Journal of Attention Disorders
Volume 10 Number 2
November 2006 160-170

10.1177/1087054706288116
http://jad.sagepub.com
hosted at
http://online.sagepub.com

Does ADHD in Adults Affect the Relative Accuracy of Metamemory Judgments?

Laura E. Knouse
Matthew J. Paradise
University of North Carolina at Greensboro
John Dunlosky
Kent State University

Objective: Prior research suggests that individuals with ADHD overestimate their performance across domains despite performing more poorly in these domains. The authors introduce measures of accuracy from the larger realm of judgment and decision making—namely, relative accuracy and calibration—to the study of self-evaluative judgment accuracy in adults with ADHD. **Method:** Twenty-eight adults with ADHD and 28 matched controls participate in a computer-administered paired-associate learning task and predict their future recall using immediate and delayed judgments of learning (JOLs). Retrospective confidence judgments are also collected. **Results:** Groups perform equally in terms of judgment magnitude and absolute judgment accuracy as measured by discrepancy scores and calibration curves. Both groups benefit equally from making their JOL at a delay, and the group with ADHD show higher relative accuracy for delayed judgments. **Conclusion:** Results suggest that under certain circumstances, adults with ADHD can make accurate judgments about their future memory. *(J. of Att. Dis. 2006;10(2) 160-170)*

Keywords: *ADHD; memory; metacognition; judgments; paired-associate learning*

In prior research, ADHD has been associated with cognitive deficits, including problems with working memory (Barkley, Murphy, & Kwasnik, 1996), behavioral inhibition (Nigg, Butler, Huang-Pollock, & Henderson, 2002), and strategy use (Hervey, Epstein, & Curry, 2004). In the present research, we investigate an issue of growing relevance in the field—namely, the accuracy of self-perceptions. This issue is of particular importance because inaccurate self-evaluations may undermine performance on a variety of tasks (Dunlosky, Hertzog, Kennedy, & Thiede, 2005). In addition, prominent theories posit that impaired self-awareness is a feature of this disorder. For example, Barkley (1998) predicts that problems with self-awareness arise from working memory impairments in those with ADHD.

In more than a decade of research, the prevailing conclusion has been that people with ADHD tend to overrate their abilities across a variety of domains compared with control participants (e.g., Hoza, Pelham, Dobbs, Owens, & Pillow, 2002; O'Neill & Douglas, 1991). These studies typically report a measure of absolute accuracy—the extent to which a person predicts his or her actual level of performance on a task. For instance, Knouse, Bagwell, Barkley, and Murphy (2005) had adults with and without ADHD first estimate their percentile ranking within the general population for driving ability. They then calculated each participant's percentile ranking within their sample based on a variety of driving indices (collisions, driving citations, and use of safe driver behavior). To examine absolute accuracy, they subtracted each participant's actual ranking from their estimated ranking. This created a discrepancy score, where a positive number indicated an overestimation and a negative number indicated an underestimation. Knouse et al. (2005) found that the group with ADHD overestimated their driving ability to a greater extent than did the non-ADHD group. This pattern of results is similar to that obtained in several other studies examining children with ADHD. Specifically, discrepancy scores are consistently larger (more positive) for individuals with ADHD across a range of domains (Gerdes, Hoza, & Pelham, 2003; Hoza et al., 2002; Hoza et al., 2004; Owens & Hoza,

Authors' Note: Address correspondence to Laura E. Knouse, Department of Psychology, 296 Eberhart Building, P.O. Box 26170, University of North Carolina at Greensboro, Greensboro, NC 27402-6170; e-mail: LEKNOUSE@uncg.edu.

2003). Other studies examining performance on laboratory tasks have not computed discrepancy scores but have replicated a pattern of overestimation by individuals with ADHD. Such tasks include recall for items from a story (O'Neill & Douglas, 1991), completion of word puzzles (Whalen, Henker, Hinshaw, Heller, & Huber-Dressler, 1991), and performance during social tasks (Diener & Milich, 1997; Hoza, Waschbusch, Pelham, Molina, & Milich, 2000; Ohan & Johnston, 2002).

Absolute accuracy is undoubtedly an important facet of self-evaluative judgments, but another aspect of predictive accuracy—called relative accuracy—is just as important and has been overlooked in the current ADHD literature. In contrast to absolute accuracy, relative accuracy is the degree to which a person's judgments discriminate performance among test items. For instance, in the domain of memory, an individual may be asked to study a set of paired associates (e.g., doctor – lobster, cat – fork). After studying each one, he or she would then predict the likelihood of recalling the response (e.g., lobster) when later shown the cue (i.e., doctor - ?) during the test. To measure relative accuracy, the person's judgments would then be correlated with test performance across items. Higher correlations indicate greater relative accuracy of the judgments at discriminating between the recall of one item compared to another. Similar to absolute accuracy, a high level of relative accuracy is important for the effective regulation of behavior (for a review, see Dunlosky et al., 2005).

Importantly, absolute accuracy and relative accuracy are computationally and functionally distinct (Lichtenstein & Fischhoff, 1977; Nelson, 1996), so outcomes concerning one measure may not generalize to the other (Koriat, Sheffer, & Ma'ayan, 2002). For instance, a person may overestimate overall performance (i.e., have poor absolute accuracy for an entire set of items) yet have perfect relative accuracy for discriminating between performance level across items. Despite this, none of the current articles on self-evaluative judgments and ADHD have examined relative accuracy. Accordingly, a major goal of the present research was to estimate relative accuracy on a laboratory task for people with ADHD.

Even though ADHD is commonly viewed as a childhood disorder, the prevalence rate in the adult population has been estimated to be 3% to 5% (DuPaul et al., 2001; Heiligenstein, Conyers, Berns, & Smith, 1998; Murphy & Barkley, 1996). Longitudinal studies indicate that 50% to 80% of children with ADHD continue to meet criteria for the disorder into adolescence and beyond (Barkley, Fischer, Smallish, & Fletcher, 2002; Weiss & Hechtman, 1993). Cognitive deficits in adult ADHD analogous to those associated with childhood ADHD, including behavioral inhibition and vigilance, have been documented in an increasing number of studies (Nigg et al., 2002; Nigg et al., in press). However, the current literature has largely ignored the accuracy of self-evaluations in adults with ADHD, and only one published study (Knouse et al., 2005) has tackled the question directly. For this reason, we chose to examine the judgment accuracy of adults with ADHD.

To do so, we used a well-established method for examining the accuracy of recall predictions from the cognitive literature on metamemory (for reviews, see Koriat, 1997; Schwartz, 1994). As in the example above, participants study noun pairs and make a judgment of learning (JOL)—a prediction of the likelihood of recalling each item on the criterion test—for each pair. After studying and judging items, participants are then tested for recall. Following each test trial, participants also judge the likelihood that their response is correct. This retrospective confidence judgment is used to calculate the accuracy with which a participant can evaluate the correctness of each response. We chose this particular task for a variety of reasons. First, this particular metamemory task is sensitive to detecting deficits in judgment ability that are manifest in other kinds of disorder, such as Korsakoff's disease and frontal lobe damage. Second, current theory of cognitive processing suggests that people with reduced attentional capacity may find this particular task challenging (Ryan, Petty, & Wenzlaff, 1982), which in turn may reveal deficits in judgment ability. Finally, and most important, this task allowed us to estimate the effects of ADHD on both the absolute accuracy and the relative accuracy of predictive JOLs as well as on retrospective judgments of confidence.

Each participant also made two kinds of JOL: immediate and delayed. When making immediate JOLs, the participant was cued by the first word of the to-be-remembered pair immediately after the offset of an item for study. For delayed JOLs, the cue was presented after a delay from the study of an item—an interval that was filled with the study and judgments of other items (Nelson, Narens, & Dunlosky, 2004). The question of delayed JOL accuracy in people with ADHD is especially important because relative accuracy for delayed JOLs is typically quite high in studies of nonclinical populations. Thus, if people with ADHD demonstrate relatively normal levels of accuracy on delayed JOLs, they may also be able to use these judgments to effectively control learning (e.g., Dunlosky, Kubat-Silman, & Hertzog, 2003).

Given the poorer absolute accuracy for people with ADHD demonstrated in previous research in other domains, we would expect similar deficits in absolute accuracy for this population. Relative accuracy has not yet been examined in the research on ADHD. Based on Barkley's (1998) conjecture that people with ADHD are impaired in self-awareness, one might expect general deficits across both absolute and relative accuracy. Nevertheless, because the two kinds of accuracy are computationally and conceptually distinct, people with ADHD may show relative accuracy equal to that of unaffected adults. Although there is no direct precedent for such a result in the ADHD literature, other populations with cognitive deficits have demonstrated relative accuracy of memory judgments similar to unaffected individuals. Such results have been obtained in studies of older adults (e.g., Connor, Dunlosky, & Hertzog, 1997), young children (Koriat & Schitzer-Reichert, 2002), people with traumatic brain injury (Kennedy & Yorkston, 2000), and people with Alzheimer's disease (Moulin, Perfect, & Fitch, 2002). Parallel findings for people with ADHD would suggest an intact aspect of self-evaluative judgment.

Method

Participants

Participants were adults reporting a current diagnosis of ADHD who met research criteria based on symptom counts, developmental deviance, and functional impairment. Twenty-eight adults meeting inclusion criteria for the group with ADHD were compared to 28 age- and gender-matched controls. Participants with ADHD were recruited from multiple sources, including the undergraduate subject pool at a public university in the southeastern United States, that institution's office of disability services, an ADHD specialty clinic at that institution, and from the offices of mental health professionals in the surrounding community. Seventy-five percent of the sample with ADHD were university students, whereas the remaining 25% were members of the surrounding community. The age- and gender-matched non-ADHD group was recruited from the undergraduate subject pool and from the local community. In most cases, age matches were available in the general non-ADHD sample. However, in a few instances, non-ADHD participants of a certain age were actively recruited to match an adult with ADHD. Although most matches were born in the same year, the maximum difference in age between matches was 3 years.

Criteria for the group with ADHD included a current ADHD diagnosis, reported ADHD symptoms above *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM-IV*; American Psychiatric Association, 1994) thresholds, evidence of functional impairment related to symptoms, and developmental deviance of symptoms—specifically, (a) self-reported history of prior ADHD diagnosis by a health care professional, (b) evidence of six or more self-reported inattentive symptoms or six or more hyperactive/impulsive symptoms from a modified version of the ADHD Rating Scale (ADHD-RS; DuPaul, Power, Anastopoulos, & Reid, 1998) or the Conners Adult ADHD Rating Scale (CAARS; Conners, Erhardt, & Sparrow, 1999), (c) evidence of functional impairment related to reported symptoms (e.g., problems in daily activities, management of money, relationships), and (d) evidence of current problematic inattentive or hyperactive symptoms compared to population norms as indicated by a t score greater than 64 (92nd percentile) on at least one subscale of the CAARS. Participants recruited for the non-ADHD group were excluded if they met any of these criteria. Participants were excluded from either group if they reported a history of mental retardation, psychosis, neurological disability, or traumatic brain injury. A single participant from the original ADHD sample reported traumatic brain injury in childhood and was therefore excluded. Five participants in the ADHD group self-reported a history of reading disorder. When we performed statistical analyses with these participants excluded, our results and conclusions remained unchanged, and thus, we retained these participants for the analyses.

Participant characteristics are displayed in Table 1 and those who participated ranged in age from 18 to 60 (M = 23.82, SD = 9.96). Each group contained 11 males (39%) and 17 females (61%). With regard to race, 100% of participants with ADHD self-identified as Caucasian compared to 89% in the control group. Two of the control group participants self-identified as African American and one as Asian. The groups did not differ significantly with respect to years of education, and there were no statistically significant differences between groups in performance on either the verbal or nonverbal sections of the Shipley Institute of Living scale (see Table 1). Self-reported use of substances on the day of the testing session did not differ between groups. Thus, it was not necessary to consider same-day substance use as a covariate in the statistical analyses. In the overall sample, 0% used alcohol, 0% used illegal drugs, 20% used tobacco, 18% used over-the-counter medications, and 32% used prescription medications (other than medications for ADHD) on the day of the testing session.

When reporting on their prior diagnosis, 5 participants in the group with ADHD reported being initially diagnosed by a pediatrician, 1 by a family doctor, 9 by a psychiatrist, and the remaining 13 by a clinical psychologist. Eighty-nine percent of the group with ADHD (25 participants) reported taking medication for their symptoms at some time during their lives, and 57% of the sample

Table 1
Comparison of Group Characteristics

	Group			
	ADHD		Non-ADHD	
Measure	*M*	*SD*	*M*	*SD*
Age	23.82	9.67	23.82	10.43
Education	14.36	1.73	14.19	1.67
SCL-90 GSI	0.90*	0.49*	0.50*	0.47*
Shipley Verbal	29.44	4.15	29.79	4.02
Shipley Nonverbal	17.19	2.20	16.29	2.76
ADHD-RS Severity Score				
IA	13.75*	4.27*	3.80*	3.27*
HY/IM	14.86*	5.13*	5.60*	4.11*
All symptoms	30.43*	8.00*	10.07*	6.90*
ADHD-RS Symptoms				
IA	5.00*	2.58*	0.82*	1.16*
HY/IM	4.68*	2.33*	1.14*	1.24*
All symptoms	9.68*	3.81*	1.90*	2.24*
CAARS				
Total scores	111.42*	29.40*	49.19*	24.55*
Total *t* scores	67.72*	7.19*	45.89*	8.83*
Areas of impairment	4.43*	2.44*	0.54*	1.14*

Note: SCL-90 GSI = Symptom Checklist-90 Global Symptoms Index; ADHD-RS = ADHD Rating Scale; IA = Inattentive; HY/IM = Hyperactive/Impulsive; CAARS = Conners Adult ADHD Rating Scale.
*$p < .01$.

(16 participants) were taking medication at the time of the study. Slightly more than half of those taking ADHD medications (56%) took them on a daily basis, whereas 25% took them as needed. The remainder took their medication only on weekdays but not weekends.

Materials

Metamemory task. The current study used a JOL metamemory methodology to compare various metamemory judgments to objective recall data in participants with and without ADHD. The JOL methodology is used to gauge a participant's accuracy in monitoring his or her memory for a particular set of stimuli and has been validated extensively in the cognitive psychology literature (Nelson & Narens, 1990).

The memory trials were administered via a computerized protocol in which 60 pairs of unrelated nouns (e.g., daffodil—blood, potato—frog, icebox—acrobat) were used as the stimuli for the paired-associate recall task (for details on stimulus construction and larger stimulus sets comparable to those used here, see Connor et al., 1997). Participants were briefed on the procedure in person by the experimenter as well as by written directions and examples presented on the computer screen. At the beginning of the task, participants were informed that they would study pairs of words (e.g., doctor – lobster) for a later recall test and that they would also be making predictions about the likelihood that they would correctly recall the matched associate for each item.

During the study phase, each item pair appeared individually on the computer screen for 8 s (e.g. doctor-lobster) in a random order. Half of the items were randomly assigned to receive immediate JOLs and half to receive delayed JOLs. Immediate JOL items were followed by a cue consisting of the first word of the pair (e.g., doctor-?). The judgment was requested using the following:

> How confident are you that in about 10 minutes from now you will be able to recall the second word of the item when prompted with the first word? (0 = *definitely won't recall*, 2 = *2% sure*, 20 = *20% sure*, 100 = *definitely will recall*)?

Participants could enter any whole number between 0 and 100. For the delayed JOL items, the cue (e.g., doctor-?) and request for judgment appeared after a random number of intervening items. Twenty seconds were allotted for each participant to make his or her rating after which an error message appeared and the trial was considered an error trial. After all items were studied and rated, participants began the recall test phase of the experiment.

During the recall phase, the first noun of each noun-noun pair appeared on the computer screen followed by a blank. Participants were prompted to type the second word of the pair, if possible. Recall trials were self-paced. After each recall attempt, a retrospective confidence judgment was requested for that item. Specifically, participants were asked to rate their confidence about the accuracy of the preceding response on the 0 to 100 scale used to rate JOL.

Two procedures were used to measure absolute accuracy: discrepancy scores and calibration curves. Discrepancy scores were calculated for each participant across judgments by converting mean JOL to a proportion (i.e., JOL / 100) and subtracting this from the actual proportion of items recalled. This procedure yields a score where a positive number represents overestimations of performance and negative numbers represent underestimations. Calibration curves were constructed by plotting the mean JOL for a group of item judgments (participants' predicted recall probabilities) against the actual recall probability within that group of items. Judgments are accurate in the absolute sense to the extent that they conform to the diagonal. Relative accuracy was examined using the Goodman-Kruskall gamma correlation (Nelson & Narens, 1990)—a correlation that assesses the correspondence between a continuous measure (0 to 100 JOL) and a dichotomous outcome (correct vs. incorrect).

ADHD symptoms. The ADHD-RS (DuPaul et al., 1998) lists the *DSM-IV* inattentive and hyperactive-impulsive symptoms for the disorder, and participants rate the frequency of each behavior on a 4-point Likert-type scale (0 = *never or rarely*, 1 = *sometimes*, 2 = *often*, 4 = *very often*). A modified adult self-report version of this scale was used in the current study similar to that used by Murphy and Barkley (1996). Ratings of current symptoms were used in this study. The measures yield both severity scores (totals across items) and symptom counts (number of symptoms rated at least often) for each symptom list and for the entire scale. This measure was used to obtain a tally of endorsed symptoms, and norms associated with this scale were not used. Internal consistency for the ADHD-RS items in this sample was very good ($\alpha = .93$). The Conners Adult ADHD-RS (Conners et al., 1999) is a normative rating scale that assesses problematic ADHD symptoms continuing into adulthood. It is a 66-item self-report measure containing four subscales: Inattention/Memory, Hyperactivity, Impulsivity/Emotional Labiality, and Poor Self-Concept. The cutoff score for problematic symptoms on this normed scale has been set at $t = 64$ (92nd percentile), and this criterion was used in defining the group with ADHD. Internal consistency for this measure in the current study was very good ($\alpha = .98$).

Functional impairment. These items were a subset of items from the ADHD Current Symptom Scale (Barkley & Murphy, 1998). Participants were asked to indicate the extent to which problems indicated on the ADHD-RS (i.e., problems with inattention, hyperactivity, and impulsivity) interfere with their ability to function in several areas of life and activities. Internal consistency for this scale in the current study was very good ($\alpha = .92$).

Psychiatric symptoms. The Symptom Checklist-90-Revised (SCL-90-R) was used to assess the level and severity of psychiatric symptoms among participants (Derogatis, 1975). It is a 90-item, multidimensional self-report inventory designed to screen for a broad range of psychological problems and symptoms of psychopathology across nine symptom dimensions. The SCL-90 yields a global index of severity of psychiatric symptoms (the Global Severity Index [total score/90], which can range from 0 to 4). Internal consistency for this measure in the current study was very good ($\alpha = .97$).

Intellectual ability. The Shipley Institute of Living Scale was used as a general screen for intellectual ability (Shipley, 1946). This timed measure includes vocabulary and pattern completion tasks designed to provide a rough estimate of verbal and nonverbal intellectual ability.

Procedure

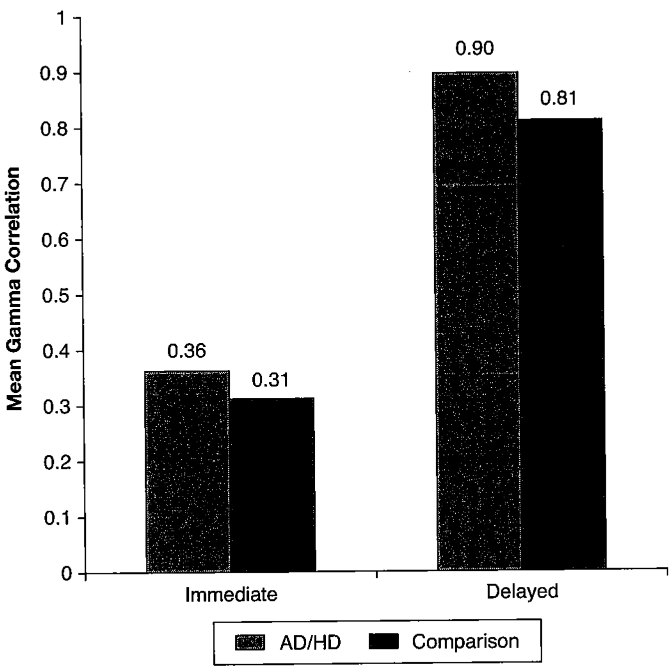
Arrangements were made with participants who took stimulant medication for ADHD to remain medication free on the day that they participated in the experiment.[1] During the 90-min experimental session, participants were greeted by the experimenter and asked to read and sign an Institutional Review Board–approved consent form. They were then seated at a computer workstation and given verbal instructions on the metamemory task. They were asked to complete the automated metamemory task described above, which included a set of instructions and examples presented on the screen prior to the experimental task. Following this, the participants were administered relevant self-report questionnaires and finally completed the Shipley Institute of Living scale. Participants received either a gift certificate or were given credit for the research participation component of an introduction to psychology course. They were then debriefed and thanked for their participation.

Results

Rating Scale Data

Total scores on the ADHD-RS, also called severity scores, were examined between groups for inattentive and hyperactive-impulsive *DSM-IV* symptoms (Table 1). Participants in the group with ADHD had higher scores across all symptom groups with the strongest difference in terms of total symptoms. The group with ADHD reported significantly more current symptoms above threshold (rated at least "often") for inattentive, hyperactive/impulsive, and combined lists with the strongest between-group effect size for combined symptoms. Adults with ADHD reported an average of 9.68 symptoms, whereas the non-ADHD group reported an average symptom count of 1.90. On the Conners Adult ADHD-RS, the mean total t scores for the group with ADHD was significantly higher than the non-ADHD group corresponding to the 96th versus the 34th percentile, respectively (see Table 1). Thus, symptom counts in the group with ADHD were developmentally deviant compared to peers. Based on symptoms reported via the ADHD-RS, 8 members of the group with ADHD met or exceeded the symptom threshold for ADHD predominantly inattentive type, and 12 met or exceeded the threshold for ADHD predominantly hyperactive/impulsive type. The remaining 8 participants in this group exceeded self-reported symptom thresholds associated with ADHD combined type. However, it is important to note that these groupings were formed based on self-reported symptom counts only and not full diagnostic criteria.

Figure 1
Mean Relative Accuracy (Gamma Correlation) of Judgment of Learning by Group and Time of Judgment



In terms of functional impairment, all participants in the group with ADHD reported that they often experienced problems in at least two different domains of functioning as a result of ADHD symptoms. The actual number of problem areas endorsed ranged from 2 to 10 ($M = 4.50$, $SD = 2.33$). The group with ADHD did have a higher mean score on the Global Severity Index of the SCL-90 indicating greater severity of overall psychiatric symptoms.

Relative Accuracy

Judgments of learning. Goodman-Kruskal gamma correlations between item-by-item JOLs and recall (correct vs. incorrect) were computed for each participant as an index of relative JOL accuracy (for detailed rationale for the use of gamma as a measure of relative judgment accuracy, see Nelson, 1984). Figure 1 presents mean gamma correlations by group for each time of JOL. The main effect of time of JOL was significant such that delayed JOLs ($M = 0.86$, $SD = 0.18$) were more accurate than immediate JOLs ($M = 0.33$, $SD = 0.34$), $F(1, 108) = 102.21$, $p = .000$. The overall main effect of ADHD status was not statistically significant, and relative accuracy was similar between groups for immediate JOLs. For delayed JOLs, the mean gamma correlation for the group with ADHD ($M = 0.90$, $SD = 0.13$) was significantly higher than that for the non-ADHD group ($M = 0.81$, $SD = 0.21$), $F(1, 54) = 4.34$, $p = .04$. This finding suggests that participants in the group with ADHD were not poorer than those in the non-ADHD group at predicting the relative probability of recalling items as reflected in their JOLs made at a delay. The main effect of time suggests that the delayed JOL effect operated similarly for ADHD and non-ADHD adults. That is, relative JOL accuracy improved for both groups when the judgments were made at a delay after study.

Table 2
Mean Discrepancy Scores for Judgments of Learning (JOLs) by Time of JOL and Group

	Group			
	ADHD		Non-ADHD	
Time of JOL	*M*	*SD*	*M*	*SD*
Immediate	.02	.28	–.02	.30
Delayed	–.05	.20	–.09	.21

Note: Positive discrepancy scores represent overestimates of performance. Negative scores reflect underestimates.

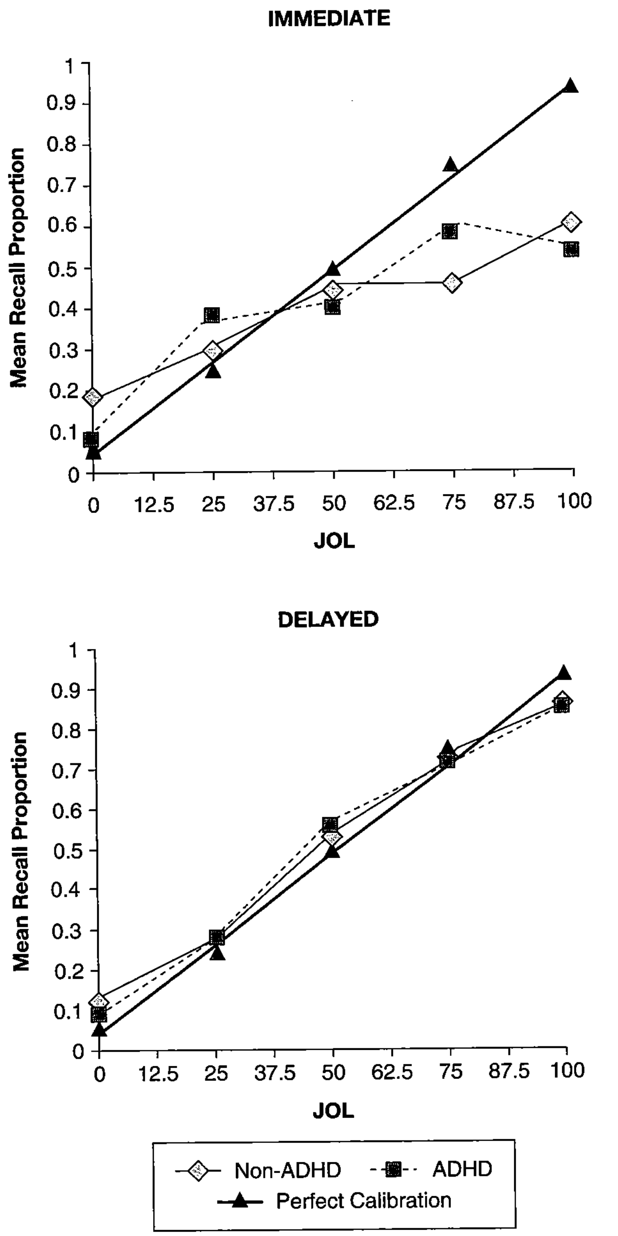
Retrospective confidence judgments. Gamma correlations between retrospective confidence judgments and recall outcome (correct vs. incorrect) were computed for each participant. These retrospective confidence judgments in the overall sample were highly accurate ($M = 0.97$, $SD = 0.04$) and did not differ between groups. Thus, compared to adults without ADHD, the adults with ADHD were equally accurate at identifying which responses were correct at the time of recall.

Absolute Accuracy

Judgments of learning. Table 2 contains the mean discrepancy scores by group. Discrepancy scores were calculated between each participant's mean JOL and his or her proportion of items recalled. A 2 (Group) × 2 (Time of JOL) mixed factor ANOVA revealed no significant main effect for group, $F(1, 108) = 0.76$, $p = .39$, or JOL type, $F(1, 108) = 2.42$, $p = .122$. The interaction between ADHD status and time of JOL was nonsignificant as well. Thus, no significant between-group differences in the absolute accuracy of JOL were evident.

Calibration curves depicted in Figure 2 provide an additional index of absolute JOL accuracy. Perfect calibration is depicted by the diagonal where the JOL value matches the actual level of recall performance. Visual inspection of these curves reveals that both groups were much better calibrated when making delayed JOLs than

Figure 2
Calibration Curves Depicting Absolute Accuracy of Judgments of Learning (JOL) by Group



Note: Curves for immediate JOL appear in the top panel, with curves for delayed JOL in the bottom panel. Items were sorted into five groups based on JOL received, and mean recall proportion across participants was calculated within each group of items. These groups consisted of JOL 0 to 12.5 ("0"), 12.5 to 37.5 ("25"), 37.5 to 62.5 ("50"), 62.5 to 87.5 ("75"), and 87.5 to 100 ("100").

immediate JOLs, as indicated by curves for delayed JOLs being closer to the diagonal (see the bottom panel of Figure 2). Participants in both groups showed the common hard-easy effect in that their immediate JOLs tended to be slightly underconfident for easier items and overconfident for more difficult ones. These calibration curves also demonstrate no substantive differences between the calibration of participants with and without ADHD.

Retrospective confidence judgments. Discrepancy scores reflecting absolute judgment accuracy were computed by subtracting each participant's recall proportion from his or her mean predicted proportion given by the retrospective confidence judgment. Judgments displayed high absolute accuracy (ADHD: $M = -0.01$, $SD = .05$; non-ADHD: $M = -0.04$, $SD = .16$) and did not differ significantly between groups, $F(1, 108) = 1.26$, $p = .27$.

Recall Performance and Judgment Magnitude

Mean level of recall. The mean proportions of words recalled by adults with and without ADHD are presented in Table 3. A 2 (Group) × 2 (JOL Type) mixed factor analysis of variance (ANOVA) did not reveal any significant main effects or interaction. The non-ADHD group and the group with ADHD recalled a similar proportion of items across those given immediate and delayed JOL ($M = 0.368$ vs. 0.391, respectively), $F(1, 108) = 0.169$, $p = .682$. In the overall sample, recall did not differ for items receiving immediate versus delayed JOLs.

Mean JOL ratings. Mean JOL ratings for adults with and without ADHD are reported in Table 3. In a 2 (Group) × 2 (JOL Type) mixed factor analysis of variance (ANOVA), there were no significant main effects or interaction. In this sample, adults with ADHD did not provide significantly overinflated JOLs compared to non-ADHD participants.

Mean retrospective confidence judgments. The retrospective judgments given by participants did not differ by group nor by the time that the item had received a JOL (Table 3). Adults with and without ADHD gave comparable confidence ratings in their recalled responses.

Discussion

A major goal of the present research was to investigate judgment accuracy for adults with ADHD. Similar to previous studies, we examined the degree to which ADHD produces overconfidence in future performance. Most important, we provide the first examination of the relative accuracy and calibration for their memory predictions. In this discussion, we briefly touch on the main outcomes relevant to absolute and relative accuracy in turn.

In contrast to prior studies that used discrepancy scores to measure absolute judgment accuracy in those

Table 3
Mean Recall and Judgment Magnitudes by Group and Time of Judgment

	Time							
	Immediate				Delayed			
	ADHD		Non-ADHD		ADHD		Non-ADHD	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Mean recall proportion	0.36	0.20	0.38	0.23	0.38	0.22	0.40	0.24
Mean JOL	38.38	19.56	36.16	19.82	32.22	22.6	30.85	21.89
Mean RCJ	34.33	18.88	36.40	21.54	37.92	20.57	38.88	23.16

Note: JOL = judgment of learning; RCJ = retrospective confidence judgment.

with ADHD, we did not obtain between-group differences in judgment accuracy on the paired-associate learning task as measured by discrepancy scores or by calibration curves. There are several possible explanations for this result. First, the vast majority of prior research on self-evaluation accuracy has been conducted with children with ADHD. Children with the disorder may display deficits in the accuracy of their domain- and task-specific judgments—specifically, they may tend to overestimate—whereas such tendencies may disappear by adulthood. Similar developmental trends appear in studies examining the judgments of nonaffected young children such that younger children tend to overestimate their performance to a greater extent than older children (Schneider, 1998). Thus, our sample may have "grown out of" their deficits in self-evaluation, leading to our inability to detect differences in absolute accuracy. Despite this possibility, at least one other study has identified significant overestimation in terms of absolute accuracy for a sample of adults with ADHD (Knouse et al., 2005).

A second possible explanation for the lack of group differences in our study relates to the way in which absolute accuracy is measured here and in past studies—the nature of the discrepancy score itself. In many studies (Hoza, Pelham, Milich, & Pillow, 1993; Ohan & Johnston, 2002; Owens & Hoza, 2003), the magnitude of the judgments given by children with and without ADHD is equivalent, whereas those with ADHD perform more poorly on the criterion test. If this specific pattern of results underlies previous findings of differences in absolute accuracy, then when groups are matched on test performance, differences in the amount of overestimation should disappear (cf., Connor et al., 1997). Fortuitously, our groups of adults obtained statistically equivalent levels of recall performance on this simple paired-associate learning task. In retrospect, this result was not entirely unexpected given findings in the ADHD and memory literature demonstrating intact basic memory processes with decrements in performance on more complex tasks that require additional processing (Barkley, DuPaul, & McMurray, 1990; Hervey et al., 2004). The present findings concerning absolute judgment accuracy raise the possibility that if children with and without ADHD were equated on levels of criterion test performance, the apparent differences in overestimation might disappear. This possibility should be investigated empirically.

Our findings on relative judgment accuracy contribute new information to the current discussion. As shown in Figure 1, both adults with and without ADHD showed the delayed JOL effect such that both groups benefited from making their predictive recall judgments at a delay. Most important, adults with ADHD were able to obtain high levels of predictive accuracy when making delayed JOLs. Moreover, the group with ADHD was highly accurate when judging the correctness of their test responses. These positive results may suggest possible mechanisms by which the self-regulated learning of adults with ADHD can be improved. Students with ADHD may benefit from strategies that use these accurate delayed judgments to regulate their learning and to enhance their scholarship. Such intervention strategies for individuals with ADHD are critical because they are likely to experience problems in academic settings (Barkley, Fischer, Smallish, & Fletcher, in press; Heiligenstein, Guenther, Levy, Savino, & Fulwiler, 1999).

One possibility is to train students with ADHD to explicitly judge their learning of newly studied materials by self-testing their memories, which is akin to making delayed JOLs. Outcomes from these self-tests could then be used to regulate learning by training the students to restudy those items that were incorrectly recalled. This kind of intervention has already shown promise for improving learning of older adults and adults with traumatic brain injury (for a review, see Dunlosky et al., 2005) and may also work for students with ADHD. Importantly, however, students with ADHD may require environmental support—such as computer programs—that help them make use of their accurate judgments. For example, a computer program could be devised that would record the delayed JOL given by a student with ADHD and would then present the lower JOL items more frequently for study.

Although the kinds of measurement procedures introduced here to explore the relation between ADHD and judgment will be widely applicable, the present findings should be considered in light of the limitations of the current study. First, the smaller size and recruitment source for this sample (college population and surrounding community) may limit the generalizability of these findings to the population of those with ADHD. Specifically, this sample may be more high functioning and well educated than the average adult with ADHD. Second, the current study relied on report of diagnosis and rating scale data for inclusion and did not include enough participants to assess the influence of ADHD subtype. Though we employed a multitiered approach, future studies should include definitions based on more rigorous criteria including diagnostic interview and should examine the effects of ADHD subtyping on judgment accuracy.

In summary, we believe the current research augments the literature and suggests new directions. From a methodological perspective, we introduce measures of accuracy from the larger realm of judgment and decision making—namely, relative accuracy and calibration—to the study of self-evaluative judgment accuracy in those with ADHD. In doing so, we present the first set of results on relative judgment accuracy in adults with ADHD on a memory task. These results suggest that, at least under specific circumstances, adult students with ADHD can make accurate judgments about their future memory. Thus, they may benefit from strategies that support the translation of their judgments into adaptive action.

Note

1. Stimulant medications prescribed to treat ADHD have been shown to improve performance on laboratory tasks, such as tapping vigilance, impulsive responding, short-term memory, reaction time, and problem solving (e.g., Tannock, Schachar, & Logan, 1995). With clearance from their prescribing physicians, participants taking daily stimulant medication or Atomoxetine (Straterra) were required to omit their morning dose prior to the procedure. Physicians were sent a cover letter describing the study, a signed Health Information Portability and Accountability Act compliant authorization for release from the patient, and a medical clearance form for them to sign and return. Following the session, participants continued their normal dosing schedule as prescribed by their physicians.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Barkley, R. A. (1998). *ADHD and the nature of self-control.* New York: Guilford.

Barkley, R. A., DuPaul, G. J., & McMurray, M. B. (1990). Comprehensive evaluation of attention deficit disorder with and without hyperactivity as defined by research criteria. *Journal of Consulting and Clinical Psychology, 58*, 775-789.

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2002). The persistence of Attention-Deficit/Hyperactivity Disorder into young adulthood as a function of reporting source and definition of disorder. *Journal of Abnormal Psychology, 111*, 279-289.

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (in press). Young adult outcome of hyperactive children: Educational, occupational, social, sexual, and financial functioning. *Journal of the American Academy of Child and Adolescent Psychiatry.*

Barkley, R. A., & Murphy, K. R. (1998). *Attention-Deficit Hyperactivity Disorder: A clinical workbook* (2nd ed.). New York: Guilford.

Barkley, R. A., Murphy, K. R., & Kwasnik, D. (1996). Psychological adjustment and adaptive impairments in young adults with ADHD. *Journal of Attention Disorders, 1*, 41-54.

Conners, C. K., Erhardt, D., & Sparrow, E. (1999). *Adult ADHD rating scales: Technical manual.* Toronto, Ontario, Canada: Multi-Health Systems.

Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging, 12*, 50-71.

Derogatis, L. R. (1975). *Symptom Checklist-90-R.* Bloomington, MN: Pearson Assessments.

Diener, M. B., & Milich, R. (1997). Effects of positive feedback on the social interactions of boys with attention deficit hyperactivity disorder: A test of the self-protective hypothesis. *Journal of Clinical Child Psychology, 26*, 256-265.

Dunlosky, J., Hertzog, C., Kennedy, M., & Thiede, K. (2005). The self-monitoring approach for effective learning. *Cognitive Technology, 10*, 4-11.

Dunlosky, J., Kubat-Silman, A. K., & Hertzog, C. (2003). Training monitoring skills improves older adults' self-paced associative learning. *Psychology and Aging, 18*, 340-345.

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD Rating Scale-IV: Checklists, norms, and clinical interpretation.* New York: Guilford.

DuPaul, G. J., Schaughency, E. A., Weyandt, L. L., Tripp, G., Kiesner, J., Ota, K., et al. (2001). Self-report of ADHD symptoms in university students: Cross-gender and cross-national prevalence. *Journal of Learning Disabilities, 34*, 370-379.

Gerdes, A. C., Hoza, B., & Pelham, W. (2003). Attention-deficit/hyperactivity disordered boys' relationships with their mothers and fathers: Child, mother, and father perceptions. *Development and Psychopathology, 15*, 363-382.

Heiligenstein, E., Conyers, L. M., Berns, A. R., & Smith, M. A. (1998). Preliminary normative data on DSM-IV attention deficit disorder in college students. *Journal of American College Health, 46*, 185-188.

Heiligenstein, E., Guenther, G., Levy, A., Savino, F., & Fulwiler, J. (1999). Psychological and academic functioning in college students with attention deficit hyperactivity disorder. *College Health, 47*, 181-185.

Hervey, A. S., Epstein, J. N., & Curry, J. F. (2004). Neuropsychology of adults with attention-deficit/hyperactivity disorder: A meta-analytic review. *Neuropsychology*, *18*, 495-503.

Hoza, B., Gerdes, A. C., Hinshaw, S. P., Arnold, L. E., Pelham, W. E., Molina, B., et al. (2004). Self-perceptions of competence in children with ADHD and comparison children. *Journal of Consulting and Clinical Psychology*, *72*, 382-391.

Hoza, B., Pelham, W. E., Dobbs, J., Owens, J. S., & Pillow, D. R. (2002). Do boys with attention-deficit/hyperactivity disorder have positive illusory self-concepts? *Journal of Abnormal Psychology*, *111*, 268-278.

Hoza, B., Pelham, W. E., Milich, R., & Pillow, D. R. (1993). The self-perceptions and attributions of attention deficit hyperactivity disordered and non-referred boys. *Journal of Abnormal Child Psychology*, *21*, 271-286.

Hoza, B., Waschbusch, D. A., Pelham, W. E., Molina, B. S. G., & Milich, R. (2000). Attention-deficit/hyperactivity disordered and control boys' responses to social success and failure. *Child Development*, *71*, 432-446.

Kennedy, M. R. T., & Yorkston, K. M. (2000). Accuracy of metamemory after traumatic brain injury: Predictions during verbal learning. *Journal of Speech, Language, and Hearing Research*, *43*, 1072-1086.

Knouse, L. E., Bagwell, C. L., Barkley, R. A., & Murphy, K. R. (2005). Accuracy of self-evaluation in adults with Attention-Deficit Hyperactivity Disorder: Evidence from a driving study. *Journal of Attention Disorders*, *8*, 221-234.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*, 349-370.

Koriat, A., & Schitzer-Reichert, R. (2002). Metacognitive judgments and their accuracy: Insights from the processes underlying judgments of learning in children. In P. Chambres, M. Izaute, & P. Marescaux (Eds.), *Metacognition: Process, function, and use* (pp. 1-18). Boston: Kluwer.

Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General*, *131*, 147-162.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance*, *20*, 159-183.

Moulin, C. J., Perfect, T. J., & Fitch, F. (2002). Metacognitive processes at encoding: Insights from Alzheimer's disease. In P. Chambres, M. Izaute, & P. Marescaux (Eds.), *Metacognition: Process, function, and use* (pp. 35-47). Boston: Kluwer.

Murphy, K. R., & Barkley, R. A. (1996). Prevalence of DSM-IV symptoms of ADHD in adult licensed drivers: Implications for clinical diagnosis. *Journal of Attention Disorders*, *1*, 147-161.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109-133.

Nelson, T. O. (1996). Gamma is a measure of the accuracy of predicting performance on one item relative to another item, not of the absolute performance on an individual item. *Applied Cognitive Psychology*, *10*, 257-260.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125-140). New York: Academic Press.

Nelson, T. O., Narens, L., & Dunlosky, J. (2004). A revised methodology for research on metamemory: Pre-judgment recall and monitoring. *Psychological Methods*, *9*, 53-69.

Nigg, J. T., Butler, K. M., Huang-Pollock, C. L., & Henderson, J. M. (2002). Inhibitory processes in adults with persistent childhood onset ADHD. *Journal of Consulting and Clinical Psychology*, *70*, 153-157.

Nigg, J. T., Stavro, G., Ettenhofer, M., Hambrick, D., Miller, T., & Henderson, J. M. (in press). Executive functions and ADHD in adults: Evidence for selective effects on ADHD symptom domains. *Journal of Abnormal Psychology*.

Ohan, J. L., & Johnston, C. (2002). Are the performance overestimates given by boys with ADHD self-protective? *Journal of Clinical Child Psychology*, *31*, 230-241.

O'Neill, M. E., & Douglas, V. I. (1991). Study strategies and story recall in attention deficit disorder and reading disability. *Journal of Abnormal Child Psychology*, *19*, 671-692.

Owens, J. S., & Hoza, B. (2003). The role of inattention and hyperactivity/impulsivity in the positive illusory bias. *Journal of Consulting and Clinical Psychology*, *71*, 680-691.

Ryan, M. P., Petty, C. R., & Wenzlaff, R. M. (1982). Motivated remembering efforts during tip-of-the-tongue states. *Acta Psychologica*, *51*, 137-147.

Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition, and wishful thinking. *Developmental Science*, *1*, 291-297.

Schwartz, B. (1994). Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic, Bulletin, & Review*, *1*, 357-375.

Shipley, W. C. (1946). *The Shipley Institute of Living Scale*. Los Angeles: Western Psychological Services.

Tannock, R., Schachar, R., & Logan, G. (1995). Methylphenidate and cognitive flexibility: Dissociated dose effects in hyperactive children. *Journal of Abnormal Child Psychology*, *23*, 235-266.

Weiss, G., & Hechtman, L. (1993). *Hyperactive children grow up*. New York: Guilford.

Whalen, C. K., Henker, B., Hinshaw, S. P., Heller, T., & Huber-Dressler, A. (1991). Messages of medication: Effects of actual versus informed medication status on hyperactive boys' expectancies and self-evaluations. *Journal of Consulting and Clinical Psychology*, *59*, 602-606.

Laura E. Knouse received her BS from the University of Richmond and her MA from the University of North Carolina at Greensboro (UNCG). She is currently a doctoral student in clinical psychology at UNCG. Her research interests include cognition and ADHD, with particular interest in metacognition and memory.

Matthew J. Paradise is an assistant professor of clinical psychology at the University of North Carolina Greensboro, who received his undergraduate degree in psychology from Stanford University and his PhD in child-clinical psychology from University of Washington. His publications often spotlight at-risk and underserved clinical populations, and his research interests focus broadly on adolescents and young adults who are at risk for psychopathology and look more specifically at the connection between risky behavior, mental health, and substance use across time. He is also concerned with the identification of biobehavioral precursors to maladaptive behavior, innovative strategies for intervention, and methodological and statistical challenges inherent in developmental research.

John Dunlosky received his PhD from the University of Washington in 1993 and then spent 2 years at the Georgia Institute of Technology. He is now an associate professor at Kent State University in the Cognitive Program. His research focuses on how individuals evaluate and control their ongoing learning of new materials, which is especially relevant to improving student scholarship. During the past decade, he has examined these metacognitive processes in many populations—most notably, older adults—and has demonstrated how adults of all ages can improve their learning through efficient control.