Time Estimation and Performance on Reproduction Tasks in Subtypes of Children With Attention Deficit Hyperactivity Disorder

José J. Bauermeister
University of Puerto Rico

Russell A. Barkley
Medical University of South Carolina

José V. Martínez, Eduardo Cumba, Rafael R. Ramírez, Graciela Reina, Maribel Matos, and Carmen C. Salas
University of Puerto Rico

This study compared Hispanic children (ages 7 to 11) with combined type (CT, n = 33) and inattentive type (IT, n = 21) attention deficit hyperactivity disorder (ADHD) and a control group (n = 25) on time-estimation and time-reproduction tasks. The ADHD groups showed larger errors in time reproduction but not in time estimation than the control group, and the groups did not differ from each other on their performance on this task. Individual differences could not be accounted for by oppositional-defiance ratings and low math or reading scores. Although various measures of executive functioning did not make significant unique contributions to time estimation performance, those of interference control and nonverbal working memory did so to the time-reproduction task. Findings suggest that ADHD is associated with a specific impairment in the capacity to reproduce rather than estimate time durations and that this may be related to the children’s deficits in inhibition and working memory.

Attention deficit hyperactivity disorder (ADHD) is characterized by developmentally inappropriate and impairing levels of inattention, hyperactive–impulsive behavior, or both that arise in childhood (American Psychiatric Association, 1994). Considerable debate exists as to whether those with ADHD who are chiefly inattentive (IT) and those who are both IT and hyperactive–impulsive (combined type or CT) represent different subtypes of ADHD or distinct disorders (Milich, Ballentine, & Lynam, 2001). Research findings seem to support the qualitative distinctiveness of these two types. The IT group, or at least a subgroup of it, is characterized by an attentional disturbance termed sluggish cognitive tempo (SCT) and passivity. This group is less likely to have comorbid disruptive disorders and externalizing behaviors and possibly has a greater risk for internalizing disorders. The CT, on the other hand, presents problems with resistance to distraction, persistence of effort, and response inhibition and is more likely to have comorbid disruptive or externalizing disorders (Barkley, 1997a; Milich et al., 2001).

One domain of cognitive functioning that has not been investigated for differences between these subtypes is the sense of time. Timing tasks are generally subdivided into those requiring retrospective or prospective responses. Both methods can be further subdivided into tasks requiring a verbal estimation or a motor reproduction of the sample interval. In retrospective paradigms, the participants are not told beforehand that they will be asked to estimate or reproduce a previously elapsed interval. Only after the interval has passed are they asked to render a judgment as to its duration. In this procedure, only one trial is likely to be given as it warns the participant to attend to the possibility that future intervals may be asked for estimation. It also seems to make few demands on inhibition, attention, and executive functioning, requiring instead primarily reconstruction of the interval from memory of the past event (Zakay, 1990). ADHD children are not expected to have difficulties on this task, given that deficits in memory storage and retrieval are not thought to be associated with the disorder. Consistent with this view, a previous study of ADHD children did
not find difficulties with this aspect of timing relative to a control group (Barkley, Koplowitz, Anderson, & McMurray, 1997).

Prospective timing methods forewarn the participants that an upcoming interval will require their attention and a response. The sample interval can be responded to either with verbal estimation, motor production, or motor reproduction methods (Zakay, 1990). In time-estimation tasks, participants verbally report how long a clearly demarcated time interval lasted. In time-production tasks, participants are verbally told the duration length and are then asked to produce the duration by some motor response, such as turning a light on and off. In the more complex time-reproduction task, participants are shown a duration by some nonverbal means, again such as turning a light bulb on and off, after which they must replicate that duration using the same or some other motor response. These are not simply different means of measuring the same timing construct. The reproduction task is believed to be more cognitively challenging because it demands a longer allocation of attention to temporal processing and makes heavier demands on working memory (holding the sample duration in mind to use in reproducing it; J. W. Brown, 1990; Mimura, Kinsbourne, & O’Connor, 2000; Zakay, 1990).

Previous studies (Gerbing, Ahadi, & Patton, 1987; Montare, 1977) have shown that poor motor inhibition is associated with sense of time deficits, particularly as assessed by time-reproduction paradigms. Less inhibited individuals make greater reproduction errors than more inhibited ones (Gerbing et al., 1987; Levine & Spivack, 1959; Levine, Spivack, Fuschillo, & Tavernier, 1959; Seigman, 1961). Given the impaired inhibition in the CT, the latter should be associated with impaired time reproduction. The CT may also impair time reproduction via the deficits that these children have in working memory (Barkley, 1997a). The nonverbal working memory may assist with timing behavior, as that is where sequences of events are held temporarily online for analysis and comparison to some metric (e.g., heart beat, respiration, a repetitive motor stereotypy, or an internal clock; J. W. Brown, 1990; Church, 1984; Michon & Jackson, 1984).

Timing may be further enhanced by verbal working memory (such as self-counting via self-speech). Research shows that working memory is important for accurately rendering temporal durations of at least 1 sec or more in duration (Mangels, Ivry, & Shimizu, 1998; Mimura et al., 2000; Rubia et al., 1998). In contrast, motor timing and time perception involving intervals less than 1 sec may be mediated more by the basal ganglia (Harrington, Haaland, & Hermanowicz, 1998; Rao et al., 1997) and cerebellum (Mangels et al., 1998). Judging durations longer than 20 to 30 sec also may need assistance from episodic, reference, or long-term memory because such durations may exceed the time frame of short-term or working memory (J. W. Brown, 1990; Kinsbourne & Hicks, 1990; Mimura et al., 2000; Nichelli, Venneri, Molinari, Tavanna, & Grafman, 1993). As applied to ADHD, this view suggests that both the poor motor inhibition and the impaired working memory associated with the disorder will interfere with timing behavior, particularly time reproduction, given its heavier demands on working memory.

Consistent with this rationale, children, teens, and adults with a clinical diagnosis of ADHD have been found to make significantly more errors of reproduction than control participants (Barkley, Edwards, Laneri, Fletcher, & Metevia, 2001; Barkley et al., 1997; Barkley, Murphy, & Bush, 2001). Group differences in time estimation, however, are not consistently found across studies (see Barkley, 1997a, for a review). Interestingly, no previous research has compared children with ADHD–IT and ADHD–CT in their sense of time. However, children with IT, especially those with SCT, have been found to present difficulties in focused attention, alertness, and information processing (Barkley, 1997a; McBurnett, Pfiffner, & Frick, 2001; Milich et al., 2001). This might mean that the IT group, especially the subset with SCT, would make qualitatively different errors in timing tasks than would the CT group.

For instance, the more perceptual or initial processing aspects of sensing time may be impaired in this IT group, but not impaired in the ADHD–CT group, as noted previously. If time estimation (the verbally reporting of an interval’s duration) can be taken as an index of such temporal processing, then the IT group might be expected to have a deficit in the accuracy of their estimations. This estimation deficit would result in deficits in time reproduction as well, because it is also dependent on the accurate perception and processing of the sample interval for it to then be replicated. Such a rationale is admittedly speculative in view of the limited information on the cognitive deficits found in the IT group and the dearth of prior research on timing behavior in this subtype.

This study, therefore, examined the extent to which children with ADHD–CT and ADHD–IT differ in their sense of time as measured by both time estimation and reproduction tasks. It also examined the degree to which performance on measures of executive functioning, particularly inhibition and working memory, were related to timing behavior in these subtypes. The study was conducted in a school sample of Hispanic (Puerto Rican) children diagnosed as ADHD to reduce potential referral biases often associated with clinical samples that may mask differences between the CT and IT groups.
Method

Participants

Participants were 79 children ages 7 to 11 years. These children were recruited from 26 elementary schools in the San Juan Metro area, Puerto Rico. They met the following core entry criteria: an IQ ≥ 80; the biological child of a Puerto Rican mother; have not lived in another country for more than 2 years; no evidence of significant sensory, language, neurological or pervasive developmental difficulties; and no history of treatment with stimulant or other psychotropic medication prior to 6 months of study participation. Thirty-three of the participants met Diagnostic and Statistical Manual of Mental Disorders (4th edition [DSM–IV]; American Psychiatric Association, 1994) diagnostic criteria for CT (22 boys and 11 girls), 21 met criteria for IT (15 boys and 6 girls), and 25 did not meet diagnostic criteria for the disorder (13 boys and 12 girls). The subset of a sample of 98 children ages 6 to 11 participated in a study that compared CT and IT groups on psychoeducational, behavioral, social, and clinical measures (Bauermeister et al., in press). For our study, 6-year-old participants were excluded. The time-estimation and time-reproduction tasks and one of the neurocognitive tests (Stroop color–word test [SCWT]) were administered only to children 7 years and older.

Procedure

A five-stage procedure was used to select participants. In the first stage, teachers were asked to identify children with attention problems, with and without high levels of activity–impulsivity, and control children. For each teacher, children were asked to complete the Distraction–Motivation (DM) and Activity–Impulsivity (AI) scales of the School Behavior Inventory (SBI; Bauermeister, 1994), which are separately normed for Puerto Rican boys and girls. A total of 1,041 children were initially identified. Those children with SBI scores above or below the established cutoff scores (see the following) passed this stage. Data on the number of participants screened at this and the following two stages are not available.

During the second stage, research assistants obtained informed consent from the mothers and administered a brief interview, the Disruptive Behavior Rating Scale (Barkley, 1998), and the Child Behavior Checklist (Achenbach, 1991a). Children were considered eligible if the screening interview and questionnaires indicated that they could meet the core entry criteria and presented inattention or hyperactivity–impulsivity behaviors at home.

At Stage 3, approximately 4 weeks after the initial screening, research assistants interviewed teachers of preselected children and administered the full version of the SBI and other questionnaires. Children who again met the DM and AI cutoff scores established in the first stage were given final assignments to three groups and scheduled for diagnostic assessment. The CT group (n = 37) consisted of children with a score greater than the 93rd percentile on both the DM and AI scales of the SBI. The IT group (n = 35) had a score greater than the 93rd percentile on the DM scale and below the 69th percentile on the AI scale. Children in the control group (n = 25) had scores below the 84th percentile on the DM and AI scales of the SBI, the total and broadband scales of the SBI, and the Teacher Report Form of the Child Behavior Checklist (Achenbach, 1991b); no parent or teacher complaints of significant behavior or academic problems; and no history of mental health services as reported by parents. The cutoffs selected were recommended by Barkley, DuPaul, and McMurray (1990) to empirically distinguish ADHD children in these two types.

During the fourth stage, the diagnostic assessment of the 97 eligible participants was made by three advanced clinical psychology students who were blind to the children’s prediagnostic group membership. Diagnoses were based on information obtained from a structured developmental interview and on Draft No. 4 of the Spanish-parent version of the Diagnostic Interview Schedule for Children (DISC–IV; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000). Each DISC–IV item was administered and scored in a structured manner. However, when necessary, the interviewers deviated from the structured format of the DISC–IV and clarified ambiguous responses through further questions and probes. Decisions as to the presence or absence of ADHD symptoms during the past 6 months were based on the information provided by mothers on the DISC–IV and the ADHD scale of the Disruptive Behavior Rating Scale, and on the information provided by teachers on the ADHD scale and SBI. As a general rule, when informants (parents, teachers) disagreed as to the presence of the number of inattentive or hyperactivity–impulsivity symptoms required for a diagnosis of CT or IT, more weight was given to teacher reports. Teachers were felt to be in better position to identify ADHD symptoms in the classroom and were thought to be generally more knowledgeable of behaviors to be expected in children within the age range studied. However, there were three instances in which hyperactivity–impulsivity symptoms were not reported by teachers but clearly and consistently reported by mothers at home, at school in the previous school year, and in other settings. In these instances, precedence was given to the information provided by mothers. All other ADHD diagnostic criteria and diagnoses (oppositional defiant disorder [ODD], conduct disorder, major depression, bipolar disorder, general-
ized anxiety disorder, and separation anxiety disorder) were based on information obtained from the DISC–IV. As a measure of quality control, 20% of the interviews were randomly selected and reviewed by different research assistants to ensure strict adherence to administration and recording procedures.

At the final stage, interviewers presented each case to two senior clinical child psychologists, who reviewed the information, asked questions, and clarified issues as needed. Following each case presentation, interviewers and clinicians independently provided DSM–IV diagnoses. The mean kappa (κ) coefficients between initial interviewers and clinicians were .96 (CT) and .93 (IT). The mean κ for the non-ADHD diagnoses ascertained was .99. The information collected for those cases for which diagnostic disagreement existed was further reviewed and discussed, and consensus diagnoses were reached. Of those classified by teachers during the third stage as inattentive and hyperactive, 81% received a diagnosis of ADHD–CT. On the other hand, of those classified as inattentive, 60% were diagnosed as ADHD–IT and 9% received an ADHD–CT diagnosis. The remaining participants of these two groups (n = 18) did not receive an ADHD diagnosis and were excluded. None of the children in the control group received a diagnosis of ADHD, but 4 received other DSM–IV diagnoses.

Each participant received a battery of behavioral and neurocognitive tests in a laboratory setting. The tests, including the time-estimation and time-reproduction tasks described later, were given in a standardized order. Research assistants administering these tests were blind to the children’s diagnostic status. The mothers and teachers were paid $10 and $15, respectively, for their participation.

Screening Measures

SBI (Bauermeister, 1994). This inventory includes narrow and broadband scales that were developed based on factor analyses of classroom behavior ratings of boys and girls. The AI and DM factors extracted are similar to the ones obtained for a previous version of the instrument (Bauermeister, Vargas, González, Colberg, & Carroll, 1987). The internal consistency coefficients (alpha) of the AI and DM, broadband (excessive control and insufficient control), and total scales for both genders ranged from .88 to .95, and the test–retest reliability from .58 (excessive control boys) to .89 over 4 to 6 weeks. Nine of the 12 AI items and 6 of the 9 DM items corresponded to DSM–IV hyperactivity–impulsivity and inattentive symptoms, respectively.

ADHD and ODD scales of the Disruptive Behavior Rating Scale (Barley, 1997b). The ADHD scale yields ratings of the nine DSM–IV hyperactivity–impulsivity and nine inattentive symptoms, respectively. The ODD scale yield ratings of the eight DSM–IV symptoms of that disorder. The internal consistency for the Spanish version of both scales completed by mothers and teachers ranged from .86 to .93; the test–retest reliability over a 4-week period analyzed with the intraclass correlation coefficient ranged from .78 to .89.

Child Behavior Checklist and Teacher Report Form (Achenbach, 1991a, 1991b). These checklists yield psychopathological symptoms along eight different dimensions, including attention, aggression, delinquency, anxiety–depression, social withdrawal, and so on. The scales have satisfactory reliability and validity. Alpha coefficients for Puerto Rican children ranged from .51 to .97 (Bauermeister et al., in press).

Draft No.4 of the Spanish DISC–IV (Shaffer et al., 2000). The DISC–IV is a structured instrument designed for the assessment of DSM–IV psychiatric and substance use disorders in children and adolescents. As previously pointed out, the DISC–IV was not used to generate diagnoses based on its computer algorithms, but to systematically survey the symptoms and criteria for the DSM–IV diagnoses assessed in the study.

The Wechsler Intelligence Scale for Children–Revised (Herrans & Rodríguez, 1992). This Spanish version of the Wechsler Intelligence Scale for Children–Revised has been adapted and normed for Puerto Rican children. The internal consistency for Total IQ is .94; the test–retest correlation over a 2-week interval is .90. The concurrent validity of the Wechsler Intelligence Scale for Children–Revised is supported by its pattern of correlations with other measures of intelligence (Herrans & Rodríguez, 1992).

Other Measures

Woodcock Psychoeducational Battery–Spanish (Woodcock, 1982). Tests used from this battery provide a measure of academic achievement in math and reading. The internal reliability for the latter measures was .92 and .88.

Sluggish Cognitive Tempo Scale. A mother version of this scale was constructed with Child Behavior Checklist items associated in the research literature with SCT symptoms (Bauermeister et al., in press). The items are: “confused or seems to be in a fog,” “daydreams or gets lost in his/her thoughts,” “stares blankly,” and “underactive, slow moving, or lacks energy.” Confirmatory factor analyses of parent ratings by our research group indicate that Sluggish Cognitive Tempo
Scale items fall in a dimension different from inattentive and hyperactive–impulsive dimensions. The alpha coefficient for the Sluggish Cognitive Tempo Scale in this sample was .82.

**SCWT (Golden, 1978).** Errors on this test are frequently found in children with ADHD (Pennington & Ozonoff, 1996). It contains three parts, each having a separate card containing 5 columns of 20 items and timed for 45 sec. Each participant is urged to read the items as rapidly as possible. In the first, the child is asked to read a list of color names (blue, red, and green). Then the participant is required to name patches of ink of the same color. In the third part (interference condition), the child is required to name the color of ink in which a word is presented while ignoring the word name. We used the color–word number of correct responses score (maximum score = 100), which is a measure of interference control or resistance to distraction (i.e., capacity to inhibit reading the word that is the dominant response).

**Hand Movement Scale (HMS; Kaufman & Kaufman, 1983).** This test evaluates the child’s ability to imitate progressively longer sequences of skilled hand movements. The score is the number of movement sequences successfully imitated. Scores are standardized with a mean of 10 and a standard deviation of 3. The HMS was used here to provide a measure of nonverbal working memory and motor sequencing. Previous studies have reported differences between children with and without ADHD on this measure (Barkley, Grodzinsky, & DuPaul, 1992; Mariani & Barkley, 1997).

**Simon Task (ST).** This is a commercially available toy that employs a plastic circular device housing four different colored keys on its top surface. When pressed, each of them emits a different tone. When activated, the game automatically presents a sequence of tones and lights up the key corresponding to those tones. To reproduce a melody, the child must press the keys in their correct sequence. With each trial, the sequence of tone and keys becomes increasingly longer and more complex. The participant’s score is the longest correctly reproduced sequence averaged across two consecutive trials (score range = 0 to 13). Mean scores on the ST administered to children ages 6 to 11 years 2 to 4 weeks apart were not significantly different ($p > .05$), and the test–retest correlation was .74 (Matos, 2000). This task appears to evaluate nonverbal (spatial) working memory in that adolescent scores on it load on a working memory factor (Barkley, Edwards, et al., 2001). However, it is also possible that participants can use the keys’ colors to memorize the sequence and thus employ some verbal cues to reproduce it. Adults with ADHD obtain lower scores than control groups (Murphy, Barkley, & Bush, 2001).

**Tower of Hanoi Test (TOHT).** The TOHT used in this study is the version described in detail by Pennington, Groisser, and Welsh (1993). It consists of two boards, one for the child and another for the examiner. Each board holds three tapered pegs and three (or four, in more difficult problems) plastic rings of graduated sizes that fit on the pegs. For each of the nine problems, the child must make the minimum numbers of moves to transfer his or her initial ring configuration into a duplicate of the examiner’s configuration. Three rules must be followed: a large ring cannot be placed on top of a smaller one, only one ring may be moved at a time, and a ring always must be moved onto a peg (not put on the table). To pass the problem, the child must achieve two consecutive correct solutions in a maximum of six trials required to solve problem types of differing difficulty. A maximum of 6 points was assigned to a problem solved in the first two trials, 5 points for solution on Trials 2 and 3, 4 points for solution on Trials 3 and 4, 3 points for solution on Trials 4 and 5, and 2 points for solution on Trials 5 and 6. The range of possible scores is 0 to 54 (Pennington et al., 1993). The TOHT is often used as a measure of planning and the ability to hold in mind the established rule to guide behavior (working memory) to solve each problem. Research shows that deficits on this task often exist in children with ADHD (Pennington & Ozonoff, 1996). Scores on the TOHT administered 2 to 4 weeks apart to children 6 to 11 years were not significantly different ($p > .05$) and correlated .69 (Matos, 2000).

**Dependent Measures**

**Time-estimation task.** Children were told they were going to play a game with a flashlight and that the examiner would turn on her lantern, turn it off, and then ask how long, in seconds, the child thought that the flashlight was on. After demonstrating how the game was played, the examiner administered a practice trial with a time duration of 5 sec. Two trials for each of six different time intervals were then presented in the same order. In the first sequence, the intervals were presented in the following order: 6, 13, 25, 10, 33, and 18 sec; in the second, the order was 33, 6, 18, 10, 13, and 25 sec. The sequence of time duration was randomly determined. During the task administration, the examiner stood at a distance of 3 feet from the participant. The light of the lantern was always directed over the child’s right shoulder. An inaudible stopwatch, whose face was not visible to the child, was used to present the time intervals.

**Time-reproduction task.** The test administration procedure and the 12 time intervals described previ-
ously were used in this task, but in a different randomly determined order (first sequence: 13, 18, 10, 25, 33, and 6 sec; second sequence: 33, 13, 6, 18, 25, and 10 sec). However, the participant was told that in this game he or she would have a flashlight to reproduce the time interval presented by the examiner. The child’s flashlight was placed on a table in front to avoid the participant’s playing with it and becoming distracted while the examiner presented the time interval for that trial. A practice trial with a time duration of 5 sec was also administered. The reproduction task always followed the time-estimation task. The child’s score on each trial was the duration of the response. Following the procedures of Barkley, Edwards, et al. (2001) and Barkley, Murphy, et al. (2001), raw scores were converted to an absolute discrepancy score by subtracting the sample duration from that estimated or reproduced by the participant and eliminating positive or negative signs (S. W. Brown, 1985). These scores measured the magnitude of the child’s errors in timing in each trial.

### Results

**Sociodemographic, Psychoeducational and Clinical Characteristics**

The three groups were compared on relevant variables using chi-square test (gender and diagnoses) and analyses of variance. Given the relatively large number of omnibus tests under each construct, a Bonferroni correction was applied to control for Type I error rate. Scheffé tests were then used to analyze potential pairwise group differences for continuous variables. As can be seen in Table 1, the three groups did not differ significantly on gender, age, parents’ education, mother’s age, or number of siblings. Children in the CT and IT groups were rated with similar levels of inattention but significantly more so than the comparison group. On the other hand, children in the CT group were rated as more hyperactive—impulsive than children in the IT and comparison groups, as expected from the selection

<table>
<thead>
<tr>
<th>Measure</th>
<th>Groups</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
<th>$\chi^2$/$F$</th>
<th>Contrasts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (% male)</td>
<td>ADHD–CT (1)</td>
<td>63.60</td>
<td>71.24</td>
<td>52.00</td>
<td>1.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td>8.42</td>
<td>9.10</td>
<td>1.09</td>
<td>1.53</td>
<td>1.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother’s age (years)</td>
<td></td>
<td>33.94</td>
<td>35.05</td>
<td>6.17</td>
<td>1.37</td>
<td>7.03</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parents’ education (years)</td>
<td></td>
<td>12.73</td>
<td>13.33</td>
<td>1.46</td>
<td>1.20</td>
<td>2.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># siblings</td>
<td></td>
<td>1.45</td>
<td>1.52</td>
<td>1.03</td>
<td>1.24</td>
<td>.78</td>
<td>.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ADHD Scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADHD Inattention–M</td>
<td></td>
<td>17.09</td>
<td>17.76</td>
<td>4.99</td>
<td>1.20</td>
<td>2.86</td>
<td>82.49***</td>
<td>1.2&lt;3</td>
<td></td>
</tr>
<tr>
<td>ADHD Inattention–T</td>
<td></td>
<td>20.70</td>
<td>19.05</td>
<td>6.11</td>
<td>4.05</td>
<td>8.87</td>
<td>9.73***</td>
<td>1.2&lt;3</td>
<td></td>
</tr>
<tr>
<td>ADHD Hyp.–Imp.–M</td>
<td></td>
<td>18.12</td>
<td>9.86</td>
<td>6.31</td>
<td>1.31</td>
<td>3.70</td>
<td>63.60***</td>
<td>1&lt;2&lt;3</td>
<td></td>
</tr>
<tr>
<td>ADHD Hyp.–Imp.–T</td>
<td></td>
<td>20.30</td>
<td>3.48</td>
<td>4.58</td>
<td>1.36</td>
<td>2.12</td>
<td>142.30***</td>
<td>1&lt;2,3</td>
<td></td>
</tr>
<tr>
<td><strong>Psychoeducational Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total IQ</td>
<td></td>
<td>101.9</td>
<td>97.10</td>
<td>10.21</td>
<td>1.10</td>
<td>10.79</td>
<td>4.59**</td>
<td>1&lt;3</td>
<td></td>
</tr>
<tr>
<td>WBP–S Reading</td>
<td></td>
<td>96.58</td>
<td>94.76</td>
<td>13.83</td>
<td>1.10</td>
<td>1.60</td>
<td>13.50***</td>
<td>1&lt;2&lt;3</td>
<td></td>
</tr>
<tr>
<td>WBP–S Math</td>
<td></td>
<td>89.27</td>
<td>84.48</td>
<td>12.08</td>
<td>1.09</td>
<td>8.18</td>
<td>18.47***</td>
<td>1&lt;2&lt;3</td>
<td></td>
</tr>
<tr>
<td><strong>Clinical variables and comorbidities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCTS–M</td>
<td></td>
<td>2.21</td>
<td>3.50</td>
<td>2.69</td>
<td>1.17</td>
<td>.48</td>
<td>15.21***</td>
<td>1.2&lt;3</td>
<td></td>
</tr>
<tr>
<td>CBCL Anx–Dep–M</td>
<td></td>
<td>60.24</td>
<td>59.05</td>
<td>10.16</td>
<td>5.45</td>
<td>5.05</td>
<td>3.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ODD Scale–M</td>
<td></td>
<td>11.00</td>
<td>8.38</td>
<td>5.84</td>
<td>1.36</td>
<td>2.93</td>
<td>18.62***</td>
<td>1&lt;2&lt;3</td>
<td></td>
</tr>
<tr>
<td>% ODD</td>
<td></td>
<td>45.50</td>
<td>19.00</td>
<td>12.00</td>
<td>1.18</td>
<td>1.24</td>
<td>9.03**</td>
<td>1&lt;2,3</td>
<td></td>
</tr>
<tr>
<td>% Any anxiety disorder</td>
<td></td>
<td>21.20</td>
<td>23.80</td>
<td>8.00</td>
<td>1.36</td>
<td>1.51</td>
<td>2.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Any depressive disorder</td>
<td></td>
<td>6.10</td>
<td>4.80</td>
<td></td>
<td></td>
<td></td>
<td>1.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Neurocognitive Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tower of Hanoi Test</td>
<td></td>
<td>20.30</td>
<td>24.00</td>
<td>10.68</td>
<td>3.07</td>
<td>9.98</td>
<td>6.49**</td>
<td>1&lt;3</td>
<td></td>
</tr>
<tr>
<td>Stroop Color–Word Test</td>
<td></td>
<td>20.21</td>
<td>19.14</td>
<td>4.75</td>
<td>25.24</td>
<td>7.07</td>
<td>7.30***</td>
<td>1&lt;2&lt;3</td>
<td></td>
</tr>
<tr>
<td>Simon Task</td>
<td></td>
<td>3.77</td>
<td>4.14</td>
<td>2.09</td>
<td>1.10</td>
<td>1.56</td>
<td>1.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hands Movement Scale</td>
<td></td>
<td>10.39</td>
<td>10.90</td>
<td>2.76</td>
<td>1.30</td>
<td>2.29</td>
<td>7.20***</td>
<td>1&lt;2&lt;3</td>
<td></td>
</tr>
</tbody>
</table>

*Note: ADHD = attention deficit hyperactivity disorder; CT = combined type; IT = inattentive type; M = mother; T = teacher; Hyp.–Imp. = hyperactivity–impulsivity; SCTS = Sluggish Cognitive Tempo Scale; WPS = Woodcock Psychoeducational Battery; CBCL = Child Behavior Checklist; Anx.–Dep. = anxious–depressed; ODD = oppositional defiant disorder; CD = conduct disorder. Scores for CBCL Anx.–Dep.–M and Hands Movement values are standardized scores. A Bonferroni adjustment was applied to omnibus analyses. Scheffé tests were used to analyze pairwise group differences.

*Degrees of freedom for chi-square analysis of gender and diagnoses = 2, N = 79; degrees of freedom for the remaining analyses of variance = 2, N = 71 to 76 depending on missing values.

**$p < .01$, ***$p < .001$.}

156
criteria for these groups. Further inspection of the data (not presented in Table 1) indicated that all children in the IT group had teacher ratings on the DM and AI scales of the SBI that were at least 1 SD apart. In terms of their psychoeducational characteristics, the groups were within the normal range of intelligence. The IT group had significantly lower total IQ scores than the comparison group. Both ADHD groups obtained significantly lower reading and math achievement scores than the control group.

As to their clinical profile, both ADHD groups presented significantly higher ratings of SCT. The CT and IT groups received significantly higher mother ratings of ODD behaviors. Psychiatric comorbidity for the clinical groups was generally low. None of the participants received a diagnosis of conduct disorder. Significant group differences were found only for ODD, where about half of the children in the CT group received this diagnosis.

### Time Estimation and Time Reproduction

Absolute discrepancy scores for the time-estimation and time-reproduction tasks were computed for each of the six time durations averaged across the two trials given for each duration. The descriptive statistics for these raw scores are available on request. Due to significant skewness on the distribution of the measures, the standard deviation exceeded the mean score for many measures when evaluated on the original scale metric. Therefore, a logarithmic transformation was applied to reduce the level of skewness and non-normality and transform the original scale values into a distribution more consistent with the assumptions of parametric statistical analyses. The mean and standard deviations for these transformed scores for the time tasks are summarized in Table 2. To examine the relation between absolute discrepancy time estimation and reproduction errors, we calculated an omnibus mean score for each measure by averaging across all of the durations. Bivariate analyses indicated a significant Pearson correlation between the time-estimation and time-reproduction measures ($r = .32, p \leq .01$). This finding indicated a significant but relatively weak association between these two different tasks.

Measures were analyzed with a 3 (groups) × 6 (duration) repeated-measures analyses of variance. Each cell of our design was tested for deviations from normality by the Shapiro–Wilk test, and for most cells there was no evidence of departures from normality. The sphericity assumption necessary for correct interpretation of $F$ values was rejected. Thus, the assumption of compound symmetry was not met, implying that equal variances and covariances between cells in our design was not a realistic assumption. Therefore, our repeated-measures analyses of variance was re-estimated with software capable of estimating exact $p$ values without making the assumption of homoscedastic error variances (XPro, 2001). This software is the first commercially available program capable of exact parametric methods (Weerahandi, 1994). Probability values produced by this program are exact and are not based on asymptotic approximations that are most likely inaccurate in small samples.

#### Time-estimation task

The main effect for group was not significant (Table 2). The main effect for dura-

### Table 2. Means and Standard Deviations for Transformed Log Absolute Discrepancy Scores for Time Estimation and Reproduction Tasks and Results for ANOVAs

<table>
<thead>
<tr>
<th>Time Durations</th>
<th>Groups</th>
<th>CT ($n = 33$)</th>
<th>IT ($n = 21$)</th>
<th>Control ($n = 25$)</th>
<th>ANOVA $F^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Time estimation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>2.80</td>
<td>1.00</td>
<td>2.75</td>
<td>.76</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>2.94</td>
<td>1.21</td>
<td>2.99</td>
<td>1.08</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>3.06</td>
<td>.98</td>
<td>3.31</td>
<td>.87</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>3.09</td>
<td>.98</td>
<td>3.44</td>
<td>.83</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>3.33</td>
<td>.94</td>
<td>3.63</td>
<td>.67</td>
</tr>
<tr>
<td>33</td>
<td></td>
<td>3.67</td>
<td>.77</td>
<td>3.62</td>
<td>.63</td>
</tr>
<tr>
<td>Time reproduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>2.16</td>
<td>1.12</td>
<td>1.80</td>
<td>1.7</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>2.61</td>
<td>.97</td>
<td>2.10</td>
<td>.79</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>2.48</td>
<td>1.23</td>
<td>2.22</td>
<td>1.06</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>2.65</td>
<td>.76</td>
<td>2.47</td>
<td>.92</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>2.97</td>
<td>.89</td>
<td>2.81</td>
<td>.63</td>
</tr>
<tr>
<td>33</td>
<td></td>
<td>3.18</td>
<td>.90</td>
<td>3.18</td>
<td>.64</td>
</tr>
</tbody>
</table>

**Note:** ANOVA = analysis of variance; CT = combined type; IT = inattentive type.

*a*df = 2, 76

***p < .001.
tion was significant. A general trend was observed in which the longer the durations to be estimated, the larger the absolute errors made by all children. Scheffé post hoc comparisons indicated that the magnitude of the time-estimation errors for the 6 sec duration was significantly smaller from the 18 sec, 25 sec, and 33 sec durations ($p \leq .05$). Also, the magnitude of the errors for the 10 sec and 13 sec durations was not significantly different from each other but was smaller than the errors for the 33 sec duration. The Group × Duration interaction was not significant.

**Time-reproduction task.** The main group effect was significant (Table 2). Additional comparisons with the Scheffé procedure indicated that the CT ($p \leq .001$) and IT ($p \leq .01$) groups made larger absolute discrepancy errors than the control group. The two ADHD groups did not differ significantly from each other. The main effect for duration was also significant. In general, as duration increased, errors tended to increase in magnitude. Scheffé post hoc comparisons indicated that the magnitude of the errors for the 6 sec, 10 sec, and 13 sec durations was not significantly different from each other but was smaller in magnitude from errors for the 25 sec and 33 sec duration ($p \leq .05$). The magnitude of the errors for the 6 sec duration was significantly smaller in comparison to the 18 sec duration, which in turn was smaller in comparison to the 33 sec duration ($p \leq .05$). The Group × Duration interaction was not significant.

**Contribution of Variables That Distinguished the Groups on Time-Reproduction Scores**

As indicated in Table 1, the CT, IT, or both groups obtained lower total IQ, reading, and math scores and higher mother ratings on the ODD and SCT scale. The CT group presented also a higher rate of comorbid ODD diagnosis than the other two groups. To examine the potential impact of these measures on time-reproduction group differences, we first computed bivariate Pearson correlations between these variables and the omnibus mean score for absolute time-reproduction measures for all participants. Total IQ and SCT scores were not significantly correlated with time reproduction and were thus excluded from further analyses.

Math scores ($r = -.33$), reading scores ($r = -.24$), and ratings of ODD symptoms ($r = .29$) were significantly correlated ($p \leq .01$) with time-reproduction errors. Also, math ($r = -.56$) and reading ($r = -.57$) scores were significantly correlated ($p \leq .001$) with mother ratings of inattention. Furthermore, hyperactivity–impulsivity ratings were significantly associated ($p < .001$) with both math ($r = -.36$) and reading ($r = -.39$). ODD symptoms were also significantly correlated ($p \leq .001$) with inattention ($r = .63$) and hyperactivity–impulsivity ($r = .69$) ratings. These relations with the core symptoms of ADHD suggest that the use of math, reading, and ODD scores as covariates in the analyses of the dependent variables is questionable. Doing so would remove variance on the time-reproduction absolute discrepancy measures possibly directly due to ADHD. Such use of math, reading, and ODD scores as covariates could also violate an assumption of that analysis (independence of the covariate and the independent variable; Miller & Chapman, 2001). Because of these issues, we did not control for these relevant variables using analysis of covariance. Instead, we proceeded to use multiple linear regression methods to determine if reading, math, and ODD ratings accounted for differences on time reproduction beyond that attributable to ADHD symptoms. To do so we entered these three scores and the ADHD inattention and hyperactivity–impulsivity scores simultaneously into the regression model. For these analyses, the dependent variable was the omnibus mean absolute discrepancy score averaged across all durations.

To examine multicollinearity between ADHD ratings and the comorbid variables, we computed correlations between these variables. All correlations were significant. Hyperactivity–impulsivity demonstrated overlapping variance with inattention ($r^2 = .51$), ODD severity ($r^2 = .48$), reading ($r^2 = .15$), and math ($r^2 = .13$). Inattention demonstrated overlapping variance with ODD severity ($r^2 = .40$), reading ($r^2 = .32$), and math ($r^2 = .31$). Therefore, the remaining nonoverlapping variance for each measure was still substantial enough to be worth examining for its contribution apart of its overlap with ADHD. Because reading and math scores were highly correlated ($r = .72$), we proceeded to estimate two regression models to predict absolute discrepancy reproduction scores using hyperactivity–impulsivity, inattention, ODD severity, and either math or reading scores. All regression models were estimated with XPro (2001) software that provided increased precision in the estimation of $p$-values in small samples. The regression model for math, $F(4, 74) = 6.12, R^2 = .25$, and for reading, $F(4, 74) = 6.05, R^2 = .25$, were significant ($p \leq .001$). The ADHD inattention score was a significant predictor ($p \leq .01$) of higher absolute discrepancy time-reproduction errors in both models. The other variables in these models were not significant predictors of time reproduction.

**Contribution of Executive Function Measures to Time-Estimation and Time-Reproduction Absolute Discrepancy Scores**

As can be seen in Table 1, the CT and IT groups presented similar performance on the SCWT and HMS but significantly lower than the comparison group. On the TOHT, only the CT obtained a poorer performance
than the control group. The three groups did not differ on the ST.

It has been hypothesized that executive functions, particularly nonverbal working memory, play a significant role in time reproduction (Barkley, 1997a). This hypothesis is consistent with the pattern of negative correlations obtained between the absolute discrepancy omnibus mean scores of time reproduction and time estimation and the executive-function measures. As judged by the magnitude of Pearson correlations, time-reproduction scores had a stronger relation with these executive-function tasks (SCWT $r = -0.56$, ST $r = -0.47$, HMS $r = -0.32$, TOHT $r = -0.41$, $p \leq 0.01$) than time-estimation scores (SCWT $r = -0.27$, ST $r = -0.24$, HMS $r = -0.27$, TOHT $r = -0.18$, $p \leq 0.05$, except for TOHT). In view of these findings, we used multiple regression analyses to examine the contribution of scores on the TOHT, SCWT, ST, and HMS to the participants’ performance on the omnibus mean time-estimation and time-reproduction tasks. All these models were estimated with XPro (2001) software. For the time-estimation task, the full regression model was significant, $F(4, 66) = 2.59, p \leq 0.05, R^2 = .14$. None of the predictor variables, however, made significant unique contributions to time-estimation performance. The regression model for the time-reproduction task was also significant, $F(4, 67) = 11.57, p \leq .001, R^2 = .41$. Scores on the SCWT ($\beta = -0.34, t = -3.68, p \leq .001$) and ST ($\beta = -0.09, t = -2.16, p \leq 0.05$) were the only variables that made significant unique contributions in this model, after controlling for all the other variables in the equation.

### Discussion

This study attempted to replicate past studies by demonstrating that ADHD is associated with a deficit in time sense, particularly in the capacity to accurately reproduce time intervals. We attempted to extend earlier research in several important respects. First, we evaluated both time estimation and reproduction in children with and without this disorder. This approach permitted us to determine if children with ADHD present a specific impairment only in the capacity to reproduce rather than estimate time durations, as research with adolescents and adults with ADHD has suggested. Second, we focused on the question of whether the two most common subtypes of ADHD (CT vs. IT) differed from each other and from a control group in their time estimations and reproductions. Third, we examined the degree to which other variables that are associated with ADHD may be responsible for any deficits observed on the timing measures. Finally, we evaluated the contribution of several measures of executive functioning to performance on these timing measures.

In keeping with a growing number of studies, we documented that those with ADHD make larger reproduction errors than the control group and that these errors increase in magnitude as the sample duration increases. This study went further than prior ones, however, in finding that this deficit is not evident on a measure of time estimation but is substantial on a measure of time reproduction. A similar set of findings was obtained in prior studies on ADHD teens (Barkley, Edwards, et al., 2001) and adults (Barkley, Murphy, et al., 2001). This suggests that the disparity between time estimation (unimpaired) and time reproduction (impaired) exists across a wide age span of people with ADHD. These results might indicate that the difficulty with timing in those with ADHD may not be one of perception of time or the accurate detection of durations (at least as indexed by verbal estimation), but in holding the time duration in working memory so as to guide reproduction via a motor response (as indexed by time reproduction). Though an intriguing idea, time-estimation tasks may index more than just the accurate detection of time (such as familiarity with verbal counting strategies to measure time durations), leaving this interpretation open to question.

It could be argued that the reproduction task is longer than the estimation task and that this may have led to our finding group differences on the latter task simply due to a possible delay aversion in the ADHD group (Sonuga-Barke, 1996). Although possible, our results do not seem consistent with that idea. The ADHD participants made larger absolute errors regardless of direction. Also, our regression analyses showed a unique contribution of both inhibition (interference control on the SCWT) and working memory (ST) to performance on this task. Neither of these constructs plays any role in the delay-aversion explanation of ADHD (Sonuga-Barke, 1996), but they do in neuro-psychological (executive functioning) theories of the disorder (Barkley, 1997a).

Our findings shed some light on the view that certain executive functions contribute to performance on the time-reproduction task. The executive-function tasks accounted for substantially more variance (41%) in time-reproduction errors than time estimation (14%), with two of these tasks making significant contributions: the Stroop task (interference part) and the ST. The former task is widely regarded as assessing the inhibitory function of interference control (preventing disruption of performance by prepotent competing responses), whereas the ST has been interpreted previously as largely evaluating nonverbal working memory (retention of tone or gesture sequences; Barkley, Murphy, et al., 2001). These results imply that both inhibition and working memory make significant contributions to the accurate performance of time reproductions and are consistent with the results of earlier studies using brain-injured and normal samples (Mimura et al., 2000).
This study goes further than previous research, however, in demonstrating that this deficit is found in ADHD children of Hispanic (Puerto Rican) ancestry. Noteworthy here is also the fact that these samples were school-identified rather than clinically referred. Such children may be less severe in their disorder and have lower rates of comorbidity than do clinically referred children having ADHD. Most recent studies on timing problems in ADHD children have employed clinic-referred samples (Barkley, Edwards, et al., 2001; Barkley et al., 1997). This study implies that the timing problems associated with ADHD are not just evident in participants of European/North American ancestry or in the more severe clinical segments of children with the disorder.

Our investigation also sought to extend previous research in this area by studying possible subtype differences in ADHD individuals’ timing behavior. Children with the CT form of ADHD were compared here to those with the IT form of the disorder. Neither subtype showed deficits in time estimation, whereas both demonstrated problems with time reproduction. One obvious interpretation of the relative lack of subtype differences in our results is that both subtypes share the same timing problem and are not qualitatively different from each other in this cognitive capacity. But it is also possible that the manner in which the IT group was formed resulted in a heterogeneous group. Research suggests that the IT group can be further subdivided into a subset (approximately 30% to 50%) with SCT. It is this subset that shows qualitative differences from the CT group that have led some to argue for its incorporation into a separate and distinct disorder (Milich et al., 2001). Therefore, subtype differences on timing and other cognitive abilities might depend on the proportion of the IT group having elevated SCT symptoms. Our results suggest that this might be so, given that inattention and hyperactivity, but not SCT symptoms, were significantly correlated with the executive-function tasks and time-reproduction errors.

ADHD is often associated with other psychiatric disorders, such as ODD, conduct disorder, and internalizing disorders, as well as with learning disabilities and low-average intelligence. We found, however, that the most common psychiatric conditions seen in conjunction with ADHD in this age group (ODD), as well as cognitive problems (math and reading), did not make unique contributions to the timing problems we found in the participants, once inattention and hyperactivity–impulsivity were also taken into account. Our results agree with two earlier studies that found no significant contribution of comorbid disorders or achievement to the timing problems found in ADHD teens (Barkley, Edwards, et al., 2001) and adults (Barkley, Murphy, et al., 2001). They are also consistent with a larger literature that finds (a) little or no association of comorbid conditions with the larger arena of executive-functioning deficits consistently seen in children with ADHD and (b) that these executive-function deficits often remain even after the questionable practice of controlling for level of IQ (Klorman et al., 1999; Mariani & Barkley, 1997; Murphy et al., 2001; Nigg, 1999; Nigg, Hinshaw, Carte, & Treuting, 1998).

This study has several limitations that should be kept in mind when interpreting its results. Our findings apply to children identified through a teacher-screening process rather than those children with ADHD who may be clinically referred. Although this approach has the advantage of avoiding treatment referral biases in the sampling process, it constrains the generalization of our findings. The fact that participants under stimulant treatment were excluded is not likely to affect our findings. The prevalence of stimulant medication treatment for Puerto Rican children with ADHD in the general population was 3.6% at the time of the diagnostic interview (Bauermeister et al., 2003). Another limitation may have arisen in the procedures for administering the timing intervals. This process involved a manual presentation of the sample durations by the examiner. This procedure is less precise both in presenting the sample duration and in timing the child’s reproductions than had the task been computer-administered. It is possible that the ADHD subtypes differ more subtly in their timing problems, a difference that could emerge had a more precise timing task been utilized.

An additional potential limitation is the small size of the IT group, which reduced the power of the analyses to detect potential group differences and to examine gender effects on the timing measures. Also, our measure of behavioral inhibition was limited to the Stroop interference score. Different results might have been evident had other measures of inhibition been employed. Finally, order effects may have influenced the performance of the participants on the time tasks, given that they were administered in the same order. Results for the time-reproduction task may have been less apparent had the time-estimation task not been given earlier in the testing sequence. However, our time-reproduction findings are similar to the ones reported in a previous study in which time estimation did not precede time reproduction (Barkley et al., 1997).

Bearing these limitations in mind, this study adds to a growing literature that demonstrates difficulties with time reproduction (but not estimation) in ADHD. It further extends these findings to both the IT and CT subtypes of the disorder and replicates and extends these results to a sample of Hispanic children with ADHD. This study also found that these timing inaccuracies were related, in part, to deficits in inhibition (interference control) and working memory. Such findings are consistent with earlier neuropsychological research on the cognitive mechanisms of timing behavior.
REFERENCES


Received June 25, 2003
Accepted August 26, 2004