Neurobiology of Learning and Memory 142 (2017) 200–208

Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

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

# Feedback-based probabilistic category learning is selectively impaired in attention/hyperactivity deficit disorder

Yafit Gabay [a,c,*], Liat Goldfarb [b,c]

[a] Department of Special Education, University of Haifa, Haifa, Israel
[b] Department of Learning Disabilities, University of Haifa, Haifa, Israel
[c] Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Haifa, Israel

ARTICLE INFO

*Article history:*
Received 22 December 2016
Revised 2 April 2017
Accepted 25 April 2017
Available online 3 May 2017

*Keywords:*
Attention-deficit hyperactivity disorder
Probabilistic category learning
Procedural learning
Feedback
Paired-associate
Weather prediction task

ABSTRACT

Although Attention-Deficit Hyperactivity Disorder (ADHD) is closely linked to executive function deficits, it has recently been attributed to procedural learning impairments that are quite distinct from the former. These observations challenge the ability of the executive function framework solely to account for the diverse range of symptoms observed in ADHD. A recent neurocomputational model emphasizes the role of striatal dopamine (DA) in explaining ADHD's broad range of deficits, but the link between this model and procedural learning impairments remains unclear. Significantly, feedback-based procedural learning is hypothesized to be disrupted in ADHD because of the involvement of striatal DA in this type of learning. In order to test this assumption, we employed two variants of a probabilistic category learning task known from the neuropsychological literature. Feedback-based (FB) and paired associate-based (PA) probabilistic category learning were employed in a non-medicated sample of ADHD participants and neurotypical participants. In the FB task, participants learned associations between cues and outcomes initially by guessing and subsequently through feedback indicating the correctness of the response. In the PA learning task, participants viewed the cue and its associated outcome simultaneously without receiving an overt response or corrective feedback. In both tasks, participants were trained across 150 trials. Learning was assessed in a subsequent test without a presentation of the outcome or corrective feedback. Results revealed an interesting disassociation in which ADHD participants performed as well as control participants in the PA task, but were impaired compared with the controls in the FB task. The learning curve during FB training differed between the two groups. Taken together, these results suggest that the ability to incrementally learn by feedback is selectively disrupted in ADHD participants. These results are discussed in relation to both the ADHD dopaminergic dysfunction model and recent findings implicating procedural learning impairments in those with ADHD.

## 1. Introduction

The ability to classify objects and events into distinct categories is important for human cognition. Our actions and decisions are based on categorization abilities that can either be based on a single past experience or be acquired in an incremental manner. A commonly used task to study categorization functions in cognitive neuroscience is the Weather Prediction Task (WPT), which is a typical probabilistic category learning task in which participants learn to classify multi-featured stimuli into one of two categories. This is typically done based on trial-by-trial corrective feedback. In the above-referenced WPT, participants predict an outcome, the weather, based on cues conveyed by a set of geometric features appearing on four individual cards presented in all possible combinations. An important aspect of the weather prediction task is its probabilistic nature. In particular, there is no one-to-one mapping between cues and outcomes. Declarative memorization is a less useful strategy in the weather prediction task because of the probabilistic relationship between cues and outcomes. Instead, the probabilities associated with particular cues and combinations of cues, acquired gradually across trials much as habits or skills are acquired, are most predictive of outcome. People with amnesia due to damage to the medial temporal lobe exhibit intact learning on the weather prediction task, although their declarative knowledge about the learning situation is impaired (Knowlton, Mangels, & Squire, 1996). By contrast, patients with basal ganglia

---

* Corresponding author at: Department of Special Education and The Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Mount Carmel, Haifa 31905, Israel.

*E-mail address:* ygabay@edu.haifa.ac.il (Y. Gabay).

http://dx.doi.org/10.1016/j.nlm.2017.04.012
1074-7427/

disorders such as Parkinson's and Huntington's disease exhibit impaired learning in the weather prediction task (Knowlton, Squire, Paulsen, Swerdlow, & Swenson, 1996; Shohamy et al., 2004). This dissociation suggests the importance of the so-called procedural learning system (including basal ganglia) for probabilistic category learning.

Recent observations are advancing our understanding about how exactly the basal ganglia contribute to incremental learning (such as the kind employed in the WPT). The basal ganglia are paramount to procedural learning, enabling, among other things, the learning and mastering of task performance automatization. Dopaminergic neurons, arising from midbrain nuclei and innervating basal ganglia, have been consistently implicated in contributing to skill learning by mediating feedback processing and reward prediction (Fiorillo, Tobler, & Schultz, 2003; Hollerman & Schultz, 1998; Schultz, 1997; Schultz, Dayan, & Montague, 1997), features that are critical to trial-and-error learning (Shohamy, Myers, Kalanithi, & Gluck, 2008). In order to investigate whether the basal ganglia are critical for learning with feedback, Shohamy et al. (2004) devised two variants of the weather prediction task. A feedback-based (FB) task mirrored the typical weather prediction task. In this variant, participants initially guess the relationship between the probabilistic cues and the outcome and subsequently learn from experimenter-provided feedback about the correct outcome that is signaled by the probabilistic cues. This corrective feedback is eliminated in a paired associate (PA) variant of the weather prediction task. In this task, participants view a cue and its outcome simultaneously and learning proceeds through observation. Thus, in the PA version of the weather prediction task no response is required, except to press a key to advance to the next trial. These two variants of the weather prediction task share the common objective of learning outcomes signaled by a set of probabilistic cues. They differ in whether learning takes place by feedback (FB task) or by observation (PA task). Human functional neuroimaging (fMRI) studies corroborate findings in animals, showing that the WPT instigates basal ganglia response, and does so to a greater extent during feedback-based training than through mere observation devoid of feedback (Poldrack et al., 2001). Similarly, patients suffering from loss of dopaminergic innervation of the basal ganglia (e.g., Parkinson's disease) exhibit impaired learning when trained under feedback-dependent tasks (Knowlton et al., 1996), while maintaining intact performance via observational training (Shohamy et al., 2004; Smith & McDowall, 2006). A recent study offers direct evidence of the significance of midbrain dopamine to feedback-based learning in the WPT. Specifically, using positron emission tomography (PET), Wilkinson et al. (2014) demonstrated dopamine release in the right ventral striatum of healthy participants when performing the WPT based on trial-by-trial feedback, but not in an observational task with no feedback. These findings that patients with Parkinson's and Huntington's disease are impaired in the FB variant of the WPT, but not in the PA variant (Holl, Wilkinson, Tabrizi, Painold, & Jahanshahi, 2012; Shohamy et al., 2004), together with the findings on the involvement of the basal ganglia and striatal DA in the FB variant (Poldrack et al., 2001; Wilkinson et al., 2014), suggest that another population associated with dopaminergic deficiency might also demonstrate this interesting disassociation: the ADHD population.

### 1.1. ADHD and related deficiencies

Attention deficit disorder is one of the most common neurodevelopmental disorders with a prevalence of 3–5% of the general population. It is characterized by age-inappropriate levels of sustained attention, or impulse control, and activity levels that are present across multiple environments (American Psychiatric Association, 1994). ADHD typically surfaces early in childhood, and more often than not persists throughout adolescence and into adulthood (Barkley & Lombroso, 2000). Those affected by ADHD often exhibit significant educational, emotional, and social developmental deficits (Loe & Feldman, 2007; Wehmeier, Schacht, & Barkley, 2010).

Despite decades of research, the source of the neurocognitive dysfunctions and causes of ADHD are still hotly debated (Johnson, Wiersema, & Kuntsi, 2009). It has been suggested that individuals with ADHD suffer from executive function impairments (but see Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005), including set shifting (Boonstra, Kooij, Oosterlaan, Sergeant, & Buitelaar, 2010), planning (Kofman, Larson, & Mostofsky, 2008), working memory (Schweitzer et al., 2000), and inhibition impairments (Barkley, 1997). Indeed, participants with ADHD demonstrate deficits in a variety of inhibition tasks such as the Simon task (Mullane, Corkum, Klein, & McLaughlin, 2009), the continuous performance test (Losier, McGrath, & Klein, 1996), and the stop signal task (Nigg, 1999). Extant literature reveals that along with executive function deficits, motivational processes, and reward-related responses are likewise affected among individuals with ADHD (Aase & Sagvolden, 2006; Luman, Oosterlaan, & Sergeant, 2005; Sagvolden, Aase, Zeiner, & Berger, 1998; Scheres, Milham, Knutson, & Castellanos, 2007; Stark et al., 2011). In particular, it appears that children and adolescents with ADHD are more sensitive to rewards than non-ADHD controls (Fosco, Hawk, Rosch, & Bubnik, 2015; Luman, van Meel, Oosterlaan, & Geurts, 2012), and prefer small immediate rewards to larger delayed rewards (Barkley, Edwards, Laneri, Fletcher, & Metevia, 2001; Demurie, Roeyers, Baeyens, & Sonuga-Barke, 2012; Tripp & Alsop, 2001).

In an attempt to account for the diverse range of deficits associated with ADHD, and in particular the motivational and cognitive impairments, a neurocomputational model was recently suggested by Frank and his colleagues (Frank, 2004; Frank, Santamaria, O'Reilly, & Willcutt, 2007; Maia & Frank, 2011). Their assumption is that striatal dopamine (DA) reduction in ADHD is the common source of both motivational (reinforcement) and cognitive deficits, observed in those with ADHD. In particular, Frank et al. (2007) stated that some of the ADHD cognitive dysfunctions may arise from dysfunctions of both the prefrontal cortex and the dopaminergic dysfunction within the basal ganglia. In support of this model Frank and his colleagues demonstrated that participants with ADHD are impaired in positive (Go) and negative (NoGo) reinforcement learning. Significantly, they found that medications improved Go reinforcement learning relative to NoGo reinforcement learning and that they were predictive of an improvement in the working memory of ADHD individuals in distracting conditions. This finding suggests the presence of common DA mechanisms in ADHD and supports a unified account of the DA function in ADHD.

In addition to the dysfunctions detailed above, procedural learning impairments have been shown to play a role in ADHD. Procedural learning ("how-to knowledge") is related to our ability to acquire skills, habits, and procedures. It is conceived as implicit as it occurs without intention or conscious awareness (Nissen & Bullemer, 1987) and is believed to be free of attentional resources (Frensch, Lin, & Buchner, 1998). Procedural knowledge is difficult to verbalize and is acquired in an incremental manner (Ashby & Casale, 2003). It has been shown that individuals with ADHD exhibit impaired performance in a variety of motor and cognitive procedural learning tasks such as motor sequence tapping (Adi-Japha, Fox, & Karni, 2011; Fox, Adi-Japha, & Karni, 2014, 2016; Fox, Karni, & Adi-Japha, 2016), serial reaction time (Barnes, Howard, Howard, Kenealy, & Vaidya, 2010; Prehn-Kristensen et al., 2011), probabilistic selection (Frank, Santamaria, O'Reilly, & Willcutt, 2007), visual category learning (Huang-Pollock, Maddox, & Tam, 2014), and artificial grammar learning (Laasonen et al., 2014; Rosas et al., 2010). ADHD impairments are evident not only during online skill learning but also during offline consolidation phases, which tend to rely upon the cortical-striatal network (Adi-Japha et al., 2011). On the neural level, neuroimaging studies have demonstrated structural and functional abnormalities in the basal ganglia and cerebellum (core structures of the procedural learning system) and, specifically, deficits in cortico-striatal loops (Berquin et al., 1998; Teicher et al., 2000). ADHD individuals have also been shown to express lower striatal dopaminergic levels (Dougherty et al., 1999; Grace, 2001; Krause, Dresel, Krause, Kung, & Tatsch, 2000) and both children and adults with ADHD have abnormally high densities of dopamine transporters (DATs) (Dougherty et al., 1999; Krause et al., 2000). These observations have led some to view ADHD as a disorder that arises from a selective disruption in the procedural learning systems (Nicolson & Fawcett, 2007, 2011; Ullman, 2004; Ullman & Pullman, 2015). The assumption is that a selective disruption in the procedural learning system leads to a fundamental impairment in the ability to "automatize" behaviors, resulting in increased demands on attentional resources. This is manifested in a learning profile characterized by reduced resistance to interference, sensitivity to distractions, and excessive fatigue, similar to behavioral symptoms observed in ADHD. Notably, procedural learning dysfunctions are intriguing since this type of learning ("how-to knowledge") is not considered an executive function. Studies show that employing a dual-task aimed at blocking executive functions does not harm procedural learning (Foerde, Poldrack, & Knowlton, 2007; Waldron & Ashby, 2001; Zeithamova & Maddox, 2006) and that procedural learning skills are preserved among patients suffering from executive function impairments due to frontal damage (Knowlton et al., 1996). That ADHD involves procedural impairments challenges the executive function framework as the sole account for the diverse range of deficits observed in this condition.

In the present study, we look deeper into the relationship between ADHD and procedural learning dysfunctions by employing the weather prediction task, which is a probabilistic category learning task. Probabilistic category learning tasks have been widely used in neuropsychological research on procedural learning impairments (Knowlton, Squire, & Gluck, 1994; Knowlton et al., 1996; Shohamy, Myers, Onlaor, & Gluck, 2004). Notably, in both the FB and PA versions of the weather prediction task, patients with Parkinson's and Huntington's disease (DA and basal ganglia-related diseases) are impaired in the FB variant of the weather prediction task, but not in the PA variant (Holl et al., 2012; Shohamy et al., 2004). Hence, the same pattern of results might also be observed in the case of ADHD individuals, who also demonstrate DA and basal ganglia-related deficiencies.

Accordingly, in the present study we examine probabilistic category learning in the FB and PA versions of the weather prediction task among ADHD adults and age-matched controls. The influence of feedback on procedural learning in ADHD and the potential disassociation between the FB and PA versions of probabilistic tasks have not yet been assessed systematically. Since striatal DA influences feedback-based learning (Shohamy et al., 2008; Wilkinson et al., 2014), we hypothesize that ADHD's learning impairment will be selective to incremental feedback-based learning in accordance with both procedural and dopaminergic accounts of ADHD (Frank et al., 2007; Nicolson & Fawcett, 2011; Ullman & Pullman, 2015). In particular, we expect a selective disruption of probabilistic category learning in the FB variant of the weather prediction task and intact learning in the PA variant. Poorer learning, relative to controls, among individuals with ADHD in both task variants would be less in favor of the dopaminergic hypothesis since the FB variant of the weather prediction task relies more on dopaminergic release. If probabilistic category learning is unimpaired among ADHD participants, then the performance of ADHD and control participants should not differ.

## 2. Methods

### 2.1. Participants

Eighteen participants with attention deficit disorder and a matched control group participated in the study for a total of 36 participants. All were students from the University of Haifa, Israel, most of whose students come from families of middle to high socioeconomic status. All participants were native Hebrew speakers and their ethnicity was Israeli Jewish. Diagnosis of a comorbid learning disability was an exclusion criterion; a well-documented history of ADHD was the inclusion criterion for the ADHD group. Each individual received a formal diagnosis of ADHD performed by a pediatric neurologist and a positive screening for ADHD-based DSM-5 criteria. Namely, all participants in the ADHD group (except for one) answered the DSM-5 criteria for ADHD, i.e., answering "YES" to at least 5 symptoms of the inattention criteria or the hyperactivity and impulsivity criteria (see Table 1 for group means). The control group was age-matched with the ADHD group, and had no attention problems and a similar level of cognitive abilities that were evaluated using a series of cognitive tests that measured general intelligence, reading comprehension, and math skills. Written informed consent was obtained from all participants. The study was approved by the Institutional Review Board of the University of Haifa and was conducted in accordance with the Declaration of Helsinki.

All participants underwent a series of cognitive tests to evaluate general intelligence (as measured by Raven's SPM tests), reading and math skills. Details about these standardized tasks are presented in Table 1. Results are shown in Table 2. Groups did not differ significantly in age, intelligence, and reading/math skills. However, the ADHD group differed significantly from the control group in the ADHD measures derived from the DSM-5 questionnaire.

### 2.2. Apparatus and materials

Testing took place in a sound-attenuated chamber with participants seated directly in front of a computer monitor during the entire experiment. Stimulus presentation and the recording of response time and accuracy were controlled by a computer program (E-PRIME; Schneider, Eschman, & Zuccolotto, 2002). The stimulus material and card arrangements were similar to those used in the study of Holl et al. (2012) and were created from a set of four tarot cards, each with a different geometric pattern (composed of triangles, circles, diamonds, or squares), arranged horizontally across the middle of the computer screen in black against a white background. See Fig. 1.

Each version of the weather prediction task (FB or PA) included 150 trials during the training phase. On each training trial, participants saw a particular arrangement of cards composed of one, two, or three of the four possible tarot cards. Four-card and no-card arrangements were not used; hence, the experiment included 14 possible card arrangements. Each arrangement was associated with one of the two weather outcomes (Rainy or Fine). Overall, outcomes were presented with equal frequency. Each individual card was associated with a particular outcome with a fixed, independent probability. The probability assigned to each card was counterbalanced and the probability of an outcome in a particular trial was based on the combined probability of the presented cards (see Table 3). Two cards were predictive of fine weather: one strongly (card 4), one weakly (card 3). Two cards were predictive of rainy weather: one strongly (card 1), one weakly (card 2). Overall, participants experienced similar card arrangements, but due to the probabilistic nature of the task the actual outcomes could differ slightly across participants.

**Table 1**
Psychometric tests.

The following tests were administered:

1. *Raven's Standard Progressive Matrices test* (Raven, Court, & Raven, 1992) – Non-verbal intelligence was assessed by the Raven's-SPM test. This task requires participants to choose the item from the bottom of the figure that would complete the pattern at the top. The maximum raw score is 60. Test reliability coefficient is 0.9
2. *One-minute Test for Words* (Shatil, 1995) – Reading skills were examined by the One Minute Test for Words which assesses the number of words accurately read aloud in the space of one minute. The test contains 168 non-vowelized words of an equivalent level of difficulty listed in columns. Words read correctly in the space of one minute are measured
3. *Arithmetic Two-Minute test* – Participants' mathematical automaticity skills were assessed using the Arithmetic Two-Minute test (Openhin-Bitton and Breznitz, unpublished). The task consists of 80 simple arithmetic calculation problems, including the four basic math operations (addition, subtraction, multiplication, and division). The problems are presented in four columns, 20 problems for each basic math operation. Participants are instructed to solve as many problems as possible, from all four types, in 2 min. Total time, accuracy and correct responses per minute are scored
4. *DSM-5 attention/hyperactivity disorder questionnaire Hebrew version* (American Psychiatric Association, 2013) – was used to verify inclusion criteria of the ADHD group. The self-report questionnaire consists of 18 items 9 regarding inattentive symptoms and 9 regrading symptoms of impulsivity and hyperactivity. Each participant is asked to indicate for each item whether he or she experienced the particular symptom

**Table 2**
Demographic and psychometric data of ADHD and control groups.

| Measure | Group | | *P* |
|---|---|---|---|
| | *ADHD*<br>*Mean (SD)* | *Control*<br>*Mean (SD)* | |
| Age (in years) | 26.23 (3.03) | 25.27 (2.46) | *0.12* |
| Raven's SPM | 52.94 (4.54) | 54.27 (5.52) | *0.43* |
| Shatil reading test | 98.5 (15.97) | 106.94 (25.03) | *0.23* |
| Math skills | 66.94 (12.05) | 73.4 (8.26) | *0.07* |
| Inattentive symptoms | 6.33 (2) | 1.5 (1.33) | *0.00* |
| Hyperactive/impulsivity symptoms | 6.72 (2.42) | 2 (2.19) | *0.00* |

Each participant completed the weather prediction task under two different conditions (FB, PA). Thus, two parallel versions of the weather prediction task were employed with different types of cards and different binary outcomes: either Rainy and Fine or Cold and Hot. For half of the participants in each group, Rainy/Fine were the two possible outcomes in the FB condition and Cold/Hot were the outcomes in the PA condition. The remaining participants experienced the reversed pairing. In addition to the set of cards defined by the arrangement of triangles, circles, diamonds, or squares, three additional sets of the four tarot cards were also employed during the experiment, with 25% of participants in each group being trained on each set per weather prediction task variant (FB vs. PA) and with the constraint that participants be trained on a different set of cards in each condition.

### *2.3. Procedure*

The procedure was similar to that of Holl et al. (2012). Participants performed both the FB and PA tasks one after the other. Task order was counterbalanced across participants.

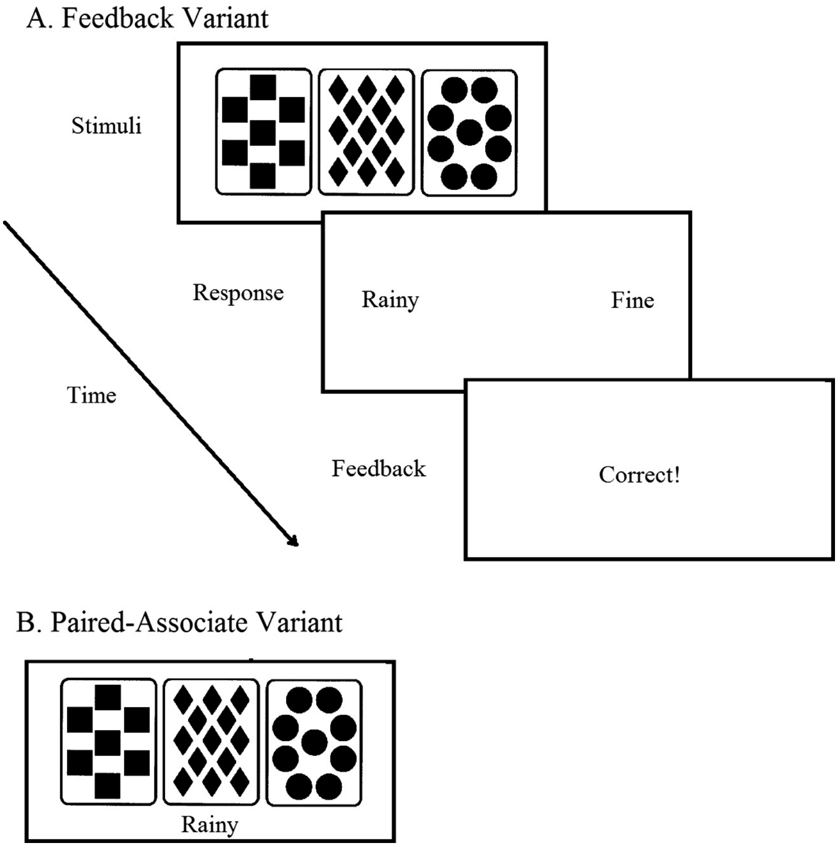

**Fig. 1.** Schematic illustration of the stimuli and tasks.

**Table 3**
Probability structure of the task.

| Pattern | Cue | | | | *P* (cue combination) | | *P* (outcome) |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | *P* (pattern) | Frequency (No. per 200 trials) | |
| A | 0 | 0 | 0 | 1 | 0.0095 | 19 | 0.89 |
| B | 0 | 0 | 1 | 0 | 0.045 | 9 | 0.78 |
| C | 0 | 0 | 1 | 1 | 0.130 | 26 | 0.92 |
| D | 0 | 1 | 0 | 0 | 0.045 | 9 | 0.22 |
| E | 0 | 1 | 1 | 1 | 0.060 | 12 | 0.83 |
| F | 0 | 1 | 1 | 0 | 0.030 | 6 | 0.50 |
| G | 0 | 1 | 1 | 0 | 0.095 | 19 | 0.89 |
| H | 1 | 0 | 0 | 0 | 0.095 | 19 | 0.11 |
| I | 1 | 0 | 0 | 1 | 0.030 | 6 | 0.50 |
| J | 1 | 0 | 1 | 0 | 0.060 | 12 | 0.17 |
| K | 1 | 0 | 1 | 1 | 0.045 | 9 | 0.55 |
| L | 1 | 1 | 0 | 1 | 0.130 | 26 | 0.08 |
| M | 1 | 1 | 0 | 1 | 0.045 | 9 | 0.44 |
| N | 1 | 1 | 1 | 0 | 0.095 | 19 | 0.11 |
| Total | | | | | 1.00 | 200 | |

In any trial, one of 14 possible combinations of four cues could appear with the probability indicated [*P*(pattern)]. Each combination of cues predicted one outcome with a probability of *P*(outcome) and predicted the other outcome with a probability of [1 − *P*(outcome)].

### *2.3.1. Weather prediction task – FB variant*

The training phase consisted of three blocks of 50 trials. In each trial participants saw an arrangement of cards and made a response to predict the weather (Rainy/Fine or Hot/Cold). Feedback appeared immediately after a response, in the form of a written indication presented on the screen to convey whether the weather prediction was correct or incorrect. Participants then requested the next trial with a key press; hence, the task was self-paced. The test phase comprised a further 42 trials with the same structure. On these self-paced trials participants predicted the weather but did not receive feedback.

### *2.3.2. Weather prediction task – PA variant*

The training phase consisted of three blocks of 50 trials. On each trial participants saw an arrangement of cards along with its weather outcome (Rainy/Fine or Hot/Cold). No classification response was required. Participants then requested the next trial, eliciting the appearance of the next card arrangement, along with its weather outcome; hence, the task was self-paced. The test phase was identical to the test phase in the FB version.

### *2.3.3. Awareness tests*

Both FB and PA tasks were followed by tests of awareness. Lagnado, Newell, Kahan, and Shanks (2006) differentiate between participants' insight into the structure of the task (Task Knowledge) and participants' insight into their own judgmental processes (Self-insight). Importantly, the two types of awareness do not necessarily agree. A participant may have an incorrect model of the task, but an accurate model of her own judgments.

### *2.3.4. Task knowledge*

Participants rated how related each card was to the weather outcome using a continuous scale ranging from 0 to 100 (e.g., 0 = definitely rainy, 50 = could be either rainy or fine, and 100 = definitely fine). After participants made a vocal response the experimenter typed the response on the keyboard.

### *2.3.5. Self-insight*

Participants then indicated how important each card was for their weather predictions by rating its importance along a continuous scale ranging from 0 to 100, with 0 = not important at all, 50 = moderately important, 100 = very important. The experimenter typed the participants' vocal response on the keyboard.

## 3. Results

### *3.1. FB vs. PA, test phase*

We first compared the accuracy of the two groups during the test phase of the FB and PA tasks. As in prior studies using the weather prediction task, the correct answer was determined according to the most probable outcome (Gluck, Shohamy, & Myers, 2002).

Preliminary analysis revealed that the order in which the two tasks were performed did not interact with the group variable (minimum $p = 0.109$). Therefore, subsequent analyses collapsed data across order. An analysis of variance (ANOVA) was conducted with Task (FB vs. PA) as a within-subjects factor and Group (ADHD vs. Control) as a between-subjects factor and with mean proportion correct weather predictions during the test phase as the dependent variable. Results are presented in Fig. 2. The main effect of task was significant, $F(1,34) = 33.306$, $p < 0.01$, $\eta_p2 = 0.49$, such that higher accuracy rates were observed during the FB task ($M = 0.85$, $SE = 0.01$) compared with the PA variant ($M = 0.73$, $SE = 0.01$) of the WPT. The main effect group was not significant, $F(1,34) = 1.32$, $p = 0.25$, $\eta_p2 = 0.03$. Importantly, the group by task interaction was significant, $F(1,34) = 4.59$, $p < 0.05$, $\eta_p2 = 0.1$. Further analysis revealed that ADHD participants performed significantly worse on the FB variant of the WPT compared with controls, $F(1,34) = 5.4$, $p < 0.05$, whereas no significant differences were observed between the two groups during the PA variant, $F < 1$. Overall, this indicates an impairment of the ADHD group

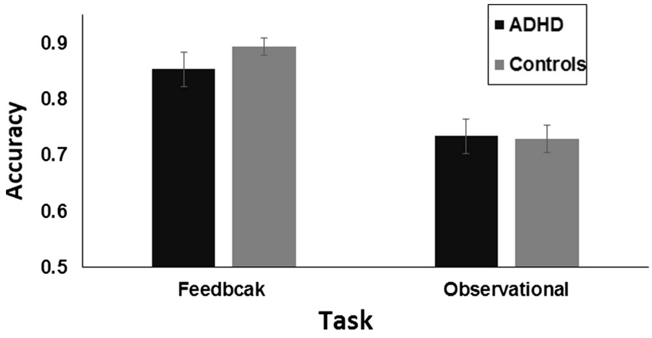

**Fig. 2.** Learning performance measured by mean proportion correct weather prediction accuracy during FB and PA tests of the Weather Prediction Task for the ADHD and Control groups. Error bars represent standard errors.

relative to the control group in feedback-based probabilistic category learning. When cues and outcomes were viewed simultaneously in the absence of corrective feedback, no group differences were observed.

We also compared the differences between the two variants of the WPT separately for each experimental group. Control participants exhibited significantly better learning in the FB task compared with the PA task, $F\ (1,34) = 31.49$, $p < 0.01$. A similar pattern was observed for ADHD participants, $F\ (1,34) = 6.49$, $p < 0.05$.

Previous studies observed a relationship between FB-based learning and IQ scores (Holl et al., 2012). In the present study, no significant differences were observed in IQ scores (as measured by participants' Raven scores) between the ADHD and control groups. Still, we conducted an ANCOVA analysis with mean proportion correct weather predictions during the test phase of the FB task as the dependent variable, group as a between-subjects factor, and IQ scores as a covariate. The main effect of group was significant, $F\ (1,33) = 4.2$, $p < 0.05$. Thus, possible differences in intellectual abilities were not the driving force behind the observed group differences.

### 3.2. FB-based learning across training-trial blocks

We also compared the learning curve of the ADHD and control groups on the FB task. (Note that the learning curve for the PA could not be evaluated because learning took place via observation and no response was required during training.) An analysis of variance (ANOVA) was conducted with Block (trials 1–50, 51–100, 101–150) as a within-subjects factor and Group (ADHD vs. Control) as a between-subjects factor and with mean proportion correct weather predictions during the learning phase as the dependent variable. Results are presented in Fig. 3. The main effect of the group was not significant, $F\ (1,34) = 1.15$, $p = 0.28$, $\eta_p2 = 0.02$. There was a significant main effect of the block, $F\ (2,68) = 5.38$, $p < 0.01$, $\eta_p2 = 0.13$, indicating that participants improved at predicting the weather across trials. The group by block interaction was not significant, $F < 1$. Out of theoretical interest, the performance accuracy in the three training blocks of the FB task was also examined separately for each experimental group. Results revealed a significant linear trend across blocks for the control group, $F\ (1,17) = 10.57$, $p < 0.01$, while no such pattern was observed for the ADHD group, $F\ (1,17) = 1.42$, $p = 0.24$. This pattern of results suggests that only the control group improves linearly with practice.

### 3.3. Awareness: task knowledge

Mean task knowledge difference scores were calculated across the four cards for each participant. A difference score was calculated for each card following the approach of Newell, Lagnado, and Shanks (2007). This was calculated as the actual probability of the negative outcome (0.2, 0.4, 0.6, 0.8, for cards 1–4 respectively) subtracted from a participant's own subjective probability estimate. A positive score is indicative of probability overestimation whereas a negative score is indicative of probability underestimation. Preliminary analysis revealed no significant main effects or interactions with the order in which the task-knowledge tasks were performed across FB and PA tasks (minimum $p = 0.107$). Therefore, the data were collapsed across task presentation order. An analysis of variance (ANOVA) was conducted on the mean difference scores with task (FB vs. PA) as a within-subjects factor and group (ADHD vs. Control) and as a between-subjects factor. Fig. 4 presents task knowledge difference scores for FB and PA tasks for each group. None of the effects were significant (minimum p = 0.08).

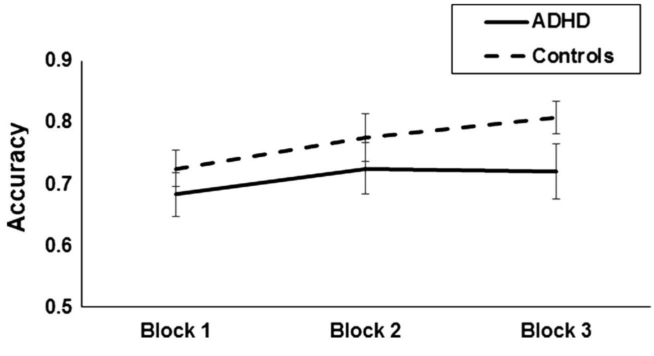

**Fig. 3.** Learning performance measured by mean proportion correct weather predictions during training for the FB task for the ADHD and Control groups. Error bars represent standard errors.

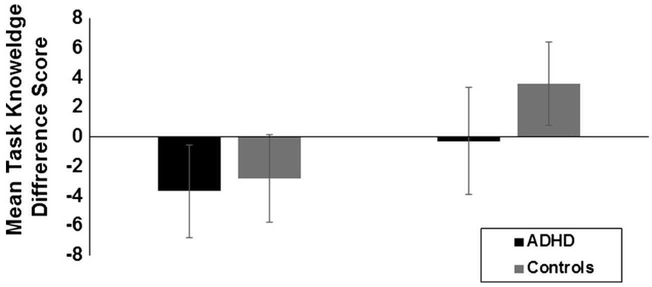

**Fig. 4.** (a) Mean task knowledge difference scores for the ADHD and Control groups. Error bars represent standard errors.

### 3.4. Awareness: self-insight

The main test of self-insight awareness is whether participants' ratings discriminate between strongly and weakly predictive cards. Ratings for the two strongly predictive cards (cards 1, 4) were combined and ratings for the two weakly predictive cards (cards 2, 3) were combined. Preliminary analysis revealed no significant main effects or interactions with the order in which the self-insight tasks were performed across FB and PA tasks (minimum $p = 0.127$). Accordingly, results were analyzed across order. An analysis of variance (ANOVA) was conducted on ratings, with Task (FB vs. PA) and Strength of association between card and outcome (Strong vs. Weak) as a within-subjects factor and Group (ADHD vs. Control) as a between-subjects factor. Fig. 5 presents participants' ratings for strongly and weakly predictive cards for the FB and PA tasks

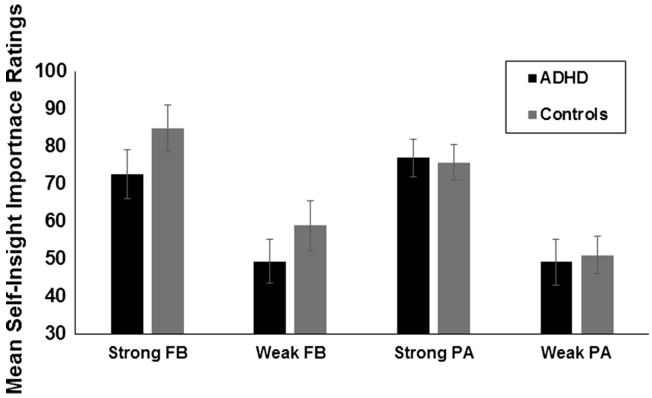

**Fig. 5.** Mean self-insight ratings for strong and weak cards for Dyslexia and Control groups. Error bars represent standard errors.

for each group. There was a significant main effect of card strength, $F(1,34) = 67.95$, $p < 0.01$, $\eta_p 2 = 0.66$, indicating that participants gave higher importance ratings to strong cards compared to weak cards. All other effects were insignificant (minimum $p = 0.167$). There were no significant group differences.

## 4. General discussion

To the best of our knowledge the present study is the first to examine systematically the role corrective feedback plays in procedural learning among participants with ADHD. We examined two versions of the weather prediction task that shared the probabilistic association of cues and outcomes, but differed as to whether learning proceeded via corrective feedback (FB version) or through observation of cues and their outcomes (PA version) between a group of adults with ADHD and matched controls.

Both versions of the weather prediction task rely on probabilistic relationships between cues and outcomes and each task requires learning across probabilistic cue-outcome relationships. The key difference between the PA and FB versions of the WPT is whether learning takes place via observation (PA) or corrective feedback (FB). In the FB task, for which responses were gathered during training trials, the performance of participants with ADHD during the learning phase did not differ significantly from that of the control group. Yet, there was a significant linear trend for the control group that was not observed in the ADHD group. It is possible that with additional training trials, significant group effects would also be observed during the training phase. Furthermore, comparison of categorization accuracy at test phase revealed that ADHD participants learned significantly less than age- and cognitive ability-matched controls in the FB variant of the weather prediction task, whereas no significant group differences were observed during the PA variant. Performance on self-insight and task knowledge tests did not differ significantly between the two groups: both ADHD and control participants accurately ranked the cues according to their predictive values and were accurate to the same degree in determining how relevant the cue was for predicting the outcome.

The observed dissociation between the FB and PA versions of the weather prediction task among the ADHD participants suggests that feedback-based procedural learning is selectively impaired in ADHD. Although procedural learning deficits were implicated in ADHD across a variety of motor, perceptual, and linguistic procedural learning tasks (Adi-Japha et al., 2011; Barnes et al., 2010; Fox, Adi-Japha, et al., 2016; Fox, Karni, et al., 2016; Fox et al., 2014; Frank et al., 2007; Huang-Pollock et al., 2014; Prehn-Kristensen et al., 2011), there were few attempts to examine the role of feedback in modulating procedural learning gains in those with ADHD. In particular, only one study employed feedback-based procedural learning in those with ADHD (Huang-Pollock et al., 2014), but there was no systematic investigation in which feedback was manipulated across tasks using the same experimental design, as was done in the present study. The present results suggest that manipulating the use of feedback in the training experience leads to performance differences between the ADHD group and the control group. Whereas control participants were able to learn better than ADHD when trial-by-trial feedback was provided in an incremental training experience, when training experience was based on cue-outcome observation no group differences were observed. These results are in accordance with previous observations suggesting that feedback processing is impaired in ADHD (van Meel, Oosterlaan, Heslenfeld, & Sergeant, 2005), including a recent study demonstrating impaired probabilistic decision making in the presence of feedback among ADHD participants compared to a situation in which feedback was absent (Pollak & Shoham, 2015).

### 4.1. Hypothesis on the neurocognitive basis of ADHD

It has been suggested that ADHD arises from selective disruption in the procedural learning system (Nicolson & Fawcett, 2007, 2011; Ullman, 2004; Ullman & Pullman, 2015), consistent with observations linking ADHD to an impaired procedural learning network (Berquin et al., 1998; Teicher et al., 2000). However, as stated earlier, it is interesting that both cognitive control deficits and procedural learning deficits coexist in those with ADHD, even though the two represent seemingly dissociated functions (Foerde et al., 2007; Waldron & Ashby, 2001; Zeithamova & Maddox, 2006).

A model that connects the two has been suggested by Frank et al. (2007). According to this neurocomputational model of ADHD, individuals with ADHD suffer from hypersensitivity to phasic DA bursts in the basal ganglia. This hypersensitivity causes impulsive and hyperactive behavior by transiently enhancing BG Go signals and suppressing NoGo signals. Therefore, inappropriate basal ganglia *gating* could lead to frontal-like symptoms such as those related to cognitive control, but could also cause problems with functions relying upon striatal DA such as feedback processing and reward prediction (Shohamy et al., 2008). In accordance with procedural learning accounts of ADHD (Nicolson & Fawcett, 2007, 2011; Ullman, 2004; Ullman & Pullman, 2015), we suggest that procedural learning is disrupted in ADHD individuals but that this difference between the ADHD and control groups is more pronounced when learning takes place on the basis of trial-by-trial feedback rather than by observation because of the involvement of striatal DA in the former type of learning (Wilkinson et al., 2014).

It should be noted that both ADHD and developmental dyslexia (DD) belong to a family of neurodevelopmental disorders that are associated with procedural learning impairments (Nicolson & Fawcett, 2007; Ullman, 2004). Indeed, procedural learning impairments have also been implicated in DD (Gabay & Holt, 2015; Gabay, Schiff, & Vakil, 2012; Howard, Howard, Japikse, & Eden, 2006; Vicari et al., 2005). Furthermore, basal ganglia/cerebellar abnormalities including disruption to cortico-striatal loops have been implicated in DD (Rae et al., 1998). There is a need to consider what distinct cortico-striatal circuit dysfunction might distinguish language disorders from different neurodevelopmental disorders such as ADHD. For example, it is possible that DD is more likely to be associated with cortico-striatal loops involving the dorsal striatum (Krishnan, Watkins, & Bishop, 2016), whereas ADHD is more related to dysfunctions of the ventral striatum and orbitofrontal/prefrontal cortices (Maia & Frank, 2011). A recent study showed that the FB and not the PA variant of the WPT is related to the release of DA in the ventral striatum (Wilkinson et al., 2014). The dissociation we found in our ADHD sample supports the assumption that ADHD is related to dysfunctions of the ventral striatum (Maia & Frank, 2011). Notably, such a dissociation was not found in a previous study investigating DD participants in both the FB and PA versions of the WPT, for whom the probabilistic nature shared by the two tasks impaired learning (Gabay, Vakil, Schiff, & Holt, 2015). Additional neuroimaging research is therefore needed in order to examine the neurobiology of different neurodevelopmental disorders and its influence on procedural learning mechanisms.

The presents study has several limitations. The study included a moderate sample size. This could potentially render group analysis underpowered. Still, the fact that we observed significant group differences with this moderate sample size indicates the robustness of our results. Another issue is related to the fact that in general there is a significant comorbidity between ADHD and reading disorders. This raises the risk that individuals with a comorbidity of reading and attention disorders were included in the ADHD sample. Yet, as noted in the results section, a diagnosis of a comorbid learning disability was an exclusion criterion. This was further verified by the fact that all ADHD participants received equivalent scores to those of the control group on the word reading test. This lessens the possibility of a comorbid reading disability in our ADHD sample.

To conclude, in the present study we observed dissociation between FB and PA variants of the WPT task in an adult sample of participants with ADHD. Feedback-based learning alongside intact PA-based learning is difficult to reconcile with a purely executive function account of ADHD, but it is consistent with a procedural learning deficit in ADHD (Nicolson & Fawcett, 2011; Ullman & Pullman, 2015). Moreover, following Frank's model we suggest that an executive problem in those with ADHD may arise at least in part from an impaired procedural learning system. If basic skills are not well acquired they will necessarily place more demands on attentional resources, leading to a more variable and less consistent learning profile such as the one that is observed in ADHD individuals.

## Acknowledgements

We would like to thank Prof. Marjan Jahanshahi for sharing a sample of cards used in her previous study.

## References

Aase, H., & Sagvolden, T. (2006). Infrequent, but not frequent, reinforcers produce more variable responding and deficient sustained attention in young children with attention-deficit/hyperactivity disorder (ADHD). *Journal of Child Psychology and Psychiatry, 47*(5), 457–471.

Adi-Japha, E., Fox, O., & Karni, A. (2011). Atypical acquisition and atypical expression of memory consolidation gains in a motor skill in young female adults with ADHD. *Research in Developmental Disabilities, 32*(3), 1011–1020.

American Psychiatric Association (1994). *Diagnostic and statistical manual (DSM-IV)*. Washington, DC: American Psychiatric Association.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Ashby, F. G., & Casale, M. B. (2003). The cognitive neuroscience of implicit category learning. *Advances in Consciousness Research, 48*, 109–142.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65.

Barkley, R. A., Edwards, G., Laneri, M., Fletcher, K., & Metevia, L. (2001). Executive functioning, temporal discounting, and sense of time in adolescents with attention deficit hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). *Journal of Abnormal Child Psychology, 29*(6), 541–556.

Barkley, R. A., & Lombroso, P. J. (2000). Genetics of childhood disorders: XVII. ADHD, Part 1: The executive functions and ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(8), 1064–1068.

Barnes, K. A., Howard, J. H., Jr., Howard, D. V., Kenealy, L., & Vaidya, C. J. (2010). Two forms of implicit learning in childhood ADHD. *Developmental Neuropsychology, 35*(5), 494–505.

Berquin, P., Giedd, J., Jacobsen, L., Hamburger, S., Krain, A., Rapoport, J., & Castellanos, F. (1998). Cerebellum in attention-deficit hyperactivity disorder: A morphometric MRI study. *Neurology, 50*(4), 1087–1093.

Boonstra, A. M., Kooij, J., Oosterlaan, J., Sergeant, J. A., & Buitelaar, J. K. (2010). To act or not to act, that's the problem: Primarily inhibition difficulties in adult ADHD. *Neuropsychology, 24*(2), 209.

Demurie, E., Roeyers, H., Baeyens, D., & Sonuga-Barke, E. (2012). Temporal discounting of monetary rewards in children and adolescents with ADHD and autism spectrum disorders. *Developmental Science, 15*(6), 791–800.

Dougherty, D. D., Bonab, A. A., Spencer, T. J., Rauch, S. L., Madras, B. K., & Fischman, A. J. (1999). Dopamine transporter density in patients with attention deficit hyperactivity disorder. *The Lancet, 354*(9196), 2132–2133.

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science, 299*(5614), 1898–1902.

Foerde, K., Poldrack, R. A., & Knowlton, B. J. (2007). Secondary-task effects on classification learning. *Memory & Cognition, 35*(5), 864–874.

Fosco, W. D., Hawk, L. W., Rosch, K. S., & Bubnik, M. G. (2015). Evaluating cognitive and motivational accounts of greater reinforcement effects among children with attention-deficit/hyperactivity disorder. *Behavioral and Brain Functions, 11*(1), 1.

Fox, O., Adi-Japha, E., & Karni, A. (2014). The effect of a skipped dose (placebo) of methylphenidate on the learning and retention of a motor skill in adolescents with Attention Deficit Hyperactivity Disorder. *European Neuropsychopharmacology, 24*(3), 391–396.

Fox, O., Adi-Japha, E., & Karni, A. (2016). Motor memory consolidation processes in young female adults with ADHD may be less susceptible to interference. *Neuroscience Letters, 637*, 91–95.

Fox, O., Karni, A., & Adi-Japha, E. (2016). The consolidation of a motor skill in young adults with ADHD: Shorter practice can be better. *Research in Developmental Disabilities, 51*, 135–144.

Frank, M. J. (2004). *Dynamic dopamine modulation of striato-cortical circuits in cognition: Converging neuropsychological, psychopharmacological and computational studies*. University of Colorado.

Frank, M. J., Santamaria, A., O'Reilly, R. C., & Willcutt, E. (2007). Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology, 32*(7), 1583–1599.

Frensch, P. A., Lin, J., & Buchner, A. (1998). Learning versus behavioral expression of the learned: The effects of a secondary tone-counting task on implicit learning in the serial reaction task. *Psychological Research Psychologische Forschung, 61*(2), 83–98.

Gabay, Y., & Holt, L. L. (2015). Incidental learning of sound categories is impaired in developmental dyslexia. *Cortex, 73*, 131–143.

Gabay, Y., Schiff, R., & Vakil, E. (2012). Dissociation between the procedural learning of letter names and motor sequences in developmental dyslexia. *Neuropsychologia, 50*(10), 2435–2441.

Gabay, Y., Vakil, E., Schiff, R., & Holt, L. L. (2015). Probabilistic category learning in developmental dyslexia: Evidence from feedback and paired-associate weather prediction tasks. *Neuropsychology, 29*(6), 844.

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learning & Memory, 9*(6), 408–418.

Grace, A. A. (2001). Psychostimulant actions on dopamine and limbic system function: Relevance to the pathophysiology and treatment of ADHD. *Stimulant Drugs and ADHD: Basic and Clinical Neuroscience, 5*, 134–157.

Holl, A. K., Wilkinson, L., Tabrizi, S. J., Painold, A., & Jahanshahi, M. (2012). Probabilistic classification learning with corrective feedback is selectively impaired in early Huntington's disease—Evidence for the role of the striatum in learning with feedback. *Neuropsychologia, 50*(9), 2176–2186.

Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience, 1*(4), 304–309.

Howard, J. H., Howard, D. V., Japikse, K. C., & Eden, G. F. (2006). Dyslexics are impaired on implicit higher-order sequence learning, but not on implicit spatial context learning. *Neuropsychologia, 44*(7), 1131–1144.

Huang-Pollock, C. L., Maddox, W. T., & Tam, H. (2014). Rule-based and information-integration perceptual category learning in children with attention-deficit/hyperactivity disorder. *Neuropsychology, 28*(4), 594.

Johnson, K. A., Wiersema, J. R., & Kuntsi, J. (2009). What would Karl Popper say? Are current psychological theories of ADHD falsifiable? *Behavioral and Brain Functions, 5*(1), 1.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*(5280), 1399.

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning & Memory, 1*(2), 106–120.

Knowlton, B. J., Squire, L. R., Paulsen, J. S., Swerdlow, N. R., & Swenson, M. (1996). Dissociations within nondeclarative memory in Huntington's disease. *Neuropsychology, 10*(4), 538.

Kofman, O., Larson, J. G., & Mostofsky, S. H. (2008). A novel task for examining strategic planning: Evidence for impairment in children with ADHD. *Journal of Clinical and Experimental Neuropsychology, 30*(3), 261–271.

Krause, K. H., Dresel, S. H., Krause, J., Kung, H. F., & Tatsch, K. (2000). Increased striatal dopamine transporter in adult patients with attention deficit hyperactivity disorder: Effects of methylphenidate as measured by single photon emission computed tomography. *Neuroscience Letters, 285*(2), 107–110.

Krishnan, S., Watkins, K. E., & Bishop, D. V. (2016). Neurobiological basis of language learning difficulties. *Trends in Cognitive Sciences, 20*(9), 701–714.

Laasonen, M., Väre, J., Oksanen-Hennah, H., Leppämäki, S., Tani, P., Harno, H., & Cleeremans, A. (2014). Project DyAdd: Implicit learning in adult dyslexia and ADHD. *Annals of Dyslexia, 64*(1), 1–33.

Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *Journal of Experimental Psychology: General, 135*(2), 162.

Loe, I. M., & Feldman, H. M. (2007). Academic and educational outcomes of children with ADHD. *Journal of Pediatric Psychology, 32*(6), 643–654.

Losier, B. J., McGrath, P. J., & Klein, R. M. (1996). Error patterns on the continuous performance test in non-medicated and medicated samples of children with and without ADHD: A meta-analytic review. *Journal of Child Psychology and Psychiatry, 37*(8), 971–987.

Luman, M., Oosterlaan, J., & Sergeant, J. A. (2005). The impact of reinforcement contingencies on AD/HD: A review and theoretical appraisal. *Clinical Psychology Review, 25*(2), 183–213.

Luman, M., van Meel, C. S., Oosterlaan, J., & Geurts, H. M. (2012). Reward and punishment sensitivity in children with ADHD: Validating the sensitivity to punishment and sensitivity to reward questionnaire for children (SPSRQ-C). *Journal of Abnormal Child Psychology, 40*(1), 145–157.

Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience, 14*(2), 154–162.

Mullane, J. C., Corkum, P. V., Klein, R. M., & McLaughlin, E. (2009). Interference control in children with and without ADHD: A systematic review of Flanker and Simon task performance. *Child Neuropsychology, 15*(4), 321–342.

Newell, B. R., Lagnado, D. A., & Shanks, D. R. (2007). Challenging the role of implicit processes in probabilistic category learning. *Psychonomic Bulletin & Review, 14*(3), 505–511.

Nicolson, R. I., & Fawcett, A. J. (2007). Procedural learning difficulties: Reuniting the developmental disorders? *Trends in Neurosciences, 30*(4), 135–141.

Nicolson, R. I., & Fawcett, A. J. (2011). Dyslexia, dysgraphia, procedural learning and the cerebellum. *Cortex, 47*(1), 117–127.

Nigg, J. T. (1999). The ADHD response-inhibition deficit as measured by the stop task: Replication with DSM–IV combined type, extension, and qualification. *Journal of Abnormal Child Psychology, 27*(5), 393–402.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*(1), 1–32.

Poldrack, R. A., Clark, J., Pare-Blagoev, E., Shohamy, D., Moyano, J. C., Myers, C., & Gluck, M. (2001). Interactive memory systems in the human brain. *Nature, 414*(6863), 546–550.

Pollak, Y., & Shoham, R. (2015). Feedback may harm: Role of feedback in probabilistic decision making of adolescents with ADHD. *Journal of Abnormal Child Psychology, 43*(7), 1233–1242.

Prehn-Kristensen, A., Molzow, I., Munz, M., Wilhelm, I., Müller, K., Freytag, D., & Baving, L. (2011). Sleep restores daytime deficits in procedural memory in children with attention-deficit/hyperactivity disorder. *Research in Developmental Disabilities, 32*(6), 2480–2488.

Rae, C., Lee, M. A., Dixon, R. M., Blamire, A. M., Thompson, C. H., Styles, P., & Stein, J. F. (1998). Metabolic abnormalities in developmental dyslexia detected by $^1$H magnetic resonance spectroscopy. *The Lancet, 351*(9119), 1849–1852.

Raven, J. C., Court, J. H., & Raven, J. (1992). *Standard progressive matrices*. Oxford, UK: Oxford Psychologists Press.

Rosas, R., Ceric, F., Tenorio, M., Mourgues, C., Thibaut, C., Hurtado, E., & Aravena, M. T. (2010). ADHD children outperform normal children in an artificial grammar implicit learning task: ERP and RT evidence. *Consciousness and Cognition, 19*(1), 341–351.

Sagvolden, T., Aase, H., Zeiner, P., & Berger, D. (1998). Altered reinforcement mechanisms in attention-deficit/hyperactivity disorder. *Behavioural Brain Research, 94*(1), 61–71.

Scheres, A., Milham, M. P., Knutson, B., & Castellanos, F. X. (2007). Ventral striatal hyporesponsiveness during reward anticipation in attention-deficit/hyperactivity disorder. *Biological Psychiatry, 61*(5), 720–724.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh, PA: Psychology Software Tools Inc..

Schultz, W. (1997). The phasic reward signal of primate dopamine neurons. *Advances in Pharmacology, 42*, 686–690.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593–1599.

Schweitzer, J. B., Faber, T. L., Grafton, S. T., Tune, L. E., Hoffman, J. M., & Kilts, C. D. (2000). Alterations in the functional anatomy of working memory in adult attention deficit hyperactivity disorder. *American Journal of Psychiatry, 157*(2), 278–280.

Shatil, E. (1995). *One-minute test for words*. Unpublished test, University of Haifa.

Shohamy, D., Myers, C., Grossman, S., Sage, J., Gluck, M., & Poldrack, R. (2004). Cortico-striatal contributions to feedback-based learning: Converging data from neuroimaging and neuropsychology. *Brain, 127*(4), 851–859.

Shohamy, D., Myers, C., Kalanithi, J., & Gluck, M. (2008). Basal ganglia and dopamine contributions to probabilistic category learning. *Neuroscience & Biobehavioral Reviews, 32*(2), 219–236.

Shohamy, D., Myers, C., Onlaor, S., & Gluck, M. (2004). Role of the basal ganglia in category learning: How do patients with Parkinson's disease learn? *Behavioral Neuroscience, 118*(4), 676.

Smith, J. G., & McDowall, J. (2006). When artificial grammar acquisition in Parkinson's disease is impaired: The case of learning via trial-by-trial feedback. *Brain Research, 1067*(1), 216–228.

Stark, R., Bauer, E., Merz, C., Zimmermann, M., Reuter, M., Plichta, M. M., & Vaitl, D. (2011). ADHD related behaviors are associated with brain activation in the reward system. *Neuropsychologia, 49*(3), 426–434.

Teicher, M. H., Anderson, C. M., Polcari, A., Glod, C. A., Maas, L. C., & Renshaw, P. F. (2000). Functional deficits in basal ganglia of children with attention-deficit/hyperactivity disorder shown with functional magnetic resonance imaging relaxometry. *Nature Medicine, 6*(4), 470–473.

Tripp, G., & Alsop, B. (2001). Sensitivity to reward delay in children with attention deficit hyperactivity disorder (ADHD). *Journal of Child Psychology and Psychiatry, 42*(05), 691–698.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*(1), 231–270.

Ullman, M. T., & Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews, 51*, 205–222.

van Meel, C. S., Oosterlaan, J., Heslenfeld, D. J., & Sergeant, J. A. (2005). Telling good from bad news: ADHD differentially affects processing of positive and negative feedback during guessing. *Neuropsychologia, 43*(13), 1946–1954.

Vicari, S., Finzi, A., Menghini, D., Marotta, L., Baldi, S., & Petrosini, L. (2005). Do children with developmental dyslexia have an implicit learning deficit? *Journal of Neurology, Neurosurgery & Psychiatry, 76*(10), 1392–1397.

Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Review, 8*(1), 168–176.

Wehmeier, P. M., Schacht, A., & Barkley, R. A. (2010). Social and emotional impairment in children and adolescents with ADHD and the impact on quality of life. *Journal of Adolescent Health, 46*(3), 209–217.

Wilkinson, L., Tai, Y. F., Lin, C. S., Lagnado, D. A., Brooks, D. J., Piccini, P., & Jahanshahi, M. (2014). Probabilistic classification learning with corrective feedback is associated with in vivo striatal dopamine release in the ventral striatum, while learning without feedback is not. *Human Brain Mapping, 35*(10), 5106–5115.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry, 57*(11), 1336–1346.

Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & Cognition, 34*(2), 387–398.