th edition), neuroscientists inform the clinical community of basic findings relevant to the definition of the disorder. The DDT provides an account of ADHD that requires careful evaluation. Hence, from a clinical neuropsychological point of view, one wishes to determine where is the convergence and where is the divergence between the DDT and, say, a cognitive-energetic model (CEM) explanation of ADHD.

Thankfully, there is convergence of evidence from the animal research reviewed in the DDT with the CEM. Both models note the variability of responding in ADHD, and both agree that reinforcement is a key factor in determining current and future behaviours in ADHD. They both appeal to a dopamine deficiency as the biochemical substrate of the disorder. The DDT and CEM instruct researchers to examine the clear association of the interval used to demonstrate deficiency: short intervals producing little or none; long intervals producing clear manifestations of deficiency compared with control children or animals. The DDT and CEM implicate widely distributed neural circuits being involved in ADHD, namely, the frontostriatal-limbic and cerebellar networks. The DDT is stronger than the CEM in its genetic predictions. For the purposes of this commentary, I briefly address four
points of divergence between the DDT and CEM: the role of state factors on ADHD, the independent role of reinforcement and inhibition, cognitive readjustment to protect future emissions of response, and motor behaviour.

State factors. The DDT contrasts with the CEM in that the former is a behavioural explanation of ADHD, and the CEM, as its name suggests, emphasizes both cognitive and energetic aspects of human behaviour. The CEM notes that behavioural overactivity of ADHD children occurs not only in the awake state but can also occur in sleep (Porrino et al. 1983). It is hard to know how an altered reinforcement mechanism could explain this finding, without having to appeal to additional biochemical mechanisms not specified in the DDT. Similarly, the DDT is unclear on how different effects which are related to behavioural activity occur when they do following midday (Porrino et al. 1983). What is the specific alteration in reinforcement that is linked to this diurnal effect? Furthermore, changes in brain state have been shown to predict the occurrence of succeeding errors (Brandeis et al. 2002).

Reinforcement and inhibition. Sagvolden et al. write “the response unit that is supposed to be inhibited is hard to define empirically (Catania 1998)” (sect. 1.2.3, para. 6). Inhibition, although a loose construct and operationalised in a variety of manners, can be measured by stop-signal reaction time (SSRT; Logan & Cowan 1984). It has been demonstrated to have high reliability (Bandy et al. 2003), associated specifically with the inferior frontal gyrus (Aaron et al. 2003), to be correlated with familial manifestations of ADHD and, in two meta-analyses, to distinguish ADHD from controls at a specific latency (Lijffijt et al., in press; Oosterlaan et al. 1998). Inhibition has been shown to be independent of reinforcement in predicting ADHD group membership (Solanto et al. 2001a). Several studies have shown that inhibition deficits in ADHD are independent of reinforcement (Oosterlaan & Sergeant 1998a; Scheres et al. 2003). One study showed an interaction between inhibition and reinforcement (Slusarek et al. 2001). These studies suggest, at the very least, that both inhibition (operationalised by the SSRT) and reinforcement are needed to explain ADHD.

Cognitive adjustment. When a human commits an error, cognitive resources are allocated to ensure that on the following trial, an error is not committed by slowing down the speed of responding (Rabbitt & Rodgers 1977). Normal children do this, but ADHD children fail to make this cognitive adjustment (Sergeant et al. 1999). This effect is independent of SSRT (Schachar et al. 2004) and can be improved by methylphenidate (Krusch et al. 1996). The DDT in its present form cannot account for this phenomenon, because it requires error detection, correction, and resource allocation – concepts not in the DDT.

Motor factors in ADHD. There has long been a clinical interest in motor functioning in ADHD (cf. Clements & Peters 1962) and even recently in differentiating ADHD children from children with a neurological disorder (Konrad et al. 2000). ADHD children can be differentiated from controls on repetitive movements (Carte et al. 1996), fine motor difficulty (Pitche et al. 2003), movement control (Eliaison et al. 2004), poor balancing (Rahberger & Wimmer 2003), and excessive overflow movements (Mottosky et al. 2003). Abnormal rhythm motor response in ADHD has been demonstrated using a tapping task (Ben-Pazi et al. 2003). ADHD children had difficulty modulating their responses with changing rhythms. Motor deficits need to be incorporated in the DDT.

Conclusion. The DDT model is an interesting contribution to the neuroscience of ADHD but requires expansion to accommodate the four areas noted here to be relevant for ADHD.

A common core dysfunction in attention-deficit/hyperactivity disorder: A scientific red herring?

Edmund J. S. Sonuga-Barke, and F. X. Castellanos

Abstract: The reinforcement/extinction disorder hypothesis (Sagvolden et al. 2004) is an important counterweight to the executive dysfunction model of attention-deficit/hyperactivity disorder (ADHD). However, like that model, it conceptualises ADHD as pathophysiologically homogeneous, resulting from a common core dysfunction. Recent studies reporting neuropsychological heterogeneity suggest that this common core dysfunction may be the scientific equivalent of a red herring.

The classical disease model of mental disorder rests on a number of assumptions (ideas taken for granted rather than tested empirically): Disorders are discrete entities, qualitatively different from normality and resulting from dysfunction (neurobiological, neuropsychological) at some level within the patient (Sonuga-Barke 1998). These assumptions have played a defining role in the contemporary neuroscience of mental disorder. They provide a meta-theoretical framework allowing shared points of reference that link science and clinical practice through common language, assumptions, and goals. They also constrain the types of questions that are deemed legitimate and the methods employed to answer them. In the neuroscience of ADHD, this has meant that one question above all has provided the ultimate challenge for researchers: Where, within the brain or mind of the ADHD child, is the site of the common core dysfunction that causes ADHD (Sonuga-Barke 1994)?

It is typical of “normal” science that one particular model garners the support of a large, cohesive, and influential group of supporters. This model then takes on the mantle of scientific orthodoxy. In the neuroscience of ADHD, this mantle has fallen on the executive dysfunction model (Arnsten 2001). This model proposes that ADHD is the result of a common core dysfunction in executive control associated with deficient inhibition-dependent processes such as working memory, planning, and interference control. These are underpinned by the prefrontal cortex and related neural circuits and neurotransmitter branches (especially mesocortical dopamine and norepinephrine pathways; cf. Roth & Saykin 2004). This “classical” executive dysfunction model, although initially based on an analogy between ADHD and the hyperactive and distractible behaviour of patients with prefrontal lesions, now receives support from (1) psychopharmacological studies highlighting the role played by catecholamines in the pathophysiology of ADHD (Bedard et al. 2004), and (2) neuroimaging studies demonstrating abnormalities within the frontal-striatal networks of children with ADHD (Castellanos 1997). Although few studies have tested its causal status, these data have been taken as compelling evidence for the executive dysfunction model of ADHD.

Challenges to this model take a number of different forms. First, there are those alternatives that call for its reinterpretation rather than its overthrow: The “field” has been looking in the right place (prefrontal cortex-executive function) for the right thing (a common core dysfunction) but needs to adjust the current model to take account of new data or ways of thinking. For example, the state dysregulation account proposed by Sergeant and colleagues elaborates the executive dysfunction model to account for the effects of factors such as reward, stimulus presentation speed, and stimulant drugs by incorporating the concept of cognitive energetic dysregulation (Sergeant et al. 1999). Second, some accounts propose a more radical departure from the dominant model. They argue that, while looking for the right sort of thing (a common core dysfunction), the field is looking in the wrong place. The model proposed by Terje Sagvolden and colleagues in the target article...
Abstract: The dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) proposes that hypodopaminergic functioning results in anomalous delay-of-reinforcement gradients in ADHD, which in turn might account for many of the observed behavioral and cognitive characteristics. However, hypodopaminergic functioning might also impair mnemonic representation of codes for spatial, motoric, and reward information and contribute to the purported shorter delay gradients in ADHD.

A variety of theoretical models of attention-deficit/hyperactivity disorder (ADHD) (or its component features) have emerged during the past 10 years, with a common aim of guiding basic and clinical research to improve diagnosis and treatment of this common but inadequately understood neurodevelopmental syndrome (reviewed by Barkley 1997b; Sergeant et al. 2003). To this end, Sagvolden et al. provide an ambitious and important synthesis of behavioral and neurobiological constructs, and they propose that hypofunctioning in one or more of the three dopamine branches alters an organism’s delay of reinforcement gradient (renders it shorter and steeper), which—depending on the immediate and past contexts—gives rise to various behavioral, cognitive, and emotive symptoms that characterize ADHD. Here, I suggest that hypodopaminergic function also influences visual-spatial selective attention and working memory, which might contribute to delay gradients. Also, I note that visual-spatial working memory correlates with behavioral symptoms of inattention, rather than hyperactivity/impulsivity—the focus of Sagvolden et al.’s model—and is enhanced by methylphenidate.

Sagvolden et al. focus on the impact of hypodopaminergic function on delay-of-reinforcement gradient and propose that delay gradients are shorter in ADHD. To understand why individuals with ADHD may (at times) have a shortened delay gradient, it is also helpful to consider how hypodopaminergic function impacts critical aspects of learning and working memory in animals and humans (e.g., Chusdasama & Robbins 2004). At a neuronal level, the psychological construct of working memory is manifest as sustained neural activity that bridges: (1) the delay between a sensory cue (a sound or flash of light) and the subsequent response to that cue (e.g., a saccadic eye movement or limb movement to the location of that sensory cue), and (2) the temporal interval between a response and a reward (e.g., Funahashi et al. 1989; Fuster & Alexander 1971; Schultz et al. 1997). In other words, this persistent neuronal activity is believed to be a mnemonic activity (i.e., mnemonic representation). Working memory allows animals and humans to temporarily hold and use information that is not currently present in the environment but is essential for adaptive behavior. However, the mechanisms of working memory overlap with those of attention to permit selective access to specific information amongst simultaneous and competing sources of information.

Emerging evidence from animal and human studies suggests that different populations of neurons (neural networks) exist in the striatum and cortex that maintain different mnemonic representational codes, such as sensory, motor, and reward (Curtis et al. 2004; Schnitzer-Torbert & Redish 2004). One reasonable speculation is that for a particular stimulus to function as a reinforcer of a specific response (or sequence of responses), the organism must aggregate mnemonic representations of at least the last motor response and its triggering sensory cue, together with that of the subsequent stimulus (the putative reinforcer). The speed of decay of these representational codes (i.e., weak mnemonic representation) might influence the delay-of-reinforcement gradient. Of note is the recent finding that pharmacological stimulation of the D1 receptor mechanism in rodents improves both selective visual attention and working memory in a dose-dependent and delay-dependent manner (Chusdasama & Robbins 2004). Also, some evidence indicates that the spontaneously hypertensive rat (SHR), which is often proposed as an animal model of ADHD, manifests deficiencies in visual-spatial attention, short-term memory, and spatial learning (De Bruin et al. 2003; Nakamura et al. 1996). Given the evidence that SHR show shorter delay gradients (as proposed for ADHD), it is plausible that impaired mnemonic codes for spatial, response, and reward cues might contribute to their anomalous delay gradients.

Several findings from the ADHD literature are of relevance to the preceding argument. First, individuals with ADHD have impairments in selective visual attention (Jonkman et al. 2001) as well as in visual-spatial working memory in ADHD (see meta-analyses by Hervey et al. 2004; Martinussen et al., in press). Second, visual-
spatial working memory correlates with behavioral symptoms of inattention, but not with symptoms of hyperactivity/impulsivity (Martinussen & Tannock, in press). Third, in molecular genetic studies, linkage has been found between specific D<sub>1</sub> haplotypes and behavioral symptoms of inattention but not hyperactivity/impulsivity (Misener et al. 2004). Fourth, methylphenidate (albeit not specific to D<sub>1</sub> mechanism) enhances visual-spatial working memory but not other aspects of visual processing (Bedard et al. 2004). Admittedly speculative, I suggest that hypodopaminergic functioning influences visual attention and working memory in animals and humans (including individuals with ADHD), which in turn contribute to delay-of-reinforcement gradients.

My final comment highlights potential problems with Sagvolden et al.’s attempt to model DSM-IV (American Psychiatric Association 1994) subtypes of ADHD (rather than its symptom dimensions). First, there is little or no evidence that these subtypes are developmentally stable or “breed true” in families (Todd et al. 2001). Rather, the putative subtypes may be source-specific syndromes, meaning that subtype-classification varies according to the informant (parent versus teacher), rater bias, and the algorithm used to combine information across informants (Gadow et al. 2004; Mitsu et al. 2002; Simonoff et al. 1998; Thapar et al. 1995). Second, in contrast to hyperactivity or impulsivity, which declines or changes in its manifestation with increasing age (e.g., change from overt motoric expression to covert inner restlessness; Downey et al. 1997), inattention persists into adulthood (e.g., Mick et al. 2004). Moreover, converging evidence from cross-sectional and longitudinal, clinical and epidemiological studies indicates that marked inattention (even at a subthreshold level for a diagnosis of ADHD) constitutes a developmental risk factor for neuropsychological and academic impairments (e.g., Chhabildas et al. 2001; Rabinder & Coie 2000; Warner-Rogers et al. 2000; McGee et al. 2002; Todd et al. 2004; Martinussen & Tannock, in press) and predicts a less robust response to psychologically based academic interventions (Rabinder et al. 2004). In particular, deficits in visual-spatial working memory are strongly associated with severity of behavioral symptoms of inattention but not with hyperactivity-impulsivity symptoms, in children with ADHD (Martinussen & Tannock, in press). Thus, although Sagvolden et al. seek to account for various behavioral characteristics of the hyperactive/impulsive and combined subtypes of ADHD via abnormal delay gradients, a more parsimonious approach might be to investigate symptom dimensions rather than categorical subtypes or ADHD itself, particularly the inattention dimension.

ACKNOWLEDGMENTS

This article is part of the project “Attention-Deficit/Hyperactivity Disorder (ADHD): From genes to therapy,” conducted at the Centre for Advanced Study (CAS) in Oslo in 2004/05. Preparation of this commentary was partly supported by funds from the Canadian Institutes of Health Research (MOP-64312).

Altered sensitivity to reward in children with ADHD: Dopamine timing is off

Jeffery R. Wicke<sup>a,c</sup> and E. Gail Tripp<sup>b,c</sup>

<sup>a</sup>Department of Anatomy and Structural Biology, University of Otago, P.O. Box 913, Dunedin, New Zealand; <sup>b</sup>Department of Psychology, University of Otago, P.O. Box 913, Dunedin, New Zealand; <sup>c</sup>Neuroscience Research Centre, University of Otago, P.O. Box 913, Dunedin, New Zealand.

jeff.wickens@stonebow.otago.ac.nz gtripp@psy.otago.ac.nz

Abstract: Despite general agreement that altered reward sensitivity is involved in attention-deficit/hyperactivity disorder (ADHD), a wide range of different alterations has been proposed. We cite work showing abnormal sensitivity to delay of reward, together with abnormal sensitivity to individual instances of reward. We argue that at the cellular level these behavioural characteristics might indicate that dopamine timing is off in children with ADHD.

We agree with Sagvolden and colleagues that ADHD involves altered sensitivity to reward. However, despite general agreement that altered reward sensitivity is involved in ADHD, a wide range of different alterations has been proposed. Wender (1971; 1972; 1974) argued that children with ADHD have a reduced sensitivity to reward and punishment; a view shared by some others (Haenlein & Caull 1987). Extending this idea, Barkley (1989b) suggested that children with ADHD do not simply show a reduced sensitivity to reward, but that reward loses its effects more quickly for them; that is, they satiate more rapidly to reinforcers and punishers than other children and they have a reduced sensitivity to partial or delayed reinforcement. In contrast, Douglas argued that children with ADHD are unusually sensitive, not less sensitive, to the effects of reward (Douglas 1989; Douglas & Parry 1994). She proposed that children with ADHD show an increased tendency to seek immediate reward, over-react to failures to obtain rewards, and are sometimes more vulnerable to the arousing and distracting effects of rewards. More recently, Barkley (1994; 1997a; 1997b) proposed that children with ADHD are neither more nor less sensitive to reward than other children, nor are they dominated by a tendency to seek immediate reward. Rather, the behaviour of children with ADHD is influenced more by “immediate events and their consequences than by those more distal in time” (Barkley 1997b, p. 77).

In order to clarify competing accounts of altered reward sensitivity in children with ADHD, we developed a signal-detection paradigm that provided direct measures of sensitivity to reward (Tripp & Alsop 1999). The sensitivity of boys with and without ADHD to differences in reward frequency was compared using a signal-detection task in which correct identification of one stimulus was rewarded three times more often than correct identification of the other. Both groups showed a consistent response bias toward the more frequently rewarded alternative, and there were group differences in the stability of this response bias. Importantly, the control children developed a stable response pattern that was governed by their overall history of reward on the task, whereas the ADHD group showed changeable patterns of response bias that were very sensitive to individual instances of reward. Our finding that children with ADHD were more sensitive to individual instances of reward, rather than the integrated history of reward, is consistent with proposals that children with ADHD are more vulnerable to the arousing and distracting effects of reward (Douglas 1989; Douglas & Parry 1994), and that their behaviour is influenced more by immediate events and their consequences (Barkley 1997b). This might also explain why the behaviour of children with ADHD is indistinguishable from that of controls when reinforcement is continuous, but differs under partial schedules of reinforcement. Under continuous reinforcement, each instance of the targeted behaviour is reinforced, so responding to individual instances of reward produces the same behavioural outcome as responding to reinforcement history.

We also agree with Sagvolden et al. that children with ADHD are particularly sensitive to reward delays. However, there is uncertainty concerning what feature of the delay produces this effect. Sonuga-Barke and colleagues (Sonuga-Barke 1994; Sonuga-Barke et al. 1992) noted that the greater sensitivity to reward delay shown by children with ADHD could have a number of explanations: (1) The children may act to increase the immediacy of individual rewards; (2) They may be generally delay averse – that is, sensitive to both pre-reward and post-reward delays – and act to reduce overall delay; or, (3) They may act to maximise the overall rate of reward by avoiding delays that decrease the overall number of rewards that can be obtained. Sonuga-Barke et al. (1992) tested these competing explanations in a series of experiments examining the sensitivity of hyperactive and control children to small immediate and larger delayed rewards. Taken together, the results of these experiments suggested that hyperactive children were neither impulsive nor reward maximisers, but rather are more generally delay averse.

The delay aversion hypothesis has been tested in two recent

BEHAVIORAL AND BRAIN SCIENCES (2005) 28:3 445
who show abnormal sensitivity to individual instances of reward. This might be important in that subset of children with ADHD. Thus, we would argue that hyperfunctioning of the dopamine system might be associated with a 3.5-second delay before reward was delivered. Children in the ADHD group showed a greater bias toward immediate reward than the controls. This result supports the hypothesis that children with ADHD are unusually sensitive to even brief pre-reward delays. It is not consistent with earlier research suggesting that they are more generally delay averse. If this were the case, ADHD children would show equal sensitivity to pre- and post-reward delays, and their bias score would be close to zero.

Hypothesising an abnormality of the dopamine system is a logical starting point to explain the above findings of abnormal sensitivity to reward in children with ADHD. We have investigated brain mechanisms for reward-related learning using intracellular electrophysiology approaches (Kerr & Wickens 2001; Reynolds & Wickens 2000). Using in vivo intracellular recording we have recently shown that behaviourally rewarding stimulation induces a dopamine-dependent potentiation of synaptic connections in the corticostriatal pathway (Reynolds et al. 2001). The degree of potentiation was correlated with the rate of learning a lever-pressing task. This dopamine-dependent potentiation of corticostriatal synapses is a potential mechanism for integrating reinforcement history, and a disorder of this mechanism might underlie the abnormal sensitivity to individual instances of reward seen in children with ADHD.

In our electrophysiological studies, the magnitude and direction of dopamine-dependent potentiation of corticostriatal synapses is sensitive to the precise timing of dopamine application (Wickens 2000; Wickens et al. 1996). At both the cellular and behavioural level (Black et al. 1985), delays of less than one second have marked effects. The brain’s mechanisms for coping with delay of reinforcement appear to rely on the brain’s ability to release dopamine in anticipation of later rewards, rather than traces at the cellular level. Thus, if the timing of dopamine signalling is off, altered sensitivity to rewards is likely, especially if these are delayed. We propose that dopamine timing is off in ADHD. A clue to how this might come about comes from the response of children with ADHD to individual instances of reward relative to the integrated history of reward. This suggests that at the cellular level there may be a greater initial magnitude of potentiation and more rapid decay. This might result from dopaminergic hyperinnervation, producing a higher-amplitude dopamine signal, with a shorter time-course because of the higher density of dopamine-transporters. Thus, we would argue that hyperfunctioning of the dopamine system might be important in that subset of children with ADHD who show abnormal sensitivity to individual instances of reward.

Attention-deficit/hyperactivity disorder (ADHD): One process or many?

A. Charles Catania
Department of Psychology, University of Maryland, Baltimore County (UMBC), Baltimore, MD 21250. catania@umbc.edu
http://www.umbc.edu/psyc/personal/catania/catanias.html

Abstract: Some commentators suggest that the attention-deficit/ hyperactivity disorder (ADHD) theory of this condition does not explain enough. Because the theory includes parameters of the delay gradient that vary across individuals and developmental modulation of behavioral outcomes by different environments, it accommodates a wide range of manifestations of ADHD symptoms. Thus, the argument could instead be made that the theory allows too many degrees of freedom. For many purposes, behavior is better defined in terms of function (e.g., consequences) than in terms of structure (e.g., muscle movements), so cognition is treated here as a variety of behavior rather than as a different category of phenomena. The commentaries are discussed in the context of these and other distinctions, including those between association and selection, between operand and respondent behavior, and between fundamental processes and those that are derivative. Other issues include: prosthetic environments, rapidity of developmental change, the concept of inhibition, the form of the delay gradient, and possible directions for experimental research.

I must begin by setting the record straight. At one point in his commentary, Karatekin refers to the theory presented by Sagvolden, Johansen, Aase, and Russell (Sagvolden et al.) as based in part on my work. If that is at all the case, it is only with regard to the fact that many years ago Sagvolden and I conducted collaborative research with pigeons on basic behavioral processes that involved delay of reinforcement (e.g., Catania et al. 1988). Sagvolden began his research with Wistar-Kyoto (WKY) rats and spontaneously hypertensive rats (SHR) independently of me, and it was he who first observed that some features of SHR behavior might be attributed to anomalies of the delay-of-reinforcement gradient, a view that I regarded with much skepticism for a considerable time. Over the years we have occasionally discussed these phenomena, in connection with extending the account to attention deficit and to individual differences, and in exploring the implications of more recent research on delay-of-reinforcement gradients that I had conducted in collaboration with my colleague, Elliot Shimoff (see Fig. 1 in my precommentary). The initiative for the ADHD enterprise treated here and the extension of its scope, however, originated with and must be fully credited to Sagvolden and his colleagues.

R1. Is one process enough?

I had expected that one criticism of the treatment would be that the account in terms of delay gradients, especially in combination with developmental contingencies, allows so many degrees of freedom that it explains too much. I had not anticipated that so many would argue that it explains too little: for example, Banaschewski, Himpel & Rothenberger (Banaschewski et al.), Carrasco, López & Aboitiz (Car-

Some of those concerns seem to arise from assumptions about whether the account should stand or fall based on its behavioral features or on correlated neurophysiological processes and the relations between them. When in my precommentary I said of the target article that “In the long run, the success of their account will depend on the adequacy with which fine details of dopamine systems are linked via grosser cellular and neuroanatomical levels to their eventual molar behavioral products,” I meant that it would be an important demonstration of how behavioral considerations could drive the search for processes at other levels. As argued in my precommentary, I maintain the assumption that behavior drives other processes and is the appropriate starting point. If certain localized physiological processes seem inconsistent with behavior, then consistent processes must be sought in other places. In response to Sonuga-Barke & Castellanos, locating a site for attention-deficit/hyperactivity disorder (ADHD) is not the “one question above all.”

Above all, we must understand how the behavior works, or we will not know what to look for in the brain (Catania 2000). Just as snow and sleet and rain have in common their origins in the properties of water molecules, behavioral phenomena that seem superficially to be vastly different may nonetheless emerge from underlying commonalities. Where Coghill and others see dichotomies, it may be useful to look instead for continua. This approach is illustrated in Figure R1, which is a hypothetical representation of the severity of attention-deficit and hyperactivity symptoms as a function of some continuous parameter that affects the steepness of the delay gradient and that varies within a population.

The figure shows the severity of attention deficit decreasing monotonically with shallower gradients (cf. Fig. 4 in my precommentary). Hyperactivity, however, is shown passing through a maximum, on the grounds that very steep delay gradients may not allow the early parts of response sequences to be captured and therefore may be less likely to engender hyperactivity than moderately steep ones. As a function of the decay parameter, therefore, attention deficit might dominate over hyperactivity (as at A), both might present themselves as significant (as at B), or hyperactivity might dominate over attention deficit (as at C). At some point, the decay parameter may approach the mean value for the population (perhaps in the neighborhood of D).

In other words, as a single continuous variable, the decay parameter can engender different patterns of symptoms with variations across the members of a population. Furthermore, if values of the decay parameter are linked to other variables, the decay parameter can handle differences in the presentation of symptoms in populations differentiated by gender or by other dimensions. This is why it would be of interest to develop methods to assess delay gradients directly in individual children.

Another degree of freedom can be added by incorporating developmental contingencies into the account. For example, suppose that attention deficit is to some extent mediated in some children by the diligent use of short delays to build conditional reinforcers capable of supporting extending sequences and higher-order skills (cf. Modicato & Presti and Sandberg), but that this intervention has little impact on hyperactivity. In this case, the balance between symptoms can vary even given the same underlying value of the decay parameter. These considerations may have implications for how other and perhaps derivative processes, such as executive function, might be determined by the interaction of the decay parameter with developmental contingencies.

Were impulsivity to be added to Figure R1, it might be assumed that it would be more likely to follow the form of the attention-deficit function than that of the hyperactivity function, because it does not depend on the maintenance of extended temporal sequences. With more reliable measures, as called for by Coghill, we can hope to refine our assessments of the correlations among these dimensions of ADHD symptoms. These are all behavioral questions, and they are orthogonal to questions of how the various aspects of ADHD behavior may be linked to physiological, genetic, or other dimensions. I am not arguing that the issue of number of processes has been settled. Rather, I am urging that for reasons of parsimony we should be sure we know how far one process can take us before we divert our efforts to the invocation of others.

Consider as an example Sonuga-Barke & Castellanos on delay aversion, which is a complex derivative of more fundamental processes. The present account in terms of delay gradients directly implies that consequences that are available only after longer delays will be less effective in maintaining behavior. An account in terms of delay aversion reaches that conclusion indirectly. In particular, it depends on avoidance behavior, which is hard to learn because nothing happens after successful avoidance responses. Those with truncated delay gradients will be exactly the ones most disadvantaged in acquiring such behavior, and, as similarly argued by Wickens & Tripp, prevention of the delay to some outcome is inevitably confounded with the more immediate reinforcement by that outcome when the delay has been avoided. With variable delays, as in most natural rather than experimental contexts, further problems of interpretation arise because delays are variable and therefore are undefined until the relevant consequence occurs.

Several commentaries have addressed extinction deficit as another proposed process to supplement the effects of delay gradients. Pending the availability of new data that may bear on its support, I have nothing more to add to the
R2. Rapid developmental changes

Banaschewski et al. imply that the dynamic developmental theory is inconsistent with rapid developmental changes when they state that according to the model “symptoms should develop more or less continuously.” But we need not assume that either the contingencies or the consequences that maintain various classes of behavior will remain invariant over time. For example, a child’s thumb-sucking may drop out fairly suddenly, not because the tactile consequences of thumb-sucking have changed, but rather because at this point in this child’s development those consequences have become less significant than the behavior that the thumb-sucking engenders in the child’s peers.

While acknowledging the potential of the theory, Kara-tekin states that the assumption is questionable that “the same mechanisms operate in the same manner from infancy through adulthood.” But changes in the significance of consequences can lead to dramatic changes in the form of behavior despite invariances in environmental contingencies.

Similarly, Levy cites “the sometimes immediate and dramatic changes in oppositional behavior of ADHD children on stimulant medication” as a problem for the theory, but given the nature of reinforcers, any medication that alters their effectiveness may well shift the relative likelihoods of different classes of behavior. This is especially worth considering because reinforcers are not absolute. They must be assessed in terms of the relative probabilities of the response to be reinforced and the responses that are occasioned by the delivery of the reinforcer. It has long been known that an opportunity to eat can reinforce drinking, or an opportunity to drink can reinforce eating, depending on which is momentarily the more probable response (Premack 1962). Consider when a parent arranges contingencies in which an opportunity to play with friends reinforces the child’s completion of a meal. Given developmental changes in the potency of different reinforcers and in the acquisition of new reinforcers, we should no more expect that varied environments will produce similar kinds of behavior within the ADHD subpopulation than we would within the population of children as a whole.

R3. Structure and function

In the analysis of behavior, it is crucial to distinguish between structure and function (Catania 1973; 1998). The former is analogous to anatomy in biology, and the latter is analogous to physiology. Diagnostic criteria and research based on medical models, as discussed by Rubia and by Banaschewski et al., often appeal to structure rather than function. For example, the head-banging of two children may be defined as self-injurious behavior even though it is maintained by contingencies such as escape from demands in one case and by unidentified endogenous sources in the other, whereas the bullying of one child and the shouting of obscurities by another may both be similarly maintained by contingent attention. The former case involves similar forms of response with different functions, and the latter involves different forms of response with similar functions.

With regard to treatment in these cases, it is more important to be concerned with function than to be concerned with form.

The issue matters because the present account brings a single functional process to bear on behavior that manifests itself in a variety of forms. Moderato & Presti’s discussion of prosthetic environments is especially welcome in this context, not only because it suggests the feasibility of behavioral interventions for ADHD but also because it discusses the advantages of functional assessment.

The point may also be relevant to whether ADHD is always disadvantageous. Carrasco et al. think it may not have been disadvantageous in phylogeny, and Sandberg thinks that, perhaps in interaction with different developmental pathways, it may not even be disadvantageous in all contemporary environments. The point is, again, that it matters more how the behavior functions than what it looks like, and variations should be no surprise. Evolutionary contingencies have selected learners, and that selection has made learning different in different situations. For example, sources of food and water vary, but prey organisms do not get repeated opportunities to learn to escape from predators. Thus, behavior shaped by consumable reinforcers is likely to be much more labile than behavior that involves escape or other interactions with potent aversive stimuli.

R4. Cognition as a variety of behavior

A recurrent theme, more explicit in some commentaries than in others, is the pitting of cognition against behavior (e.g., Banaschewski et al., Oades & Christiansen, Rubia, Sergeant, Tannock). That contrast establishes a false dichotomy. I argued against it in my precommentary, and I add to that argument here. For example, the essence of research on sensory development is that it depends on interactions with the environment or, in other words, on contingencies. The crucial difference between the active and passive kittens in the classic experiment by Held and Hein on visually guided behavior was that the contingent relations between what the active kitten did and what it saw did not exist for the passive kitten.

As another example, shifts in visual attention need not involve eye movement (Sperling 1960; Sperling & Reeves 1980). The criterion for whether something we do counts as behavior is not form (e.g., whether muscles are involved); rather, it is function or, in other words, sensitivity to environmental contingencies. The favoring of stimuli in different parts of the visual field may be spoken of in terms of automatic attentional processing, but whether one learns to read text in a language written from left to right or from right to left or from top to bottom determines an extensive history of contingencies that will be relevant to such processing. Thus, when impairment of processing and redistribution of attentional resources are suggested as alternative explanations of attention deficit, what is on offer is not a different sort of process but a different way of describing the process.

Similar arguments can be provided for more complex cases such as executive function. They are higher-order cases, in the sense that to engage in such behavior individuals must be sensitive not only to features of environments
but also to subtle properties of their own behavior with respect to those environments. It would be a mistake to assume that such capacities can develop independently of interactions with relevant contingencies (cf. Catania 1995). The treatment by Wixted and Gaitan of cognitive theories as surrogates for histories of reinforcement is also directly relevant here.

It is curious that some who dichotomize behavior and cognition suggest that behavioral accounts imply a passive organism (cf. Karatekin). If the organism whose behavior acts on its world by producing consequences is not active, then I cannot imagine what other capacity could make it so. If assumptions about passive organisms are anywhere, they are in treatments such as that of Björne & Balkenius, where so much is internalized that little if any behavior that acts on the world is evident. The model Björne & Balkenius cite is about “how context is formed and maintained,” but it doesn’t tell us where or by whom. Even the voice is typically passive, as when a child might “acquire a response that is unwanted” (by whom, one wonders). Contexts are features of environments, and saying they are internalized does not explain how they work.

**R5. Association, selection, inhibition, and the operant-respondent distinction**

Some commentators, and in particular Björne & Balkenius, are predicated on the assumption that the present account is based on an associative theory of learning rather than a selectionist one. A selectionist account holds that behavior is selected by its consequences within the lifetime of the individual organism, much as organisms are selected over generations by evolutionary contingencies (Catania 1978; 1987; Skinner 1953; 1975; and see Skinner 1981; 1984a). Though associationist theories are typically called S-R, or stimulus-response theories, the associations to which such theories appeal are located neither in behavior nor in the environment. It may be worth noting that associationism is not equivalent to connectionism, because associationism assumes behavioral units such as stimuli or responses whereas the units involved in connectionism operate at a very different level (cf. Donahoe & Palmer 1994). Furthermore, selectionism accommodates associationism because associations are as amenable to selection as other sorts of behavioral units, whereas an accommodation does not easily work in the other direction. After all, associationist accounts require a mechanism by which some associations are selected over others, but selectionist accounts need not be restricted merely to associations.

A further problem is the extension to operant or instrumental phenomena of concepts that have been derived historically from respondent or Pavlovian ones. For example, respondent procedures provide much of the precedent for stating that “extinction is not the unlearning of previously learnt associations between a response and a reinforcer . . . [but] is the learning of a new association that masks previously reinforced behavior” (Björne & Balkenius). Operant extinction, however, need not involve active suppression. It is mainly an outcome of the fact that the effects of reinforcement are temporary rather than permanent, and the extension of the Pavlovian inhibitory language was based on phenomena such as spontaneous recovery that have alternative interpretations (Catania 1998, pp. 71–77).

Accounts implicitly derived from respondent vocabularies often invoke the language of inhibition (e.g., Nigg), but that language can be ambiguous, especially in the specification of what is inhibited and what does the inhibiting. In such cases, it can be difficult to distinguish active suppression of behavior from failures to maintain it. For example, when Björne & Balkenius appeal to “enhanced contextual understanding and maintenance, thus enabling the child to inhibit responses to irrelevant stimuli,” we must be assured that the data cannot be accounted for instead in terms of enhanced responses to relevant stimuli.

Figure R2 illustrates the problem by comparing two different interpretations of a single hypothetical data set. At the start of both graphs of Figure R2, each of two responses (R1 and R2) has been maintained by reinforcers. In both graphs, one of them (R2) is subjected first to extinction and then to the reinstatement of reinforcement. In A, the starting performance at X is taken as the baseline, and the decrease of R2 in extinction is taken as an inhibitory process that is accompanied by an excitatory effect on R1. When reinforcement is reinstated for R2, both responses return to their baseline levels at X. In B, the extinction of R2 is taken as the removal of an effect of R2 reinforcers on R1 responding, allowing R1 to return to its baseline level at Y. When reinforcement is reinstated for R2, it again reduces R1 responding. If the language of inhibition is applied to A, it is not clear what is doing the inhibiting; if it is applied to B, the interaction involves specified and observable units: The reinforcement of R2 has inhibited R1.

The phenomenon illustrated in Figure R2 occurs both with concurrent schedules and with successive conditions arranged in multiple schedules (e.g., Catania 1969; Catania & Gill 1964), when it is sometimes called behavioral contrast. Based on various lines of evidence, it is more appropriately treated in terms of interactions among reinforced responses (as in graph B of Fig. R2) than in terms of side-effects of inhibitory processes during extinction. The issues have much in common with those raised earlier in regard to the status of delay aversion (see sect. R1) and with arguments against the concept of inhibition in cognitive contexts that have been offered by MacLeod et al. (2003). They may

![Figure R2](Catania). Two interpretations of changes in responding. In both graphs, reinforcement (rft) is maintained for Response 1 (R1), while Response 2 (R2) is subjected first to extinction (ext) and then to the reinstatement of reinforcement. Depending on whether the baseline is regarded as both responses reinforced together (at X) or as R1 reinforced alone (at Y), the change in responding can be interpreted as an excitatory side-effect of extinction, as in A, or as an inhibitory effect of the reinforcement of another response, as in B.
also raise questions about arguments in support of extinction deficit in ADHD that appeal to failures of an inhibitory process.

As Donahoe & Burgos have demonstrated, both operant and respondent phenomena are consistent with a selectionist account. In this context it is important to note that the distinction between associationist and selectionist theory is orthogonal to the distinction between operant processes, in which responses produce consequences, and respondent processes, in which conditional and unconditional responses (CRs and URs) do not affect presentations of conditional and unconditional stimuli (CSs and USs). The model presented by Donahoe & Burgos is significant because it suggests that delay contingencies have general enough relevance that they can operate within both operant and respondent procedures. Furthermore, research by Donahoe and Vegas has experimentally separated the role of the CS-US relation from that of the CS-UR relation, and has demonstrated that the latter, not the former, is critical for conditioning. As when the relative probabilities of the reinforced response and the responses occasioned by the reinforcer are crucial to the effectiveness of a reinforcer, as noted in section R2, here again, the significance of stimuli can be assessed only in terms of their relations to responses. The extension of the Donahoe & Burgos model to ADHD is therefore noteworthy.

R7. Theory and data: More calls to the laboratory

Several commentators ask for more experimental data (e.g., Coghill, Moderato & Presti). In response to questions such as Karatekin’s about the appropriateness of research on rats for the human problem of ADHD, we could appeal to the many ways in which nonhuman preparations have shed light on human disease (and see Catania 1983). In fact, I would argue that, except where verbal behavior is involved, a vocabulary drawn from behavioral phenomena observed over a range of species may have advantages over one drawn solely from human research. But, in the case of the dynamic developmental theory, such questions are probably better addressed more directly by parallel experiments with SHR and WKY rats and with ADHD and non-ADHD children.

An obvious start would be the direct assessment of delay gradients within response sequences and in attentional contexts, as illustrated by the pigeon data in Figure 1 of my precommentary. The procedures for the rat experiments are straightforward enough and would mainly require pilot efforts to identify relevant stimuli and experimental parameters. Detailed quantitative data would also allow comparison with the output of a mathematical model that assumes that responses weighted according to a decay function by the delays that separate them from a reinforcer contribute to a reserve of potential behavior, and that subsequent responding depends on the magnitude of that reserve, which is then depleted by responding (Catania 2005). The fits of the model to actual SHR and WKY data with different decay functions and parameters should help to narrow down the range of possible functions and, therefore, to resolve questions raised by Killeen.

Parallel work with children could use computer games that extend response sequences or that assess attention over increasing times between signals and events to be observed and acted upon. Data from such games might then provide the foundations for programs that could be arranged on laptop computers and used as a relatively simple and inexpensive component of a battery of diagnostic procedures.

The comments by Rothenberger & Kirov and by Sergeant on ADHD and sleep, taken together with literature on the behavioral effects of sleep deprivation in rats (e.g., Kennedy 2002; Kennedy et al. 2000), suggest that it may be informative to compare SHR rats with sleep-deprived WKY rats. The sleep phenomena correlated with ADHD cited in these commentaries should command special attention because their relation to delay gradients and other possible sources of ADHD seems at best remote.

Sergeant discusses relations between delays of reinforcement and errors. When sequences of errors are followed by an eventual correct response, the errors, weighted by delay, may be maintained to some extent by the reinforcers produced by subsequent correct responses. The interactions of reinforcers with correct responses and errors presumably will be influenced by the sequencing of trials, the use of correction procedures, and other procedural variables. Like the questions raised by Wickens & Tripp about experiments designed to assess sensitivity to reward and punishment, these considerations are reminders of the powerful effects that small differences in the scheduling of contingencies can have on ongoing operant behavior (Ferster & Skinner 1957). Wickens & Tripp’s use of signal-detection criteria to tease apart the sources of such differences seems a major step in an appropriate direction.

Killeen provides interesting experimental suggestions and some tantalizing data, thereby illustrating the broad reach of his model (Killeen 1994), the arousal and reinforcement components of which provide an alternative to Nigg’s treatment of arousal and reinforcement in sustained attention. Killeen’s suggestion about the use of bridging stimuli to mark the response has both theoretical and practical implications and promises to add to the arsenal of interventions suggested by Moderato & Presti. In contrast to Björne & Balkenius, who speak of “highlighting of the stimuli,” Killeen’s emphasis on the response rather than on the stimulus is particularly cogent, but the literature on the effects of such response feedback on maintained operant behavior is sparse.

Killeen suggests that the delay gradients proposed here should be reflected from left to right, so that they are converted to memory functions in which the effect of each response decays over time. With regard to that directionality, his account shares some properties with Tannock’s impaired “mnemonic representation of codes,” except that Tannock makes different assumptions about the relations between neural processes and working memory and implies a different direction of effect, seeing selective attention contributing to delay gradients rather than the other way around. Killeen also suggests a different form of the gradient, but if the same decay functions are used, the height of the memory curve where it intercepts the y-axis is the moment of reinforcement in his model is equal to the height of the delay function at a given time before the delivery of the reinforcer in the present one (Catania 2005). Thus, when they use the same decay function, the two models are formally equivalent. It will be of interest to see if the two formulations have other implications that will allow them to be differentiated experimentally.
Authors’ Response

The dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD): Present status and future perspectives

Espen Borgå Johansen, a,b Terje Sagvolden,a,b Heidi Aase,a,c and Vivienne Ann Russell[1,4]

a Centre for Advanced Study, Norwegian Academy of Science and Letters, NO-0271 Oslo, Norway; b Department of Physiology, University of Oslo, NO-0317 Oslo, Norway; c Norwegian Centre for the Studies of Behavioural Problems and Innovative Practice, UNIRAND LTD, University of Oslo, NO-0118 Oslo, Norway; d Department of Human Biology, University of Cape Town, ZA-7925 South Africa. e.b.johansen@medisin.uio.no terje.sagvolden@medisin.uio.no heidiaa@atferd.unirand.no russell@curie.uct.ac.za http://folk.uio.no/terjesa/

Abstract: The dynamic developmental theory (DDT) has benefited from the insights of the commentators, particularly in terms of the implications for the proposed deepened delay gradients in attention-deficit/hyperactivity disorder (ADHD). The introduction of modified memory processes as a basis for the delay gradients improved the links to aspects of ADHD. However, it remains unclear whether the hyperactive-impulsive and inattentive subtypes are separate subgroups or may be explained as different outcomes of the same genetic factors and thus explicable by the same principles. The DDT suggests that altered reinforcement and extinction processes define an endophenotype in ADHD that can be related dimensionally to inattention, hyperactivity, and impulsivity. The relation between the suggested endophenotype, characterized by changes in basic learning mechanisms, and other endophenotypes characterized by delay aversion or response disinhibition, needs to be tested in future studies.

R1. Basic assumptions

We first wish to thank the commentators for their valuable insights regarding attention-deficit/hyperactivity disorder (ADHD) and the present dynamic developmental theory (DDT). Any theory of ADHD will be faced with the multitude of research findings pertaining to this group of children. Given the diversity of findings, it is unlikely that they can all be explained by any new theory. However, consider the heterogeneity of ADHD findings as well as possibly etiology and underlying deficits: A good, strictly defined, and operationalized theory of ADHD should probably not try to be all-encompassing, because this would necessarily be at the expense of clarity and clear predictions. Concomitantly, in accordance with principles of parsimony, we argue that a theory should be fully explored for its explanatory power before additional explanatory processes are invoked; such exploration should not be mistaken for an attempt to explain all symptoms in ADHD.

An overarching theory such as DDT is needed to integrate the abundance of data from behavioral, genetic, neurobiological, and clinical studies of children with ADHD. It is based on general principles that are valid for human as well as animal behavior. The DDT is built on an extensive research literature (including our own studies) suggesting that reinforcement processes are altered in ADHD (Douglas 1983; Sagvolden & Archer 1989; Sagvolden & Sergeant 1998; Sagvolden et al. 1998; Sonuga-Barke 2002; 2003; Tripp & Alsop 1999; 2001; Wender 1971), as well as on studies of the animal model – the spontaneously hypertensive rat (SHR) (Sagvolden 2000; Sagvolden et al. 2005).

ADHD is a behavioral disorder with no established biological marker. We wanted to explore the extent to which ADHD symptomatology could be described and explained starting with a limited set of assumptions. Consequently, DDT is a theory that uses a few stringent behavioral principles to explain behavioral changes in ADHD predominantly hyperactive/impulsive and combined subtypes. It describes how these principles predict the way in which relatively small differences early in development might prompt behavioral interactions leading to substantial effects later on. Thus, we suggest that altered reinforcement and extinction processes combine to form the endophenotype (measurable components between cause and disorder) of ADHD. The DDT shows how alterations in these processes can account for symptoms observed in ADHD. This thinking is followed up by Catania in his precommentary. Further, through the relatively well-established link between reinforcement and dopamine function, the DDT explores how behavioral changes may be linked to a putative dopamine dysfunction in ADHD. The DDT assumes overall dopamine hypofunction in ADHD on a systems level, i.e., the net effect of the various dopamine neuronal activities on the nondopaminergic receiving cells is less than normal. In such an overall hypofunctioning system, it is likely that some aspects of the dopamine system would be upregulated because of compensatory mechanisms. Thus, it is important to emphasize that the DDT is not primarily a theory of dopamine function in ADHD. Our neurobiological understanding of the disorder will undoubtedly change as more knowledge is acquired. The DDT does acknowledge that other systems are involved, arguing that a change in one system also will affect the other transmitter and modulator systems. Other neurotransmitters or modulators are not covered in the DDT, to limit the number of assumptions. We challenge the commentators who regard the theory as too simple to expand it to deal with the acknowledged complexity of ADHD (e.g., Banaschewski, Himpel & Rothenberger [Banaschewski et al.], Carrasco, López & Aboitiz [Carrasco et al.], Coghill, Karratekin, Levy, Rubia, and Sadile & Viggiano).

R2. The DDT, a dimensional approach to ADHD

Tannock questions our attempt to model DSM-IV subtypes of ADHD and suggests that a more parsimonious approach might be to investigate symptom-dimensions rather than categorical subtypes of ADHD itself; particularly the inattentiveness dimension. We regard the DDT as a dimensional approach to ADHD that explains seemingly unrelated behavioral symptoms within the same theoretical framework. The target article explains how levels of hyperactivity and impulsivity may depend on changes in basic learning mechanisms in ADHD and how the behavioral symptoms will be dimensionally related to the shapes of individual delay gradients in ADHD. Catania elegantly shows in his precommentary how differences in delay gradients dimensionally can be related to hyperactivity, inattention, and impulsivity.

However, our suggestion that the DDT mainly applies to the hyperactive-impulsive and combined subtypes may be
unnecessarily restrictive. Hence, it is possible that the suggested endophenotype related to basic learning processes transcends the current subtypes of ADHD and includes both hyperactive and impulsive children as well as predominantly inattentive children. This is consistent with the comments on a common genetic component in the hyperactive-impulsive and the inattentive subtypes (Carrasco et al.) and the lack of ADHD “breed-trueness,” suggesting that the subtypes may change within families (Tannock).

A debate has recently started in this field on the possibility of several endophenotypes underlying the disorder. Several commentators mention that there might be multiple pathways to ADHD (Banaschewski et al., Coghill, Sergeant, Sonuga-Barke & Castellanos) consistent with the proposition that inhibition and delay aversion are two factors independently associated with ADHD (Solanto et al. 2001). In our view, what appears to be “disinhibition” is partly rooted in deficient extinction of previously acquired behavior and partly caused by impaired modulation of motor functions in terms of poor timing of starting and stopping responses, deficient acquisition, retrieval, and relearning of programs for sequential motor tasks. In this way, the vague and difficult-to-operationalize concept of inhibition is parted into definable and testable behavioral entities. In a recent review, MacLeod (MacLeod et al. 2003) provides a thorough discussion of the inhibition concept, supporting our standpoint that the term inhibition is poorly defined and difficult to test experimentally. Also, we view delay aversion as secondary to more basic behavioral processes associated with a shorter delay-of-reinforcement gradient (see target article, sect. 4.1). When the delay gradient is short and steep, even short reinforcer delays may be too long for establishment of stimulus control (target article Fig. 10, sect. 3.7). We suggest that it is aversive not to master or understand a situation partly because choices may be perceived to be forced, not free (Catania & Sagvolden 1980). Further, lack of stimulus control may lead to non-compliant and inadequate behavior met with negative consequences by parents and custodians. Thus, instead of viewing disinhibition and delay aversion as two independent factors in ADHD, the DDT adopts a dimensional approach explaining both factors as related to individual differences in basic learning processes.

R3. Delay-of-reinforcement gradients and memory

There is now general agreement that ADHD involves altered sensitivity to reinforcers (reward). Donahoe & Burgos support the DDT by means of computational modeling. Catania’s elaboration of the various implications of delay gradients is very welcome. It provides insight into many underlying processes and suggests that the DDT has even more explanatory power than laid out in the target article.

We do not agree with Coghill’s claim that there is a lack of reliable measures of reinforcement gradients and validation of a shortened delay gradient as an endophenotype. Catania points out that delay gradients are not theoretical constructs, but they may be measured experimentally. Reinforcers affect the behavior preceding them. Therefore, effects of reinforcers are plotted with the y-axis on the right side of the graph. However, this may give the impression that reinforcers work backward in time. This is obviously not the case. We agree with Killeen and Tannock that the delay gradient is a result of pairing the reinforcer with the fading of precursors, for example, the fading of memory traces of the behavior.

It is not so much that a delayed reinforcer weakens over time, as that the memory of the initiating response weakens, giving reinforcement less signal on which to operate among the buzz of other traces (Killeen, page 432).

Mnemonic representations will of course to some extent depend on attention. Karatekin asks “why not, for example, propose that deficits in attention or habit learning can explain response to reinforcers?” There does not appear to be a substantial discrepancy between this position and what we suggest in the target article: “Synaptic transmission undergoes plastic changes when presynaptic (glutamatergic) input, postsynaptic activation, and the dopamine signal occur simultaneously at the same neuron. Thus, the homogeneous dopamine signal associated with reinforcement will selectively reinforce the weights of synapses that are active around the time of behavioral reinforcement” (target article, sect. 2.5.1).

Looking on delay gradients as expressions of underlying memory processes introduces novel perspectives into the DDT. Altering short-term or working memory of events may change effects of delayed reinforcers by increasing memorability of these events when the reinforcer is delivered (see Killeen, Tannock). Memory traces will decay over time after the event and give rise to delay-of-reinforcement gradients like the ones depicted in the target article and Catania’s precommentary.

R4. Units (elements) of behavior and impulsivity

Delay gradients may be expected to vary as a function of a variety of parameters affecting memory processes. Catania points out that delay gradients may be influenced by such factors as whether response sequences are homogeneous or heterogeneous and whether the responses that make up those sequences are relatively simple units or are instead integrated higher-order units, perhaps temporally extended. The DDT predicts that, at least early in training, there will be fewer units in a chain of behavior in ADHD than normally (see target article Fig. 5, sect. 2.5.1). The slower acquisition of long behavioral chains will give rise to variable, less orderly behavior that might appear impulsive. This might also provide some of the reason why ADHD is associated with language problems which typically involve sentences with many behavioral elements (Lashley 1951). Thus, dopamine hypofunction will give rise to fewer units in a chain of behavior and will have consequences for how the child with ADHD understands and talks about his or her environment.

Impulsivity has many forms and is sometimes described in terms of executive dysfunction, disinhibition, or failure to withhold behavior, and it is typically regarded as the inverse of self-control. One variety of impulsivity may be a consequence of having too few behavioral elements in a chain of behavior, that is, the child may leave a task unfinished because the complete response sequence necessary for completing the task is not part of his or her behavioral repertoire. Another aspect of impulsivity and self-control is illustrated in Figure 3 in Catania’s precommentary in terms of hypothetical delay gradients (Ainslie 1975). Steep-
ened gradients create a greater preference for a more immediate, smaller reinforcer over a later, larger reinforcer. Such *discount functions* are used extensively in economics as well as in substance abuse research. We agree with Killeen that the choice of larger, delayed reinforcers occurs either because the organism has a history of such a delay in the present context or has been promised a delayed reward and infers its immediate value from personal histories of such delays. Further, optimal choice behavior will be hampered if memory processes underlying delay-gradients are steepened in ADHD. Thus, steep gradients can account for premature responding in several ways: (1) short IRTs (inter-response times), (2) more variable and disorganized behavior, as well as (3) preference for immediate reinforcers.

R5. Inattention and stimulus control

The explanation of inattention (poorer stimulus control), in terms of a steepened delay gradient (faster decay of the memory trace left by the onset of the stimulus), is an important aspect of the DDT. Poorer stimulus control is an important characteristic of ADHD behavior. This is pointed out by Björne & Balkenius ("weakened ability to code and maintain a context") and by Carrasco et al., suggesting that there is a different distribution of attention resources in ADHD, with a wider spatial attention framework and with a narrower time constant. The DDT, including Catania’s precommentary, discusses how stimulus control (context) is formed and maintained, as well as how goal directed behavior is guided.

Catania points out in his precommentary an important fact that is easily overlooked:

One simple but exceedingly important response that is maintained by . . . a stimulus is that of attending to it. A stimulus in the presence of which an opportunity for reinforcement is likely to arise very soon is more likely to be observed or looked at or attended to than one in the presence of which that opportunity is still some time away.

And further:

A substantial body of evidence demonstrates that organisms work to observe discriminative stimuli correlated with the delivery of reinforcers; they do not work to observe discriminative stimuli that are equally informative but are instead correlated with extinction or aversive events.

The dynamic developmental theory points out that the reinforcing value of the conditional stimulus decays by time from the onset of this stimulus to the delivery of the reinforcer (target article Figs. 5, 10, 11; sects. 2.5.1, 3.7, 4.2: Catania’s precommentary).

Although hyperactivity and impulsivity are explained in a straightforward way in terms of altered delay gradients, within the field of attention the arguments of the DDT apply mainly to sustained attention (target article and Catania’s precommentary and response). As Banaschewski et al. and Tannock point out, there may be other attention problems associated with ADHD that cannot be analyzed in terms of altered delay gradients. Tannock points out that hypodopaminergic function also influences visual-spatial selective attention, and it may involve color perception (Tannock et al. 2000). Thus, hypodopaminergia may give rise to poorer stimulus control of ADHD behavior as well as perceptual anomalies, each contributing to the sustained attention problems.

R6. Motor problems and timing in ADHD

The DDT suggests that a dysfunctioning nigrostriatal dopamine branch will cause several extrapyramidal symptoms (so-called neurological soft signs) associated with ADHD in the form of clumsiness, poor motor control, longer and more variable reaction times, poor response timing, poor handwriting, and poor correlation of the activity of different body parts (target article Fig. 1, sects. 1.1.1, 2.4, 3.2, 3.6, and 4.1). Research has also shown that dopamine and the basal ganglia are involved in timing in patients with Parkinson’s disease or schizophrenia (Packard & Knowlton 2002; Yang et al. 2004). Hence, dopamine dysfunction may not only affect motor control but also timing, as pointed out by Banaschewski et al. and Coghill.

Nigg claims that the DDT only addresses indirectly the phenomenon of slow, inaccurate responding in fast response contexts in ADHD. Sergeant refers to evidence that ADHD can be differentiated from controls on repetitive movements, fine motor difficulty, movement control, balancing, excessive overflow movements, and abnormal rhythmic motor responses.

The target article may not have sufficiently explored the argument that dopamine dysfunction related to the basal ganglia also affects timing, time pacing, and rhythm. Coghill suggests that timing deficits may be a separate pathway to ADHD. Whether timing deficits characterize only a subgroup of ADHD, constituting a separate pathway, or apply to more than one ADHD subgroup needs to be further explored in future studies.

R7. Heterogeneity, variability, individual differences, and development

The DDT describes how the same etiological factor, changes in basic learning mechanisms, may produce a wide range of symptoms or behaviors depending on how these initial changes interact with the environment (target article Figs. 2, 7, sects. 3 and 4). Donahoe & Burgos draw attention to the fact that complexity can result from repeated action of relatively simple processes, which is in accordance with Catania’s point that “as we know from the analysis of nonlinear systems, very small differences in initial conditions can result in exceedingly large long-term differences” (Gleick 1987).

Coghill and Karatekin point out that ADHD may be the developmental outcome of a variety of deficits, consequently producing different developmental trajectories in the various ADHD children who will then be heterogeneous in terms of behavior and comorbidity. Additionally, children with the same underlying deficit may show great variability in symptoms across situations and time. Sadile & Viggiano suggest that a common underlying cause such as dysfunctioning dopamine systems may lead to symptom heterogeneity. The functional state of the dopamine branches may not be the same and may be responsible for heterogeneity in ADHD.

It is important to note, as Sandberg points out, that years of active interaction have taken place before the child’s behavior reaches the level of abnormality needed to reach ADHD diagnostic criteria. Hence, the exact development and symptom profile in each individual will differ because it will depend not only on the initial abnormalities.
in learning processes but also on the environment. We agree with the argument forwarded by Banaschewski et al. that developmental effects and compensatory processes need to be taken into consideration in a developmental theory of ADHD, and we challenge other researchers to complement these aspects of the DDT. Combining these insights with behavior analysis and behavioral techniques, as advocated by Moderato & Presti, will hopefully enable the design of more effective intervention programs for these children.

R8. The functional levels of dopamine and other neurotransmitter systems

Wickens & Tripp and Sadile & Viggiano argue for hyperfunctioning dopamine systems in ADHD and animal models of the disorder. Our position is that the system as a whole might still be less efficient than normal, which probably can best be measured in neurons receiving the dopamine input (and further downstream in behavior). Hence, dopamine neurotransmission can be impaired if tonic dopamine levels are increased or decreased. However, dopamine neurotransmission (i.e., stimulus-evoked release) is decreased in SHR, which can result if extracellular dopamine is increased, or alternatively, it can explain decreased extracellular dopamine. Animal models provide clues and, at best, testable hypotheses, so we do not claim to provide an all-embracing explanation of all of the complex behavioral abnormalities of ADHD but rather attempt to provide a working hypothesis on which to base future investigations.

Banaschewski et al. suggest involvement of dysregulated noradrenergic networks in ADHD associated with impairment of automatic attention processing. We agree that noradrenergic networks may be implicated in ADHD, and there also is convincing evidence that the noradrenergic system is disturbed in SHR. One of the most robust findings is a twofold increase in glutamate-stimulated release of norepinephrine from prefrontal cortex slices of SHR compared to controls (Russell 2001; Russell & Wiggins 2000). Further, the DDT does not address a possible role of serotonin, as noted by Oades & Christiansen. As stated earlier (sect. R1), we wanted DDT to start with a simple set of assumptions and explore the explanatory power. However, we acknowledge the potential importance of other systems, including the noradrenergic and serotonergic systems. It is our belief that the explanation of ADHD may progress more quickly against the error of parsimony than against the confusion of multiple causation. Once the limits of DDT are understood, the need for alternate causal systems will also be understood.

Disturbances in sleep and diurnal rhythm in ADHD (Rothenberger & Kirov; Sergeant) are not addressed by the DDT. Sleep disturbances are, however, found in Parkinson’s disease (Chandhuri 2003), and infusion of a dopamine D1 or D2 receptor agonist in rats increases time spent awake and suppresses rapid eye movement and slow-wave sleep as a function of increasing dose (Isaac & Berbridge 2003). It is interesting that the animal model for ADHD, SHR, has poorer sleep quality and less sleep time than control rats, suggesting that SHR have increased D1 receptor activation, consistent with the hyperdopaminergia hypothesis. Hence, existing findings on sleep and diurnal rhythm indicate the involvement of dopamine. However, the possible link between dopamine dysfunction and sleep problems in ADHD has not been explored in the DDT and is left for future studies.

R9. Conclusions

The DDT offers plausible explanations for a large number of ADHD behaviors both from a purely behavioral point of view and from a neurobiological (hypodopaminergic) point of view. We believe that the commentaries have contributed significantly to elaborating the DDT by directing attention to relationships not previously considered and pointing out aspects of the theory that need to be developed further. The introduction of modified memory processes as a basis for the delay gradients improves the link to aspects of ADHD, and these areas need to be subjected to further research. Important areas for future studies include the relationship between reinforcement decay functions and memory, and between motor functions and timing in ADHD, as well as how initial neurobiological changes in ADHD affect development. Also, advances in neuroscience will produce more precise knowledge about the neurobiological changes in ADHD. These insights will offer a more solid understanding of how altered neurobiological functions interacting with environmental factors produce ADHD symptoms. Currently, it is still unclear whether the hyperactive-impulsive and inattentive subtypes are separate subgroups or may be explained as different outcomes of the same genetics and thus explained by the same principles. Future studies need to investigate the relationship between the ADHD endophenotype suggested in the DDT, involving changes in basic learning processes, and other suggested endophenotypes characterized by delay aversion and impaired response inhibition.

Summary overviews such as the present one are valuable if substantiated and almost equally valuable if some of their claims can be refuted, thus eliminating profitless lines of inquiry. The present commentators, to whom we are most grateful, have added value in both of these ways.

ACKNOWLEDGMENTS

The manuscript was prepared when the authors were Fellows at the Centre for Advanced Study (CAS) at the Norwegian Academy of Sciences and Letters, NO-0271 Oslo, Norway. We thank CAS for generous support.

References

Letters “a” and “r” appearing before authors’ initials refer to target article and response, respectively.


References/Sagolden et al.: A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD)
References/Sagolden et al.: A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD)


References/Sagovlen et al.: A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD)


Holme, E., Nafstad, I., Skare, J. U. & Sagstuen, T. (1998) Behavioural hyperactivity in rats following postnatal exposure to sub-toxic doses of...


References/Sagvolden et al.: A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD)


References/Sagvolden et al.: A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD)


References/Sagvolden et al.: A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD)


Russell, V. A. (2001) Increased AMPA receptor function in slices containing the prefrontal cortex of spontaneously hypertensive rats. Metabolic Brain Disease 16:3–143. [TS]


References

Sagvolden et al.: A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD)


Wright, S. (1939) Statistical genetics in relation to evolution. Hermann. [JWD]


