Rule-based and Information-Integration perceptual category learning in children with Attention Deficit Hyperactivity Disorder

Cynthia L. Huang-Pollock¹²

W. Todd Maddox³

Helen Tam¹

In press Neuropsychology

Abstract

Objective. Suboptimal functioning of the basal ganglia is implicated in ADHD. These structures are important to the acquisition of associative knowledge, leading some to theorize that associative learning deficits might be expected, despite the fact that most extant research in ADHD has focused on effortful control. We present two studies that examine the acquisition of explicit rule-based (RB) and associative information integration (II) category learning among school aged children with ADHD. Methods and Results. In Study 1, we found deficits in both RB and II category learning tasks among children with ADHD (n=81) vs. Controls (n=42). Children with ADHD tended to sort by the more salient but irrelevant dimension (in the RB paradigm) and were unable to acquire a consistent sorting strategy (in the II paradigm). To disentangle whether the deficit was localized to II category learning vs. a generalized inability to consider more than one stimulus dimension, in Study 2 children completed a conjunctive RB paradigm that required consideration of two stimulus dimensions. Children with ADHD (n=50) continued to underperform Controls (n=33). Conclusions. Results provide partial support for neurocognitive developmental theories of ADHD that suggest associative learning deficits should be found, and highlight the importance of utilizing analytic approaches that go beyond asking whether an ADHD-related deficit exists, to why such deficits exist.

Keywords: ADHD, explicit learning, implicit learning, COVIS

Introduction

ADHD is a behavioral syndrome marked by age-inappropriate levels of sustained attention, impulse control, and activity level that is present across multiple environments (APA, 1994). Theoretical approaches to understanding the cognitive mechanism involved in the development of ADHD have primarily focused on deficits in higher order executive control (Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005). Indeed, self-regulatory functions are

¹ Department of Psychology, The Pennsylvania State University, University Park, PA
² This work was supported in part by National Institute of Mental Health Grant R01 MH084947 to Cynthia Huang-Pollock. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health. We thank the parents, teachers, and children who participated, and tireless research assistants who helped in the conduct of the study.
³ Department of Psychology, Institute for Mental Health Research and Institute for Neuroscience, University of Texas, Austin
generally conceptualized as effortful processes used to inhibit counterproductive, negative, and impulsive responses.

However, work in the social and cognitive neurosciences has demonstrated that much of human perception, cognition, and behavior in everyday life can be performed automatically including: goal directed behavior (Bargh & Ferguson, 2000); social perception and interaction (Adolphs, 2009; Bargh & Williams, 2006); and the development and expression of stereotypes (Wheeler & Petty, 2001). Furthermore, even difficult cognitive or sensorimotor procedures that initially require attentional control become automatized following extensive practice (Anderson, 1982; Logan, 1988). In fact, in everyday life, the time or attentional capacity is often not available for conscious deliberative decision making, and there is a growing body of literature that implicit learning and other bottom-up processes are critical to the cognitive underpinnings of intuitive social cognition and behavior (Adolphs, 2009; Frith & Frith, 2006).

**Implicit Learning in ADHD**

The possible presence of an implicit learning deficit in ADHD has not been well evaluated in part because it is conceptualized as a less effortful, more automatic process. However, there is a strong theoretical rationale and converging evidence to suggest that implicit learning may in fact be impaired in ADHD.

First, suboptimal functioning in basal ganglia and frontostriatal neural loops, upon which implicit learning is dependent, is implicated in ADHD (Bush, Valera, & Seidman, 2005; Dickstein, Bannon, Castellanos, & Milham, 2006). Sagvolden et al (2005) specifically argued that the involvement of a hypofunctioning dopaminergic nigrostriatal system would result in impairments in habit learning, a class of implicit learning (Seger, 1994). Nigg and Casey (2005) similarly argued that weaknesses in frontal-striatal and fronto-cerebellar neural loops would result in ADHD-related difficulty predicting both contextual and temporal structures within the environment, structures that develop under the control of an implicit learning system.

The sparse empirical literature on implicit learning in ADHD that exists (Barnes, Howard, Howard, Kenealy, & Vaidya, 2010; Karatekin, White, & Bingham, 2009; Prehn-Kristensen et al., 2011; Vloet et al., 2010; Waber et al., 2003), tends to rely on accuracy rates as a primary performance indicator (but see: Weigard & Huang-Pollock, in review). However, qualitatively different strategies can lead to identical accuracy rates, so the analysis of accuracy alone is quite limited in its ability to explain the presence or absence of individual differences in skill acquisition.

Models of skill acquisition have greatly expanded to include functional (e.g., rule-based models, signal detection models, selective attention models), computational (e.g., COVIS, ALCOVE, ITAM, see Palmeri, Wong, & Gauthier, 2004 for review), neuronal (e.g., Hebbian model, Rescorla-Wagner model), and molecular (e.g., dopamine model) models. Category learning in particular has been extensively studied within the cognitive literature because it is critical for survival, allowing people to make adaptive responses across a variety of situations. Because of its strong theoretical and evidentiary base in the cognitive literature; its ability to shed insight into the strategies that are adopted by participants beyond accuracy analyses; its validation in school aged children (Huang-Pollock, Maddox, & Karalunas, 2011); and its use of novel stimuli to ensure performance is independent of prior learning or academic experiences; we utilize a perceptual categorization paradigm and the associated COVIS (COmpetition between Verbal and Implicit Systems) model herein.

COVIS is a neurobiologically-inspired multiple systems model of category learning that proposes the presence of at least two distinctive systems: an (explicit) verbal rule-based (RB) system and an (implicit) information-integration (II) learning system that learns by association
Category learning in ADHD

(Ashby, Alfonso-Reese, Turken, & Waldron, 1998). Neuropsychological predictions based on the COVIS model have been validated in Parkinson’s (Ashby, Noble, Filoteo, Waldron, & Ell, 2003), Huntington’s (Filoteo, Maddox, & Davis, 2001), Amnesia (Maddox & Filoteo, 2007), and Anorexia Nervosa (Shott et al., 2012).

Verbal RB category learning is dependent on the anterior cingulate and prefrontal cortices, and operates via a hypothesis testing procedure where a verbalized rule is continuously tested and updated until an optimal rule is identified. Examples of verbalizable rules include single feature (if-then), conjunctive (if-and-then), disjunctive (if-or-then), and exceptions (if-then-except) (Minda, Desroches, & Church, 2008b). Verbal skills, working memory, and attentional control facilities play important roles on RB tasks (Maddox, Ashby, & Bohil, 2003).

In contrast, II category learning, mediated predominantly by a set of frontal-striatal structures (i.e. caudate, putamen, and globus pallidus: Keri, 2003; Seger, 1994), learns not by active hypothesis testing, but by gradually recognizing subtle covariations within the environment. It is theoretically best suited for situations in which hundreds if not thousands of exemplars exist, and for which the relation among them cannot easily (if at all) be expressed using verbalizable rule-based algorithms (Ashby et al., 1998). Recent neuroimaging and behavioral studies have also suggested that this system also mediates the learning of unstructured categories. These are categories for which no systematic covariation or relations exist amongst the category exemplar and thus the category associations are arbitrary (Crossley, Madsen, & Ashby, 2012; Seger, 2008).

Research in animal, neuropsychological, and human fMRI work have all presented convergent evidence that RB and II category learning represent distinct systems that partially overlap, and which together guide correct decision making in day to day life (Maddox & Ashby, 2004; Poldrack & Foerde, 2008). Importantly, both humans and non-human primates are biased towards the RB system, even when the optimal strategy requires an II approach. This is likely because RB learning follows a hypothesis testing approach that leads to all-or-none mastery and is therefore faster than the incremental trial-and-error associative process of II learning. This bias must first be inhibited before an II strategy is adopted (Filoteo, Lauritzen, & Maddox, 2010; Zeithamova & Maddox, 2006).

Both RB and II category learning show developmental changes (Huang-Pollock et al., 2011). Compared to adults, typically developing school aged children persist in sorting stimuli along a more salient but irrelevant dimension in a RB task presumably due to immature executive control processes (Huang-Pollock et al., 2011). In turn, ADHD is a disorder that is practically synonymous with deficits in executive functions, and particularly with respect to inhibitory control (Nigg, 2001, 2003). Thus, we would similarly expect children with ADHD to have lower accuracy rates than their same aged peers because of the persistent use of the more salient but irrelevant dimension when sorting stimuli.

Under II conditions, and compared to adults, typically developing school aged children (Huang-Pollock et al., 2011) and older adults (Maddox, Pacheco, Reeves, Zhu, & Schnyer, 2010) show delays in transitioning from RB to II strategy use. The ability to navigate this transition relies on the effective use of feedback to inhibit the bias towards RB learning, which is at least partially mediated by the prefrontal cortex (Ashby & Maddox, 2011). Thus, we might predict that compared to their same aged peers, children with ADHD would likewise be delayed in transitioning from an RB to an II approach under II conditions. However, if children with ADHD have more fundamental deficits in associative learning, they may never transition to II strategy use at all, or may perform in a random fashion. We test these predictions in Study 1.

**Study 1**
Methods

Participants

Sample ethnicity reflected regional demographics: 76% Caucasian, 9% African American, 4% Hispanic, 2% Asian, 4% mixed, 5% unknown. Children aged 8-12 were community recruited from radio, magazine, and internet ads, as well as public flyers posted in the community and provided to local schools in the Central and Dauphin counties of Pennsylvania. Exclusion criteria included: (1) current non-stimulant medication treatment (e.g., neuroleptics, antidepressants), (2) diagnosis of pervasive developmental disorder, intellectual disability, sensorimotor disability, psychosis, or other parent-reported neurological disorder, and (3) estimated Full Scale IQ (FSIQ)<70.

Children with ADHD. Children with ADHD (n=81) met DSM-IV criteria for ADHD including age of onset, duration, cross situational severity, and impairment as determined by parental report on the Diagnostic Interview Schedule for Children-IV (DISC-IV: Shaffer, Fisher, & Lucas, 1997). At least one parent and one teacher report of behavior on the Attention, Hyperactivity, or ADHD subscales of the Behavioral Assessment Scale for Children-2 (BASC-2: Reynolds & Kamphaus, 1992) or the Conner’s Rating Scale (Conners, 1997) was required to exceed the 85th percentile (T-score>61). Both are commonly used and well-validated for the evaluation and diagnosis of ADHD. Following DSM-IV field trials (Lahey et al., 1994), an “or” algorithm integrating parent report on the DISC and teacher report on the ADHD Rating-Scale (DuPaul, Power, Anastopoulos, & Reid, 1998) was used to determine symptom count and subtype (Table 1). Children prescribed stimulant medication (n=28, 34%) were asked to discontinue medication use for 24-48 hours (mean=66 hours).

Controls. Controls (n=42) did not meet ADHD criteria on DISC-IV, had T-scores below the 80th percentile (T-score<58) on all listed parent and teacher rating scales, and had never been previously diagnosed or treated for ADHD. All had ≤4 total symptoms and ≤3 symptoms per ADHD dimension according to the “or” algorithm. The presence of anxiety, depression, oppositional defiant, and conduct disorders were not exclusionary.

Procedures

Informed written consent from parents and verbal assent from children was obtained in accordance with IRB procedures approved by The Pennsylvania State University. Parents were given monetary compensation and relevant clinical feedback. Children were given a small prize.

Intellectual ability. Estimated FSIQ was obtained using a 2-subtest short-form (vocabulary and matrix reasoning, test-retest reliability=0.93, predictive validity=0.87) of the Wechsler Intelligence Scale for Children (WISC-IV: Wechsler, 2003).

Categorization learning paradigms. Gabor patches (Figure 1A) were created using MATLAB routines to randomly sample from two bivariate normal distributions of spatial frequency and orientation (Maddox et al., 2003). All participants received the same randomized order of stimuli. Each task had 5 blocks of 80 trials (40 per category) with visual feedback (“correct” or “wrong”) presented below the stimulus immediately after each response for 500 ms. The task was self-paced. One of the methodological advantages of this approach is that unlike other associative learning paradigms (e.g. probabilistic learning), the stimuli are numerous and varied enough to make the success of memorization unlikely.

The conditions were (1) an RB task that could be solved by sorting along spatial frequency (Figure 1B) and (2) an II task in which both spatial frequency and orientation were needed to achieve optimal performance (Figure 1C). The tasks were counterbalanced and spaced one week apart. Primary results did not vary by counterbalancing order, and counterbalancing order was equivalent between groups, $X^2(1,123)=0.47, p=0.49$. 
In both, participants determined whether the images belonged in Category A or Category B (the “z” and “/” keys; stickers were placed over appropriate keys). They were told that though they might at first be guessing, at some point they would start to get an idea about the right answer. They were instructed to be as accurate as possible, and not to be concerned with speed. Instructions were repeated as necessary; optional rest periods were offered between blocks of 80 trials. Once task instructions were provided, research assistants removed themselves from the child’s direct line of sight but remained present in the room to provide behavioral oversight. Task length was ~20 minutes.

Data analysis

A 2 (Group) x 5 (Block) repeated measures GLM was conducted for accuracy analyses. Because qualitatively different strategies can lead to identical accuracy rates, models were fit to each participant’s set of responses on a block by block basis to determine the type of strategy used. All models assume that participants partition the stimulus space into response regions. Each response region was assigned to a category (e.g., respond A if the line is thick) and the location of each decision criterion was a free parameter.

Five models were applied to the RB data: (1) an RB unidimensional model along the correct frequency dimension that used the optimal decision criterion, (2) an RB unidimensional model along the correct frequency dimension that estimated the decision criterion from the data, (3) an RB unidimensional model along the incorrect orientation dimension that estimated the decision criterion from the data, (4) an II general linear classifier (GLC) model that partitions the stimulus space using a linear decision bound that includes both dimensions, and (5) a random responder model.

Five models were applied to the II data: (1) an II GLC model that assumed the optimal slope and intercept, (2) an II GLC model that estimated the slope and intercept from the data, (3) an RB conjunctive (CJ) model that partitions the stimulus space into four perpendicular regions by frequency and orientation, (4) an RB unidimensional model (along either dimension), and (5) a random responder model. A 2 (Group) x 5 (Block) repeated measures logistical regression was conducted for analyses of strategy use.

All models included a “noise” parameter which provides an estimate of the perceptual and criterial noise associated with classification. The models were fit using maximum likelihood procedures (Ashby & Maddox, 1992) and Akaike Information Criterion (AIC). The AIC statistic penalizes models for each additional free parameter with AIC=(2*n)+(2L), where n equals the number of parameters and L is the maximum likelihood estimate of the data given the model (Akaike, 1974).

The experiments generated a two factor Block (5) x Group (2) design. Dependent variables were accuracy rate, and the proportion of children using a given sorting strategy (i.e. RB, II, etc). SPSS’s GenLin procedure was used to test the effects of Block and Group on strategy use as a dichotomous variable.

Results

Children with ADHD were more hyperactive, impulsive, and inattentive than Controls (all \( p < 0.001, \eta^2_p > 0.30 \); Table 1). There were no group differences in IQ or age (both \( p > 0.26, \eta^2_p < 0.001 \)).

**RB Category Learning.** Main effects of Group, Block, and a Group x Block interaction were found for accuracy (all \( p < 0.03, \eta^2_p > 0.03 \); Figure 2A, Table 2). Controls outperformed children with ADHD, and both groups improved their performance over time. Whereas Controls showed a quadratic early steep learning trajectory that reached asymptote after Block 3,
Analyses of strategy use found that a larger proportion of Controls than children with ADHD sorted stimuli along the correct dimension, $X^2(1, N=99) = 7.08, p = 0.01$ (Figure 2B). The main effect of Block and the Block x Group interaction were not significant (both $p > 0.12$). Conversely, a larger proportion of children with ADHD sorted along the irrelevant dimension $X^2(4, N=99) = 11.99, p = 0.02$, but the Group x Block interaction was not significant, $X^2(4, N=99) = 3.18, p = 0.53$. The proportion of non-ADHD Controls attempting to sort stimuli using an II model was higher than children with ADHD, $X^2(1, N=99) = 7.08, p = 0.01$ (Figure 2C), but the main effect of Block and Block x Group interaction were not significant (both $p > 0.12$). The proportion of children in both groups using a random strategy was essentially zero (Figure 2D).

II Category Learning. Main effects of Group and Block (both $p < 0.02$, $\eta^2_p=0.06$), were found in which Controls outperformed children with ADHD, and accuracy improved over time (Figure 2E; Table 2). The Block x Group interaction was not significant ($p = 0.24$, $\eta^2_p=0.01$).

There was no main effect of Group or Block for the proportion of participants sorting by the correct II strategy (all $p > 0.12$), but a significant Group x Block interaction was found in which more Controls were increasingly best fit by an II strategy than children with ADHD, $X^2(4, N=96) = 14.18, p = 0.01$ (Figure 2F). In contrast, a larger proportion of children with ADHD than Controls were best fit with a random responder model, $X^2(1, N=96) = 4.39, p = 0.04$ (Figure 2H). The main effect of Block and Block x Group interaction for random model use were not significant (both $p > 0.21$). There was no main effect of Group or a Group x Block interaction for use of a unidimensional rule-based strategy (both $p > 0.43$), though unidimensional rule use decreased with Block, $X^2(4, N=96) = 15.05, p = 0.01$ (Figure 2G).

Discussion

In an earlier developmental study comparing typically developing school aged children to adults (Huang-Pollock et al., 2011), children underperformed adults because they persisted in sorting stimuli along the irrelevant dimension. Likewise, in the current study, we found that children with ADHD underperformed Controls for the same reason. Such a finding is consistent with documented ADHD-related weaknesses in interference control (Lansbergen, Kenemans, & van Engeland, 2007; Mullane, Corkum, Klein, & McLaughlin, 2009), as well as evidence for ADHD-related deficits in well-known simple explicit category learning tasks (e.g. WCST: Romine et al., 2004).

With respect to II category learning, typically developing children underperform adults because they have more difficulty inhibiting the bias for the RB system and transitioning to the II system (Huang-Pollock et al., 2011). Thus, children with ADHD might also have been expected to demonstrate this same pattern of performance when compared against non-ADHD controls. Instead, they were best fit by a random responder model even after extensive training. In comparison, with each successive block, more non-ADHD controls were sorting according to the correct II strategy. Thus, both explicit RB and associative II learning were impaired in ADHD, but the cause for poor performance on the II task could not be attributed to an ADHD-related developmental delay in the inhibition of the verbal RB system.

A prominent difference between the RB and II tasks, other than the explicit-associative distinction, is that the latter requires the integration of two dimensions, whereas the former requires attention to only a single dimension. Thus, the apparent ADHD-related deficit in II
Category learning in ADHD

learning, though predicted by theoretical models of ADHD, might instead reflect a general weakness in the simultaneous consideration of two stimulus distributions. With this in mind, in Study 2, we used a task that followed an explicit conjunctive (CJ) rule “If X and Y, then,” that required consideration of two continuous dimensions. To encourage use of both stimulus dimensions, the stimulus variance on the orientation and spatial frequency was equated so the salience of the two dimensions was approximately equal.

**Study 2**

**Methods**

**Participants**

A new sample of children aged 8-12 was recruited, screened, and identified in the same manner as in Study 1. Sample ethnicity was 75% Caucasian, 10% African American, 3.6% Hispanic, 1.2% Asian, and 10.2% mixed ethnicity. **Children with ADHD.** There were 50 children with ADHD; children currently taking stimulant medication (n=11, 22%) discontinued medication use for 24-48 hours (mean=73 hours). **Controls.** Thirty-three typically developing children were recruited. See Table 1.

**Procedure and Data Analysis**

All stimuli, procedures, and instructions were identical to Study 1. In the CJ task, the optimal strategy was an explicit rule that partitioned the stimulus space into four perpendicular regions (Figure 1D). The stimulus variances on the orientation and spatial frequency dimensions were equated so that the salience of the two dimensions was approximately equal. Items in Category A (large spatial frequency, shallow slopes) were over sampled to equate base rate of occurrence. There were 7 blocks of 54 trials. Model fitting procedures were identical to Study 1. Task length was ~20 minutes. Six models were applied to the CJ data: (1) an RB conjunctive model that assumes the optimal decision criteria along the frequency and orientation dimensions, (2) an RB conjunctive model that estimates the decision criteria along the frequency and orientation dimensions from the data, (3) an RB unidimensional model along the frequency dimension, (4) an RB unidimensional model along the orientation dimension, (5) an II general linear classifier (GLC) model that partitions the stimulus space using a linear decision bound that includes both dimensions, and (6) a random responder model.

**Results**

Children with ADHD were more hyperactive, impulsive, and inattentive than Controls (all \( p<0.001, \eta_r^2>0.41; \) Table 1). There were no group differences in IQ or age (both \( p>0.06, \eta_r^2=0.04 \)).

A significant main effect of Group (\( p=0.001, \eta_r^2=0.13 \)) was found in which Controls were more accurate than children with ADHD (Figure 3A; Table 2). The main effect of Block and Block x Group interaction were not significant (both \( p>0.10, \eta_r^2<0.02 \)). However, there was a significant linear trend for increasing accuracy among Controls, \( F(6,192)=2.44, p=0.03, \eta_r^2=0.07 \), but not children with ADHD, \( F(6,294)=0.83, p=0.55, \eta_r^2=0.02 \), which can be seen in Figure 3A.

Analyses of strategy use found no main effect of Group, Block, or Group x Block interaction on the proportion of individuals utilizing an optimal CJ strategy (all \( p>0.06; \) Figures 3B and C). However, a larger proportion of Controls were best fit with a one-dimensional model than children with ADHD, and a larger proportion of children with ADHD were best fit with a random responder model (both \( p<0.001; \) Figures 3C and D). There were no main effects of Block or a Block x Group interaction on model use, except that random model use decreased
with Block, $X^2(6,83) = 17.66, p=0.01$.

**Discussion**

Controls maintained a suboptimal one-dimensional approach and outperformed children with ADHD, who, as in the II task of Study 1, were again best fit with a random responder model. The ADHD-related performance deficit is therefore not specific to II learning, but is seen when the consideration of more than one stimulus dimension is required. We might therefore wonder why children with ADHD would not resort to a single dimensional strategy in the same manner as non-ADHD controls, particularly as they were able to identify a reliable (though incorrect) one dimensional strategy in the RB task of Study 1. We believe that the answer lies in the fact that unlike the RB task of Study 1, in the CJ task, the stimulus variances on the orientation and spatial frequency dimensions were approximately equal so that no single dimension was more salient than the other. It remains to be seen whether children with ADHD would have been able to identify a coherent sorting strategy had the stimulus dimensions been equated in the RB task of Study 1.

Controls do not demonstrate this general inability to incorporate information across multiple dimensions. Instead, their response strategies were responsive to category structure. Use of a unidimensional rule in the CJ task (although less optimal than the correct CJ strategy) could still result in an 83% accuracy rate overall, with 100% accuracy for Category A. In contrast, in the II condition, a unidimensional rule would only yield a maximum 65% accuracy. Unidimensional rule use was therefore not as reinforcing in the II condition, leading to the gradual adoption of the two dimensional II approach for the majority of Controls.

**General Discussion**

Category learning is a fundamental decision process needed for survival, and its underlying neuroanatomical substrates have been extensively studied. Among its many methodological strengths, the paradigm produces a large number of continuous stimuli that prevents participants from memorizing discrete responses. Because the stimuli and paradigm are novel to children, performance is not influenced by the quality, level, or extent of previous academic instruction. Most importantly, however, how a child or any individual solves a given task is an empirical question, and no task is process pure. The computational COVIS model provides important insight into participant response strategies that goes well beyond simple accuracy analyses. Across both studies, we found evidence that children with ADHD underperformed Controls on both explicit RB and associative II category learning tasks.

With respect to explicit learning, a process that optimally results in the creation of a verbalizable, abstract, flexible, and context-independent rule or set of rules (Ashby & Ell, 2001; Ashby & Maddox, 2005; Maddox et al., 2003; Maddox & Ing, 2005), children with ADHD struggled to acquire both a single unidimensional rule (Study 1) as well as a conjunctive rule (Study 2), but for strikingly different reasons. In Study 1, increased variance along the irrelevant dimension interfered with their ability to sort along the correct but less salient dimension. In Study 2, they were not only unable to consider both dimensions to optimize performance, but were also unable to identify a suboptimal one-dimensional approach. Instead, they were best fit with a random responder model.

The ability to learn increasingly complex explicit rules is dependent upon the health and development of the neuroanatomical structures needed to represent such rules, including the prefrontal and medial-temporal cortices, anterior cingulate cortex, and the head of the caudate (Gabrieli, Brewer, Desmond, & Glover, 1997; Minda, Desroches, & Church, 2008a). As such, the ability to explicitly learn is intimately and neuroanatomically related to executive functionality (e.g. Bunge & Zelazo, 2006). The moderating effect of executive demands during
explicit skill acquisition in ADHD, focusing in that case on working memory, has previously been demonstrated (Huang-Pollock & Karalunas, 2010). Many of the fundamental building blocks of academic skill depend upon learning explicit rules that gradually build and increase in complexity (e.g. multiplication and division building upon addition and subtraction). Our findings suggest that impairments in rule-based learning, may represent one important cognitive mechanism that contributes to academic underachievement among children with ADHD (Frazier, Youngstrom, Glutting, & Watkins, 2007).

But what about associative learning processes, which are not solved using rule-based verbal strategies and are less dependent upon executive processes? Although children with ADHD were best fit with the random responder model in the II task, this was also the case for the CJ task, suggesting a broad weakness when required to consider more than one stimulus dimension during category learning. Importantly, results are not due to problems with sustained attention. Problems sustaining attention would be identified if performance deteriorated over time (Huang-Pollock, Nigg, & Halperin, 2006; Huang-Pollock, Karalunas, Tam, & Moore, 2012), a pattern that was not seen for either group of children. Note, however, that the analysis of multiple sorting strategies may have inflated our Type I error rate. It also bears mentioning that the model-fitting process requires that the individual apply a fixed strategy across multiple trials. Therefore, when the random responder model proves the best fit, COVIS is unable to determine whether a participant is in fact responding randomly, or whether s/he is rapidly switching between strategies. Although we are unable to disentangle whether the adoption of random response strategies under II and CJ conditions can be attributed to separate mechanisms, such a question could be addressed in future studies.

Considering CJ first, when either verbal or visual spatial resources are reduced under dual task conditions, adults revert to random or unidimensional strategies, presumably because their ability to analyze the stimuli and their individual dimensions have been disrupted (Miles & Minda, 2011; Zeithamova & Maddox, 2006, 2007). Thus, for the CJ task, reduced WM capacity among children with ADHD (compared to their non-ADHD counterparts), may have similarly disrupted the process of rule discovery and lead to the adoption of a random response strategy.

In contrast, and unlike RB category learning, the learning of II category structures is highly dependent upon feedback timing (Maddox et al., 2003; Maddox & Ing, 2005). In adults, a delay of even a few seconds is enough to disrupt II learning because nonverbal learning is dependent upon the availability and timing of dopamine release (Ashby & Maddox, 2005), which is also believed to be altered in ADHD (Sagvolden et al., 2005). In the paradigms used in the current study, visual feedback (“correct”/”wrong”) was presented immediately following the response and remained onscreen for 500 ms. However, it may be that children with ADHD were unable to fully process this feedback within this period of time, which would have lead them to have an altered feedback experience when compared to non-ADHD Controls. Indeed, organic evidence of error monitoring deficits in ADHD have been found in functional (Depue et al., 2010; Durston et al., 2007), structural (Castellanos et al., 2002; Seidman, Valera, & Makris, 2005), and neural connectivity studies (Konrad & Eickhoff, 2010). Thus, for the II task, the rise of random strategy use might be due to a delayed or an otherwise altered feedback experience.

To better disentangle the mechanisms leading to the adoption of random response strategies in both the CJ and II learning tasks, future studies should consider either prolonging the feedback period or the manner in which feedback is delivered. Such a manipulation would not be expected to improve performance on the CJ task, which, as an explicit rule-based task is robust to disruptions in (or even absence of) feedback (Maddox et al., 2003; Maddox & Ing, 2005). But, if impaired II learning in ADHD was related to feedback processing, such a
manipulation would be expected to improve II category acquisition. We chose to examine performance among children aged 9-12 given previous validation of the COVIS model in this age group (Huang-Pollock et al., 2011), and because it is at this age that children are most commonly identified with ADHD. However, future studies might seek to examine performance in a population of adolescents or adults, as well as preschool children with ADHD to better map the developmental progression of category learning deficits in this population.

**Conclusions**

Overall, we found evidence that children with ADHD have deficits in both explicit RB and associative II perceptual category learning. Deficits in the former are understood to reflect underdeveloped executive systems critical to explicit learning, and deficits in the latter appear to reflect a decreased ability to utilize multiple continuously distributed streams of information in the service of decision making. Our results provide partial support for neurocognitive developmental theories of ADHD that suggest associative learning deficits should be found. That being said, the construct of “implicit learning,” like that of executive function (Miyake et al., 2000), is not unitary (Gebauer & Mackintosh, 2007), and evidence for or against one type of implicit learning process would not necessarily be expected to generalize to others. What is clear, however, is that the COVIS framework allows us to move beyond asking whether an ADHD-related deficit in performance exists, to understanding why such deficits exist.

**Table 1.** Demographic data with means and SD.

<table>
<thead>
<tr>
<th></th>
<th>Study 1</th>
<th></th>
<th>Study 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>ADHD</td>
<td>Control</td>
<td>ADHD</td>
</tr>
<tr>
<td>N (Males:Females)</td>
<td>42 (16:26)</td>
<td>81 (52:29)</td>
<td>33 (14:19)</td>
<td>50 (29:21)</td>
</tr>
<tr>
<td>#Subtypes (H, I, C)</td>
<td>--</td>
<td>5,42,34</td>
<td>--</td>
<td>5,21,24</td>
</tr>
<tr>
<td>Age</td>
<td>10.51 (1.37)</td>
<td>10.35 (1.32)</td>
<td>10.23 (1.20)</td>
<td>9.78 (1.14)</td>
</tr>
<tr>
<td>Estimated Full Scale IQ</td>
<td>106.07 (12.18)</td>
<td>103.20 (13.75)</td>
<td>108.39 (10.53)</td>
<td>103.08 (13.54)</td>
</tr>
<tr>
<td><strong>Hyperactivity/Impulsivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total # of symptoms</td>
<td>0.27 (0.50)</td>
<td>5.10 (2.94)***</td>
<td>0.23 (0.50)</td>
<td>5.72 (2.64)***</td>
</tr>
<tr>
<td>Parent BASC-2 T score</td>
<td>42.79 (5.76)</td>
<td>65.32 (13.47)***</td>
<td>44.52 (6.08)</td>
<td>62.20 (12.11)***</td>
</tr>
<tr>
<td>Parent Conners T score</td>
<td>46.31 (2.91)</td>
<td>67.06 (14.36)***</td>
<td>47.00 (3.23)</td>
<td>66.54 (13.12)***</td>
</tr>
<tr>
<td>Teacher BASC-2 T score</td>
<td>44.05 (5.05)</td>
<td>60.12 (11.83)***</td>
<td>43.06 (3.40)</td>
<td>61.48 (11.58)***</td>
</tr>
<tr>
<td>Teacher Conners T score</td>
<td>46.12 (3.40)</td>
<td>60.16 (11.57)***</td>
<td>45.79 (3.30)</td>
<td>61.46 (11.59)***</td>
</tr>
<tr>
<td><strong>Inattentive symptoms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total # of symptoms</td>
<td>0.74 (0.91)</td>
<td>7.89 (1.40)***</td>
<td>0.70 (0.99)</td>
<td>7.58 (1.54)***</td>
</tr>
<tr>
<td>Parent BASC-2 T score</td>
<td>43.81 (5.88)</td>
<td>66.27 (6.10)***</td>
<td>44.61 (5.78)</td>
<td>66.88 (5.89)***</td>
</tr>
<tr>
<td>Parent Conners T score</td>
<td>46.02 (3.63)</td>
<td>69.47 (11.19)***</td>
<td>46.09 (4.01)</td>
<td>69.64 (12.43)***</td>
</tr>
<tr>
<td>Teacher BASC-2 T score</td>
<td>43.33 (6.59)</td>
<td>61.30 (6.30)***</td>
<td>43.64 (5.02)</td>
<td>62.34 (5.43)***</td>
</tr>
<tr>
<td>Teacher Conners T score</td>
<td>46.64 (4.68)</td>
<td>59.49 (11.11)***</td>
<td>45.30 (3.50)</td>
<td>59.82 (10.71)***</td>
</tr>
</tbody>
</table>

Note: ***p<0.001. * p<0.05
<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>MainEffect/Interaction</th>
<th>F (dfw,dfb), η² or Wald X² (dfw,N), φc</th>
<th>p</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rule-Based Task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>Group</td>
<td>F(1,97)=7.80,0.07</td>
<td>0.01** Control&gt;ADHD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>F(4,388)=10.99,0.10</td>
<td>&lt;0.001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BlockxGroup</td>
<td>F(4,388)=2.73,0.03</td>
<td>0.03*</td>
<td></td>
</tr>
<tr>
<td>Proportion using relevant RB dimension</td>
<td>Group</td>
<td>X² (1,99) = 7.08,0.27</td>
<td>0.01** Control&gt;ADHD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>X² (4.99) = 7.32,0.14</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BlockxGroup</td>
<td>X² (4.99) = 4.11,0.10</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Proportion using irrelevant RB dimension</td>
<td>Group</td>
<td>X² (1,99) = 6.13,0.25</td>
<td>0.01** ADHD&gt;Control</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>X² (4.99) = 11.99,0.17</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BlockxGroup</td>
<td>X² (4.99) = 3.18,0.09</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Proportion using II approach</td>
<td>Group</td>
<td>X² (1,99) = 7.08,0.27</td>
<td>0.01** Control&gt;ADHD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>X² (4.99) = 7.32,0.14</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BlockxGroup</td>
<td>X² (4.99) = 4.11,0.10</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td><strong>Information Integration Task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>Group</td>
<td>F(1,94)=5.62,0.06</td>
<td>0.02* Control&gt;ADHD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>F(4,376)=6.56,0.06</td>
<td>0.001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BlockxGroup</td>
<td>F(4,376)=1.37,0.01</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Proportion using II approach</td>
<td>Group</td>
<td>X² (1,96) = 1.64,0.13</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>X² (4.96) = 7.36,0.14</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BlockxGroup</td>
<td>X² (4.96) = 14.18,0.19</td>
<td>0.01**</td>
<td></td>
</tr>
<tr>
<td>Proportion using 1D approach</td>
<td>Group</td>
<td>X² (1,96) = 0.18,0.04</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>X² (4.96) = 15.05,0.20</td>
<td>0.01**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BlockxGroup</td>
<td>X² (4.96) = 3.82,0.10</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>Proportion using Random approach</td>
<td>Group</td>
<td>X² (1,96) = 4.39,0.21</td>
<td>0.04* ADHD&gt;Control</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>X² (4.96) = 4.61,0.11</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BlockxGroup</td>
<td>X² (4.96) = 5.87,0.12</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td><strong>Conjunctive Rule-Based Task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>Group</td>
<td>F(1,81)=11.99,0.13</td>
<td>0.001** Control&gt;ADHD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>F(6,486)=1.78,0.02</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GroupxBlock</td>
<td>F(6,486)=1.77,0.02</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Proportion using optimal CJ approach</td>
<td>Group</td>
<td>X² (1,83) = 3.61,0.21</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>X² (6,83) = 10.76,0.15</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GroupxBlock</td>
<td>X² (6,83) = 7.56,0.12</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Proportion using 1D approach</td>
<td>Group</td>
<td>X² (1,83) = 12.50,0.39</td>
<td>0.001*** Controls&gt;ADHD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>X² (6,83) = 6.25,0.11</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GroupxBlock</td>
<td>X² (6,83) = 6.49,0.11</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Proportion using Random approach</td>
<td>Group</td>
<td>X² (1,83) = 10.35,0.35</td>
<td>0.001*** ADHD&gt;Controls</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Block</td>
<td>X² (6,83) = 17.66,0.19</td>
<td>0.01**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GroupxBlock</td>
<td>X² (6,83) = 7.34,0.12</td>
<td>0.29</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Stimulus and stimulus plots.

A. Gabor Patch

B. Unidimensional Rule

C. Information Integration

D. Conjunctive Rule
**Figure 2.** Study 1 accuracy and proportion of participants using any given strategy by Block. Solid line=Control, Dashed=ADHD.

**Rule-Based Category Learning**

**A. Accuracy**

![Graph showing accuracy for Rule-Based Category Learning](image)

**B. Relevant (■) vs. irrelevant (●) strategy**

![Graph showing relevant vs. irrelevant strategy](image)

**C. II strategy**

![Graph showing II strategy](image)

**D. Random strategy**

![Graph showing random strategy](image)

**Information Integration Category Learning**

**E. Accuracy**

![Graph showing accuracy for Information Integration Category Learning](image)

**F. II strategy**

![Graph showing II strategy](image)

**G. 1D strategy**

![Graph showing 1D strategy](image)

**H. Random strategy**

![Graph showing random strategy](image)
**Figure 3.** Study 2 accuracy and model use over Block where Squares=Controls, Circles=ADHD.

A. Accuracy CJ

![Graph A](image1)

B. Proportion best fit with optimal CJ model

![Graph B](image2)

C. Proportion best fit with 1D model

![Graph C](image3)

D. Proportion best fit with Random model

![Graph D](image4)
Category learning in ADHD

References


