Journal of Child Psychology and Psychiatry 46:4 (2005), pp 420–435 doi: 10.1111/j.1469-7610.2004.00362.x

Interactions of task and subject variables among continuous performance tests

Colin B. Denney,[1] Mark D. Rapport,[2] and Kyong-Mee Chung[3]

[1]Department of Psychological Sciences, Purdue University, USA; [2]Department of Psychology, University of Central Florida, USA; [3]Department of Psychology, University of Hawaii, USA

Background: Contemporary models of working memory suggest that target paradigm (TP) and target density (TD) should interact as influences on error rates derived from continuous performance tests (CPTs). The present study evaluated this hypothesis empirically in a typically developing, ethnically diverse sample of children. The extent to which scores based on different combinations of these task parameters showed different patterns of relationship to age, intelligence, and gender was also assessed. **Methods:** Four continuous performance tests were derived by combining two target paradigms (AX and repeated letter target stimuli) with two levels of target density (8.3% and 33%). Variations in mean omission (OE) and commission (CE) error rates were examined within and across combinations of TP and TD. In addition, a nested series of structural equation models was utilized to examine patterns of relationship among error rates, age, intelligence, and gender. **Results:** Target paradigm and target density interacted as influences on error rates. Increasing density resulted in higher OE and CE rates for the AX paradigm. In contrast, the high density condition yielded a decline in OE rates accompanied by a small increase in CEs using the repeated letter CPT. Target paradigms were also distinguishable on the basis of age when using OEs as the performance measure, whereas combinations of age and intelligence distinguished between density levels but not target paradigms using CEs as the dependent measure. **Conclusions:** Different combinations of target paradigm and target density appear to yield scores that are conceptually and psychometrically distinguishable. Consequently, developmentally appropriate interpretation of error rates across tasks may require (a) careful analysis of working memory and attentional resources required for successful performance, and (b) normative data bases that are differently stratified with respect to combinations of age and intelligence. **Keywords:** Continuous performance test, sustained attention. **Abbreviations:** CPT: Continuous Performance Test; AX: CPT based on A-Then-X target sequence; RL: CPT based on repeated letter targets; OE: omission error rate (i.e., complement of hit rate); CE: commission error rate (i.e., false alarm rate).

The Continuous Performance Test (CPT) is a popular laboratory paradigm used to measure vigilance, or the maintenance of attention for infrequent but critical events over time (Davies & Parasuraman, 1982). CPTs have found application across a wide spectrum of scientific and clinical arenas, including child (Banaschewski et al., 2003; Barkley, Grodzinsky, & DuPaul, 1992; Halperin et al., 1993; Lindsay, Tomazic, Levine, & Accardo, 2001; Okazaki, Maekawa, & Futakami, 2001; Pascualvaca et al., 1997) and adult (Elvevag, Weinberger, & Suter, 2000; Epstein, Conners, Sitarenios, & Erhardt, 1998; Liu et al., 2002) psychopathology, and teratology (Eghbalieh, Crinella, Hunt, & Swanson, 2000; Fried & Watkinson, 2001; Grandjean et al., 2001; Till, Koren, & Rovet, 2001; Walkowiak et al., 1998). A recent review suggests that it is the most frequently used and cited measure of attention in research and clinical literatures (Riccio, Reynolds, Lowe, & Moore, 2002).

Despite the burgeoning application of CPTs, published literature pertaining to their use with both normal and clinical child populations is weakened by theoretical and methodological considerations. Most notably, a variety of CPT paradigms differing widely in design have been utilized as if the responses they evoke can be interpreted independently of task conditions (Losier, McGrath, & Klein, 1996). Refined theoretical models of attention (Lovett, Reder, & Lebiere, 1999) and substantial empirical data call this practice into question (Corkum & Siegel, 1993). Furthermore, the possibility that norms used in the interpretation of CPT scores derived from different tasks may have to be stratified along different dimensions has received little empirical scrutiny. The following investigation examines these issues in detail.

Interactions of task parameters

Task parameters in CPTs may interact in complex ways that limit the generality of scores (Corkum & Siegel, 1993; J. P. Leung, Leung, & Tang, 2000; P. W. Leung & Luk, 1988; Rose, Murphy, Schickedantz, & Tucci, 2001). For example, some investigations find that the impact of display time on performance differs in a non-linear manner across event rates (Chee, Logan, Schachar, Lindsay, & Wachsmuth, 1989). Other investigations demonstrate that the effect of event rate depends on target probability (Rose et al., 2001). Findings of this nature suggest that performance on CPTs defined by different combinations of

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

task parameters may depend upon separable mechanisms or different facets of attention (Borgaro et al., 2003).

The type of signal embedded in the stimulus stream to which subjects are asked to respond illustrates this principle. Two approaches are of particular interest for purposes of the present study. The term 'target paradigm (TP)' will be employed to describe the type of stimulus or stimulus sequence that children are to detect during a CPT. The AX target paradigm (AX) is frequently employed in both research and clinical settings. It entails emission of a simple motor response following the sequential presentation of the letters 'A' and 'X'. A less frequently cited task that we will refer to as a repeated letter (RL) paradigm requires a response whenever the same letter occurs twice in succession. Specific features of these target paradigms and mechanisms through which they may affect performance are examined below.

Effects of target stimuli on working memory and attentional load

Working memory and attention. The present investigation is predicated on Cowan's (Cowan, 1988, 1997, 1999) embedded process model of working memory. This view conceptualizes WM as a subset of long-term declarative memory representations that are in a highly accessible state. Declarative memory representations refer to discrete units of information (i.e., facts) such as names of letters or numbers. Accessibility refers to the speed with which a unit of information can be retrieved for purposes of problem solving. A highly accessible representation is described as being in an activated state. Activation levels rise in response to environmental or cognitive events, but decline exponentially over time unless maintained through active rehearsal processes. Activation is also distributed across working memory representations, so accessibility falls in proportion to the number of representations being maintained. The number of representations active at a given point in time will be referred to as working memory load.

The focus of attention refers to the subset of working memory representations within conscious awareness. Thus, the focus of attention is narrower in conceptual scope than working memory itself. The central executive refers to a set of cognitive operations that are presumed to govern the interaction of attentional focus with working memory and long-term store (Baddeley, 1997; Baddeley & Logie, 1999). This construct is narrower in scope but similar in meaning to the broad concept of executive functioning in clinical neuropsychology (Morris, 1996).

Target paradigms and working memory load. The AX CPT imposes minimal working memory load. Specifically, only two declarative memory representations must be persistently maintained over the course of the task, those of the letters 'A' and 'X'. Representations of other letters that are not part of the target pattern need only be activated and maintained transiently during specific decisions involving them. In contrast, the repeated letter model requires activation of working memory representations of all the alphabetic letters since any letter can occur as part of either a target or non-target pattern.

Target stimuli and attentional demand. AX and repeated letter CPTs are also distinguishable in terms of their demands on attentional focus. In the AX task, focus of attention initially subsumes only the working memory representation of 'A'. When a stimulus letter is displayed, attentional focus expands to include the features of the letter for purposes of comparison to the characteristics of 'A' represented in working memory. If the presented letter is not an 'A' no shift of attention is required. Conversely, if the stimulus is an 'A' attentional focus must shift to the working memory representation of 'X', but only until the subsequent letter is evaluated. At that point, attention must shift back to the representation of 'A' regardless of whether the presented letter was an 'X' or not. Thus, the proportion of stimuli comprised of 'A's controls the level of demand for shifts of attentional focus. Furthermore, the proportion of stimuli comprised of 'A's can be manipulated through increases or decreases of (a) target density (TD: the proportion of presented stimuli comprised of targets), (b) the frequency of 'A's not followed by 'X's, or (c) combinations of both.

In contrast, in the repeated letter model no shift of attentional focus is required on presentation of the second member of an identical pair (i.e., a target sequence) since the letter's representation enters the focus of attention when it is presented the first time. In contrast, when two non-identical letters occur in sequence attention must shift from the first to the second member of the pair in order to judge whether the subsequent stimulus matches the second member of the previous pair. Thus, shifts of attentional focus are only required following the sequential presentation of two non-identical letters. Consequently, the frequency of demand for shifts of attentional focus should decline as target density increases.

Summary. The foregoing formulation implies that the target stimuli incorporated into a CPT determine working memory load and conditions under which attentional focus must shift. More specifically, it suggests that increasing target density should amplify attentional demands for the AX CPT but reduce them for repeated letter tasks. This begs the question of whether scores generalize across combinations of target paradigm and target density. Furthermore, this raises the possibility that norms for CPTs based on different combinations of target paradigm and density must be stratified in different ways. We now examine this psychometric issue.

Subject attributes as dimensions of psychometric analysis

The dimensions along which norms are stratified constitute a fundamental facet of an instrument's psychometric profile because individual differences in the skills targeted for measurement may be confounded by one or more broader characteristics such as age, intelligence, or gender. Control for such confounds ensures that variations in performance can be unambiguously interpreted in terms of the target construct. This implies that thorough psychometric evaluation of CPTs varying in design must incorporate detailed scrutiny of relationships among performance indices and subject attributes such as age, intelligence, and gender to ensure that scores from different tasks are interpreted within developmentally and conceptually appropriate frames of reference.

The relationship of subject characteristics to children's performance on continuous performance tests has not been analyzed systematically to date, despite its implications for methodological controls in research and normative interpretation of scores. Research pertaining to relationships among CPT scores, age, intelligence, and gender in children is reviewed below to clarify the nature and scope of the conceptual and psychometric concerns at stake. Studies involving adolescents and adults are selectively cited to illustrate or amplify issues of particular salience.

Development and CPT performance

Empirical data have consistently shown that older children make fewer errors of both omission and commission than their younger peers. Furthermore, age-related improvement in performance among children can vary across tasks based on different target stimuli (Lin, Hsiao, & Chen, 1999). At this juncture, however, relationships to age have been primarily examined within single CPTs or across two levels of a single variable. Conversely, no investigations have compared the magnitude of age effects across combinations of multiple task parameters. Such analyses are imperative to establish the extent to which age-standardized scores have equivalent conceptual and psychometric meaning across different CPTs. This issue can be addressed most efficiently through analysis of relationships between performance indices and age across experimentally varied combinations of task variables.

Intelligence and CPT performance

Relationships between intelligence and CPT scores have not received systematic or comprehensive scrutiny. Findings vary widely across investigations. Some unequivocally fail to find significant relationships between IQ and CPT performance (Porrino et al., 1983; Schachar, Logan, Wachsmuth, & Chajczdy, 1988), whereas others find clear evidence of significant, albeit moderate correlations between IQ scores and error rates (Aylward, Gordon, & Verhulst, 1997; Seidel & Joschko, 1990) or IQ scores and signal detection indices (Swanson & Cooney, 1989). Others (Pascualvaca et al., 1997) report data suggesting that the relationship between intelligence and vigilance is stronger for boys than girls.

Collectively, these data suggest that IQ may be related to CPT performance under some conditions but not others, and more strongly for boys than girls. This raises the question of what variables control the relationship of intelligence to CPT scores, and hence, the circumstances under which interpretation of CPT scores must take individual differences in intellectual functioning into account. Analysis of relationships among performance indices and IQ scores across tasks defined through experimental manipulation of task parameters offers an efficient means of examining this issue.

Gender and CPT performance

Recent investigations of CPT scores obtained by adolescents and young adults suggest that the likelihood of gender differences on CPTs is controlled by the visual complexity of stimuli (Dittmar, Warm, Dember, & Ricks, 1993; Prinzel & Freeman, 1997). A recent study suggests that task complexity may moderate the likelihood of gender differences on auditory CPTs among children (Hatta, 1993). In contrast, findings have been inconsistent when alphanumeric stimuli have been employed. Some have found no gender differences (Horn, Wagner, & Ialongo, 1989; Loge, Staton, & Beatty, 1990), whereas others report statistically non-significant or untested trends suggestive of gender-related variation (Matier-Sharma, Perachio, Newcorn, Sharma, & Halperin, 1995). No investigations have examined gender differences using multiple tasks based on alphanumeric stimuli but differing systematically with respect to multiple other parameters.

Collectively, these data do not provide a strong foundation for anticipating gender differences in performance using CPTs based on alphanumeric stimuli. The variability in findings, however, raises the possibility that the likelihood of gender effects may be controlled by interactions of task properties that differ idiosyncratically across studies without being explicitly characterized. As such, exploratory examination of gender effects across different CPTs administered within single samples may prove worthwhile.

Purposes of study

The model of working memory described by Cowan (Cowan, 1999) suggests that increasing target density should increase the frequency with which attentional focus must shift when the AX paradigm is employed, but decrease it when a repeated letter CPT is used. This implies that (a) target paradigm and target density should interact as influences on error rates, and (b) normative data bases used for interpretation of performance within or across combinations of these task variables may require different patterns of stratification with respect to subject attributes such as age, gender, and intelligence. These issues are investigated empirically in the present investigation.

Method

Participants

The sample was comprised of children attending elementary school in Honolulu (Oahu), Hawaii. Approximately 68% of the State's population and 90% of the population of Oahu reside in the city and county of Honolulu (United States Bureau of the Census, 1990). Two schools were selected for recruitment based on available data suggesting their ethnic and demographic composition closely resembled that of the broader population of children residing in Hawaii (Ifuku, 1996). One of these is a publicly funded research arm associated with the University of Hawaii. Its primary mission is to develop and test curricula suitable for children of differing abilities and backgrounds. Children are admitted to the school based on ethnicity, gender, parent socioeconomic and marital status, residence location, and academic achievement to approximate the State's census.

A private school was also selected for participation to obtain a sample reflecting the relatively large number of children attending private schools in the State (Ifuku, 1996). The school admits students from throughout the State, although the majority of children reside in the urban Honolulu area.

An informational letter accompanied by consent and demographic information forms were mailed to parents of children attending both schools. The letter provided a basic description of the research project. The latter two forms were used to obtain written consent for children's participation and demographic information concerning family members, respectively. Parental consent was obtained for 100% and 54% of the children attending the University-affiliated public school and private school, respectively. The obtained consent rate compares favorably with that reported in other studies based on school samples (Kearney, Hopkins, Mauss, & Weisheit, 1983). Complete data sets were obtained for all children whose parents granted consent.

Children were categorized with respect to ethnicity on the basis of parental report. The sample was comprised of the following groups: East Asian (36%), Part-Hawaiian (23%), Caucasian (11%), Southeast Asian (4%), Pacific Islander (<1%), and Mixed (25%). Subjects were considered 'Part-Hawaiian' if their ethnic background included any Hawaiian ancestry. Subjects were considered 'Mixed' if they could not be unambiguously assigned to one of the foregoing categories. None of the children were of African-American descent.

Instruments

Kaufman Brief Intelligence Test (K-BIT). The Kaufman Brief Intelligence Test (K-BIT) consists of two subtests (Vocabulary and Matrices) that combine to yield a composite IQ that was used to provide an estimate of children's intelligence. The Vocabulary task was developed to represent the construct of Crystallized Intelligence, whereas the Matrices subtest was intended to reflect Fluid Intelligence. The psychometric properties of the K-BIT are well established and detailed in the instrument's manual (Kaufman & Kaufman, 1990).

The distribution of composite intelligence test scores across combinations of gender and age are summarized in Table 1. These data suggest that the sample included children with above-average IQ scores (i.e., >115), especially in the 8–10-year age range. Consequently, all data analyses were cross-validated using subsamples of children with composite IQ scores in the average (i.e., 90–114) and above average (i.e., >114) ranges.

Child Behavior Checklist (CBCL). The CBCL is a widely used rating scale designed to measure multiple dimensions of child behavior. The properties of this instrument have been thoroughly evaluated and reported (Achenbach, 1991; Biederman et al., 1993; Chen, Faraone, Biederman, & Tsuang, 1994). The CBCL was utilized to screen children for features suggesting need for more comprehensive clinical assessment. Children's mean scores on all scales fell in the normal range within and across combinations of age and gender. Thirteen children (3.7% of the sample) were identified as having clinically elevated scores. Nine (2.6%) of these showed difficulties with disruptive behavior. Four (1.1%) had elevated scores on scales assessing anxious or depressive features, social withdrawal, or somatic complaints. Elevations across both domains were evident in one case. These children were

Table 1 Distribution of intelligence by gender and age

	Boys		Girls	
Age	n	IQ (sd)	n	IQ (sd)
7	16	109.19 (12.00)	26	108.85 (8.943)
8	23	114.13 (10.46)	25	106.08 (9.032)
9	17	113.47 (11.41)	27	112.03 (11.85)
10	18	119.22 (10.70)	22	116.82 (10.20)
11	25	111.20 (12.90)	33	111.75 (11.86)
12	13	107.92 (7.783)	11	107.82 (15.11)
13	24	108.13 (11.64)	22	110.14 (9.498)
14	17	107.82 (11.28)	15	107.80 (13.74)
15	9	97.22 (9.897)	9	98.89 (15.10)

Note: Age rounded to nearest whole year; IQ = Mean composite IQ score based on the Kaufman Brief Intelligence Test; sd = sample standard deviation for the composite IQ score; Normative mean for the composite IQ = 100; Normative standard deviation for the composite IQ = 15.

not univariate or multivariate outliers, so all were included in analyses.

Continuous performance tests (CPT)

Target paradigms. Two distinct target paradigms were chosen for inclusion in the current study based on a comprehensive literature supporting their conceptual linkage to different modes of information processing (Sergeant & van der Meere, 1990). The first was an AX model that required a response whenever the letter 'A' was immediately followed by the letter 'X'. The second was a repeated letter version that required a response whenever any letter occurred twice in immediate succession.

Target density. Each of the two target paradigms was combined with two levels of target density (8.3% and 33.3% density, respectively). This variable was manipulated as a means of controlling the frequency with which shifts of attentional focus were required. In the AX paradigm the number of required attentional shifts increases as the proportion of target sequences increases. In sharp contrast, shifts of attention are required following the presentations of non-targets in the RL task. Consequently, decreases in target density are expected to increase the level of this demand and thereby hamper performance.

Shared task parameters. All task parameters other than target paradigm and target density were identical across conditions. Capitalized alphabetic letters (A–Z) measuring 3.5 cm in both height and width were utilized as stimuli. Five hundred and forty letters were presented on the center of a computer monitor at a rate of one stimulus per second with a display time of .2 s and an interstimulus interval of .8 s. Each task was performed without breaks and required nine minutes to complete.[1] Targets comprised 45 (8.3%) and 180 (33.3%) of the 540 stimuli in low and high density conditions, respectively. Children were instructed to press a button on a trackball whenever targets occurred and to withhold responses to nontargets.

Prior to administration of the experimental tasks children were required to: (a) identify letters of the alphabet to ensure letter recognition, and (b) participate in 1-min practice sessions with each instrument until a criterion of 80% correct target identification was met. Different, randomly determined sequences of target and non-target were used during each successive testing session. Children were seated such that the computer monitor was approximately .5 m away, and stimuli were centered around a horizontal line between children's eyes and the center point of the computer screen. Variation in children's heights precluded precise estimation or standardization of the visual angle subtended by stimuli. Every effort was made to ensure that the visual angle was reasonably comparable across children. An experimenter was present throughout all testing, situated approximately 3 m behind the child.

[1] In the AX paradigm, each 3-min block also contained 10 'A's not followed by 'X's and 10 'X's not preceded by an 'A'.

Dependent variables

Rates of omission errors (OE) and commission errors (CE) summed across 3-minute blocks of time were used as dependent variables in this investigation. Each experimental condition yielded three such scores. Raw OE and CE rates were normalized[2] across conditions to reduce the influence of a small number of outliers. OEs and CEs were transformed separately. Normalized scores were then placed on a *t*-score scale (i.e., mean = 50, sd = 10) to facilitate analysis and interpretation. Scores pooled across blocks within combinations of target paradigm and target density were utilized for statistical analyses. These variables were chosen for study due to the frequency with which they are reported in studies of CPT performance in children. Responses had to be emitted during the interstimulus interval. Consequently, some delayed response may have been classified as commission errors. Response time data were unavailable so this possibility could not be evaluated. Decrement scores (i.e., the difference between final and initial error rates) were not employed due to their typically poor reliability[3] as indices of individual differences in performance (Linn & Slinde, 1977).

Perceptual sensitivity (d') scores as defined by signal detection theory (SDT) were highly correlated with omission error rates (range .95–.98, median = .96) because the OE rate was substantially higher than the CE rate. Observed error rates and the magnitude of correlation between OEs and d' were highly comparable to those reported in other investigations of CPT performance among typically developing samples (Conners, Epstein, Angold, & Klaric, 2003; Lin et al., 1999). Beta score reliability was poor as is frequently observed in clinical research using CPTs (Cornblatt, Lenzenweger, & Erlenmeyer-Kimling, 1989). This performance index is beset with psychometric limitations (Balakrishnan, 1998; See, Warm, Dember, & Howe, 1997) and was not subjected to analysis in the present study.

Procedures

Trained graduate students assessed children's performance on the computerized CPTs once per week over a 2-week period at the Children's Learning Clinic. These instruments were administered as part of a larger battery of tests (e.g., intelligence testing, short-term memory) that required the child's presence for approximately 1.5 hours per session. Breaks were scheduled between tests to minimize fatigue. Each child was administered a total of four CPTs (AXL: AX-low density; AXH: AX-high density; RLL: repeated letter-low density; and RLH: repeated letter-high density) across the two testing sessions (two each session, one week apart) in counterbalanced order. A minimum of 45 min ensued between CPTs.

[2] Normalization refers to a procedure whereby scores are ranked and then assigned the standard score corresponding to that rank in a normal distribution (Wilkinson, Blank, & Gruber, 1996).

[3] Reliability estimates for decrement scores ranged from .03 to .13 for OEs and .09 to .11 for CEs.

Evaluation of model fit

Hypotheses regarding relationships among subject attributes and CPT performance were evaluated using a nested series of structural equation models. Model fit was evaluated using both absolute and incremental indices. Absolute indices evaluate the proportion of variation among observed scores accounted for by a latent variable model, whereas incremental indices reflect the comparative fit of a proposed model to a baseline standard. Specific metrics were chosen on the basis of recommendations published in the modeling literature (Bentler, 1990; Kline, 1998; Maruyama, 1998).

Absolute indices

Goodness of Fit Index (GFI). The GFI indicates the proportion of covariance among observed variables accounted for by relationships embodied in a model (Kline, 1998). Values range from 0 to 1.0. A value of 1.0 indicates perfect fit. Values >.90 are indicative of adequate fit.

Root Mean Square Error of Approximation (RMSEA). The traditional chi-square statistic used in model testing is weakened by excessive sensitivity to sample size (i.e., it tends to lead to model rejection even given good fit if sample size is moderate to large). The RMSEA is a transformation of the chi-square statistic that corrects for this tendency (Browne & Cudeck, 1993). This index also includes an adjustment for the number of variables incorporated into a model such that more parsimonious models are favored over less parsimonious ones. RMSEA values falling below .05 indicate adequate fit (Kline, 1998). Departure of this index from the desired range can be evaluated for statistical significance (i.e., the magnitude of difference can be evaluated against the range of variation observed under chance conditions).

Incremental indices

Chi-square difference test (chi-square diff). This index tests the null hypothesis that a model of interest fits as well as a pre-defined baseline standard (Kline, 1998). This test can only be utilized for models which are nested under the model used for comparison. Two models are said to be in a nested relationship if one can be derived from the other by modifying values of regression weights (Kline, 1998).

Comparative Fit Index (CFI). The CFI expresses the proportional improvement in the overall fit of a proposed model relative to a baseline standard (Kline, 1998). For example, a CFI value of .98 indicates that the fit of a proposed model is 98% better than that of a pre-defined baseline standard estimated with the same sample data. Its values are constrained to fall between 0 and 1.0. Typically, the baseline standard assumes that all variables in the proposed model are uncorrelated with each other. Other null models, however, can be specified to provide more stringent evaluative standards.

Results

Interactions of target paradigm and target density

Mean error rates. Normalized OE and CE rates based on 3-minute blocks for each experimental condition are shown in Figure 1. Each data point is banded by a 99% confidence interval. Thus, any pair of points whose confidence intervals do not overlap are significantly different at an alpha level of .01. Raw error rates corresponding to each data point are expressed as percentages to facilitate comparisons with previous studies. Three-way (Target Paradigm (2) × Target Density (2) × Block (3)) analysis of variance (ANOVA) yielded significant two-way interactions of target paradigm and density for both OEs ($F(1{,}351) = 335.02$, $p < .000001$) and CEs ($F(1{,}351) = 197.57$, $p < .000001$). The high density condition yielded a mean increase in omission error rate of approximately 3% over the low density condition for the AX CPT. In contrast, an average decrease in OE rate of nearly 10% was observed for the repeated letter task at high density compared to low density. Increased density was associated with increased commission error rates for both AX and repeated letter tasks, but the magnitude of the effect was nearly twice as large for the former as for the latter. Thus, as anticipated, increasing target density exerted markedly different influences on performance across AX and repeated letter tasks.

Mean vigilance decrements.[4] The foregoing three-way ANOVA also revealed significant interactions of target paradigm, target density, and block for both OEs ($F(1{,}351) = 20.55$, $p < .000001$) and CEs ($F(1{,}351) = 7.54$, $p < .0006$). These finding indicate that mean performance decrements varied systematically as joint functions of target paradigm and target density.

Comparisons of initial and final mean scores (i.e., scores observed at three and nine minutes, respectively, see Figure 1) illustrate the nature of these interactions. Omission error rates increased significantly over time for all conditions except for the combination of the AX paradigm at low target density. Moreover, whereas OE trends were linear and comparatively shallow in slope for the two AX CPTs, quadratic trends emerged for the repeated letter tasks. Detailed analysis of the repeated letter tasks showed that OE rates increased more rapidly at low than at high density over the initial six minutes of testing ($F(1{,}351) = 27.17$, $p < .00001$), but no further

[4] Mean decrements are not the same as decrement scores. The latter term refers to a difference score computed at the individual level and which is used to assess variation in rates of performance decline across persons. Mean decrements refer to performance declines averaged over a group of individuals. Decrement scores are rarely reliable as indices of individual variation but mean decrements are typically useful as descriptors of grouped data.

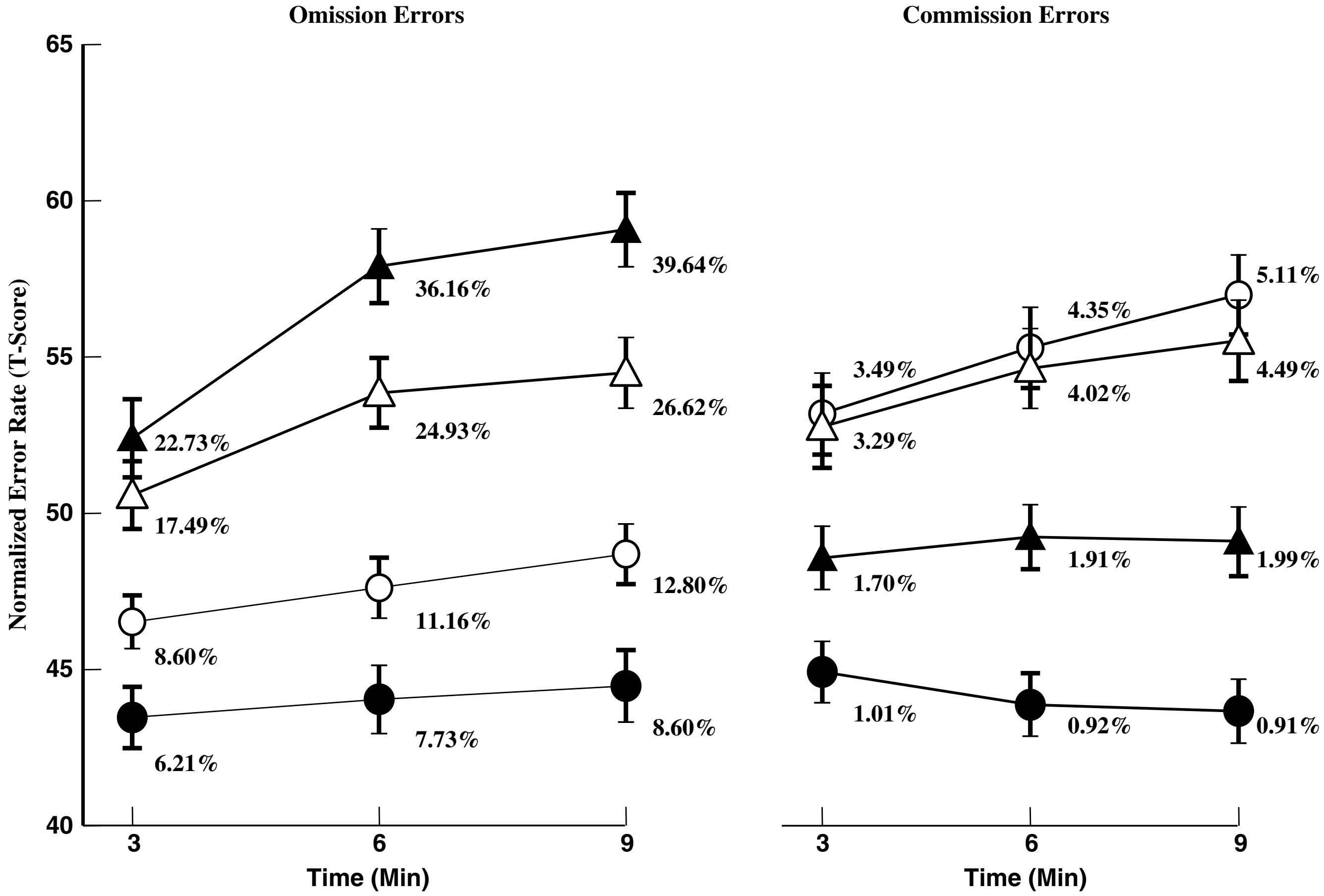


Figure 1 Mean rates of normalized omission and commission errors. Closed circles = AX target paradigm, low density condition; Open circles = AX target paradigm, high density condition; Closed triangles = repeated letter target paradigm, low density condition; Open triangles = repeated letter target paradigm, high density condition; Omission errors = failures to respond to presented targets; Commission errors = responses to non-targets; Error rates are scaled as normalized *t*-scores (mean = 50, sd = 10); Percentages refer to raw percent error rates corresponding to normalized *t*-scores; Error bars represent 99% confidence intervals around mean normalized scores; Time = time of measurement expressed in minutes

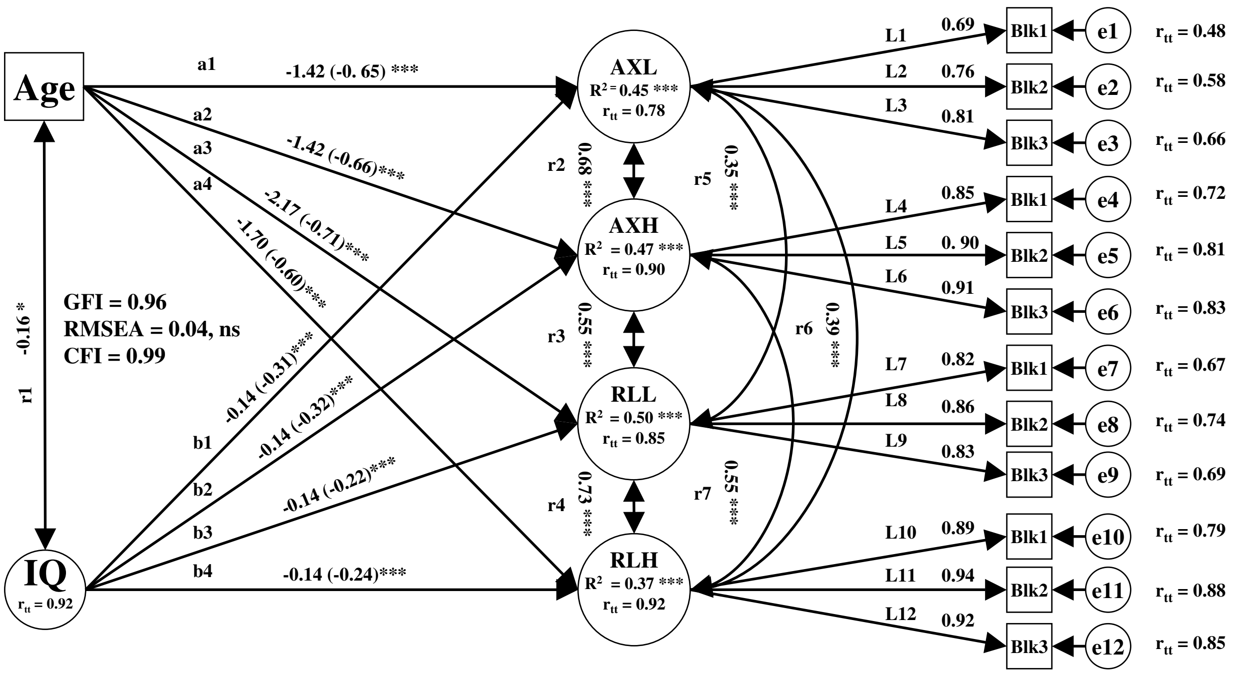
changes in performance were observed thereafter for either of these CPTs ($F(1{,}351) = 1.58$, *ns*).

The three-way interaction of target paradigm, target density, and block with respect to commission errors was largely explained by a two-way interaction of density with block ($F(2{,}702) = 48.80$, $p < 00001$). Specifically, CE rates increased substantially under high density conditions for both AX and repeated letter paradigms. Commission error rates remained constant across time for the RL instrument under the low density condition, whereas a non-significant decrease emerged for the corresponding AX task. The latter effect was just large enough to render the three-way interaction statistically significant, but was so small in magnitude as to be conceptually meaningless.

The functional distinctions among tasks illustrated by the preceding analyses beg the question of whether they also show different patterns of relationship to subject attributes relevant to construction of norms. This issue is examined in the subsequent series of analyses.

Relationships among subject attributes and continuous performance tests

Overview of structural equation models. Models used to analyze omission and commission errors are illustrated in Figures 2 and 3, respectively. Observed scores based on 3-minute blocks of time within combinations of target paradigms and densities are shown as squares in the figures (see squares labeled Blk1–Blk3 in Figures 2 and 3). Error variances associated with these scores are shown as circles labeled e1–e12. Latent variables representing scores pooled across blocks within conditions are shown as large circles (see circles labeled AXL, AXH, RLL, and RLH in Figures 2 and 3). Factor loadings of observed scores on the latent variables they define are shown as single-headed arrows in the figures (see arrows L1–L12 in Figures 2 and 3). Age expressed in whole years is shown as a square in the upper left corners of Figures 2 and 3 (see square labeled 'Age'). Intelligence is represented as a latent variable defined by the composite score from the KBIT (see circle labeled



Age Model	Chi-Sq Diff (df)	CFI_{inc}	Interaction
I (No Constraints)	-	0.21	Age X TP X TD
II (a1=a2, a3=a4)	11.02 (2) **	0.14	Age X TP
III (a1=a3, a2=a4)	23.80 (2) ***	0.05	Age X TD
IV (a1=a2)	0.78 (1), *ns*	0.21	(Age X TD) X TP
V (a3=a4)	10.31 (1) **	0.14	(Age X TD) X TP
VI (a1=a3)	22.79 (1) ***	0.03	(Age X TD) X TP
VII (a2=a4)	4.77 (1) *	0.18	(Age X TD) X TP
VIII (a1=a2=a3=a4)	29.15 (3) ***	-	Main Effect of Age
IX (a1=a2=a3=a4=0)	213.39 (4) ****	-	No Effect of Age

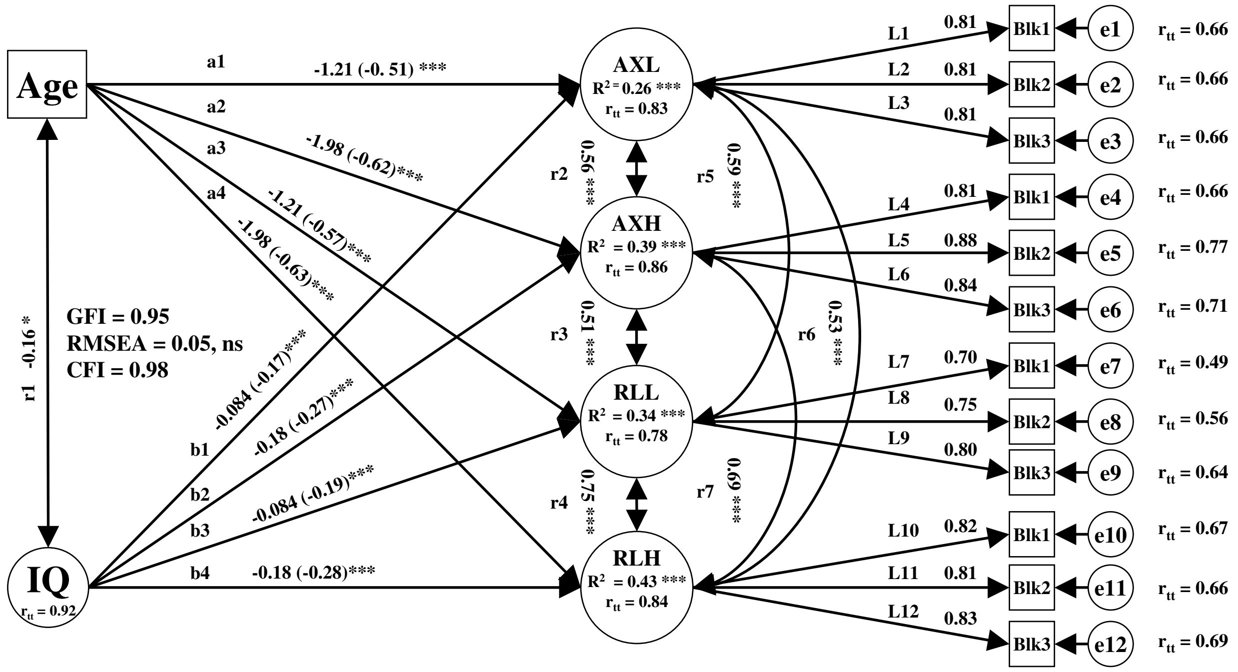
IQ Model	Chi-Sq Diff (df)	CFI_{inc}	Interaction
I (No Constraints)	-	0.02	IQ X TP X TD
II (b1=b2, b3=b4)	1.33 (2), *ns*	0.04	IQ X TP
III (b1=b3, b2=b4)	3.05 (2), *ns*	0.00	IQ X TD
IV (b1=b2)	0.76 (1), *ns*	0.03	(IQ X TD) X TP
V (b3=b4)	0.53 (1), *ns*	0.03	(IQ X TD) X TP
VI (b1=b3)	1.51 (1), *ns*	0.01	(IQ X TD) X TP
VII (b2=b4)	2.52 (1), *ns*	-0.02	(IQ X TD) X TP
VIII (b1=b2=b3=b4)	3.85 (3), *ns*	-	Main Effect of IQ
IX (b1=b2=b3=b4=0)	69.21 (4) ****	-	No Effect of IQ

Figure 2. Relationships of age and intelligence to omission error rates. R^2 values = proportions of variation in latent omission error variables accounted for by age and intelligence; r_{tt} values inside latent variables = reliability estimates for latent omission error variables; Double-headed arrows r2–r7 linking latent omission error variables = correlations among tasks after accounting for their relationships to age and intelligence; r_{tt} values next to error variances = reliability estimates for observed omission error *t*-scores; Non-parenthetical values above arrows = unstandardized regression weights; Parenthetical values above arrows = standardized regression weights; a1–a4 in tables refer to regression weights shown in the figure; b1–b4 in tables refer to regression weights b1–b4 shown in the figure; TP = Target paradigm (i.e., type of CPT); TD = Target density (proportion of stimuli comprised of targets); *** $p < .0001$; ** $p > .001$; * $p < .05$; *ns* = not significant result

IQ in Figures 2 and 3). Age and intelligence are related to latent CPT variables by regression weights (see arrows a1–a4, and b1–b4, respectively).

Hypotheses and corresponding constraints on regression weights. A hierarchically nested series of models was used to contrast relationships among age, intelligence, and performance within and across the four CPTs. Properties of these models are detailed in the tables shown at the bottom of Figures 2 and 3. These tables share a common organizational scheme. Age effects were evaluated first and the results of these analyses are shown in the table at the bottom left corner of each figure. Relationships involving intelligence were evaluated after first accounting for individual differences in age. Results of these analyses are shown in the table in the bottom right hand corner of each figure.

Models in each table are listed in order of increasing stringency. Those at the top impose few restrictions on regression weights whereas those listed toward the bottom impose strong restrictions. Models in the intermediate rows impose constraints that represent tests of main effects and various two- and three-way interactions of subject attributes and task variables. The specific relationship(s) tested by each model are summarized in the first column of each table (e.g., see column 1 of each table in Figures 2 and 3). Note that three-way interactions of subject attributes and performance may assume several different forms, each of which has implications for normative interpretation. For example, model I in the age effects table assumes that age is differentially related to performance across all combinations of task variables. This implies that different normative standards are required for all CPTs under study.



Age Model	Chi-Sq Diff (df)	CFI_{inc}	Interaction
I (No Constraints)	-	0.27	Age X TP X TD
II (a1=a2, a3=a4)	30.82 (2) ***	0.00	Age X TP
III (a1=a3, a2=a4)	1.16 (2), *ns*	0.28	Age X TD
IV (a1=a2)	15.19 (1) ***	0.14	(Age X TD) X TP
V (a3=a4)	21.08 (1) ***	0.08	(Age X TD) X TP
VI (a1=a3)	0.22 (1), *ns*	0.28	(Age X TD) X TP
VII (a2=a4)	1.06 (1), *ns*	0.27	(Age X TD) X TP
VIII (a1=a2=a3=a4)	32.31 (3) ***	-	Main Effect of Age
IX (a1=a2=a3=a4=0)	156.54 (4) ****	-	No Effect of Age

IQ Model	Chi-Sq Diff (df)	CFI_{inc}	Interaction
I (No Constraints)	-	0.14	IQ X TP X TD
II (b1=b2, b3=b4)	12.71 (2) **	0.01	IQ X TP
III (b1=b3, b2=b4)	1.01 (2), *ns*	0.15	IQ X TD
IV (b1=b2)	4.13 (1) *	0.10	(IQ X TD) X TP
V (b3=b4)	10.93 (1), *ns*	0.00	(IQ X TD) X TP
VI (b1=b3)	0.70 (1), *ns*	0.14	(IQ X TD) X TP
VII (b2=b4)	0.13 (1), *ns*	0.15	(IQ X TD) X TP
VIII (b1=b2=b3=b4)	13.01 (3) **	-	Main Effect of IQ
IX (b1=b2=b3=b4=0)	36.42 (4) ****	-	No Effect of IQ

Figure 3 Relationships of age and intelligence to commission error rates. R^2 values = proportions of variation in latent commission error variables accounted for by age and intelligence; r_{tt} values inside latent variables = reliability estimates for latent commission error variables; Double-headed arrows r2–r7 linking latent commission error variables = correlations among tasks after accounting for their relationships to age and intelligence; r_{tt} values next to error variances = reliability estimates for observed commission error *t*-scores; Non-parenthetical values above arrows = unstandardized regression weights; Parenthetical values above arrows = standardized regression weights; a1–a4 in tables refer to regression weights shown in the figure; b1–b4 in tables refer to regression weights b1–b4 shown in the figure; TP = Target paradigm (i.e., type of CPT); TD = Target density (proportion of stimuli comprised of targets); ****p* < .0001; ***p* > .001; **p* < .05; *ns* = not significant result

The chi-square difference tests evaluate the extent of deterioration in fit relative to the least restrictive model (i.e., model I) brought about by imposition of restrictions on regression weights. Conversely, the CFI_{inc} statistic describes proportional improvements in fit relative to model VIII resulting from relaxation of constraints. Model VIII assumes that subject attributes exert main effects but do not interact with task variables. It was used as the baseline standard for these comparisons due to the very strong evidence that the assumption embedded in model IX (i.e., subject attributes bore no relationship to error rates) was unreasonable (see fit indices for model IX under age and IQ effects in Figures 2 and 3).

Correlations among age, intelligence, and omission error rates are shown in Table 2. Corresponding data for commission errors are shown in Table 3. Covariances are more appropriate for structural equation modeling than correlations (Loehlin, 1992). Consequently, correlations were converted to covariances for purposes of analysis.

OE age effects. Model IV provided the best fit to the observed covariances among age, target paradigm, and target probability (see model IV in age effects table of Figure 2). This model imposed the restriction that age effects be equal for the low- and high-density versions of the AX CPT, whereas no constraint was imposed across target probability levels for the repeated letter paradigm. This model fit the data as well as the least restrictive case (i.e., model I, see chi-square difference value for model IV in age effects table of Figure 2), and provided a 21% improvement in fit over the most restrictive model (i.e., model VIII, see CFI_{inc} value for model IV in age

Table 2 Correlations among age, IQ, and normalized omission error rates

	Age	IQ	ALB1	ALB2	ALB3	AHB1	AHB2	AHB3	RLLB1	RLLB2	RLLB3	RLHB1	RLHB2	RLHB3
Age	1.00													
IQ	−0.15	1.00												
ALB1	−0.48	−0.16	1.00											
ALB2	−0.41	−0.14	0.50	1.00										
ALB3	−0.46	−0.09	0.54	0.63	1.00									
AHB1	−0.54	−0.22	0.54	0.54	0.56	1.00								
AHB2	−0.54	−0.20	0.49	0.53	0.57	0.77	1.00							
AHB3	−0.56	−0.15	0.55	0.57	0.62	0.76	0.83	1.00						
RLLB1	−0.56	−0.17	0.44	0.41	0.45	0.54	0.57	0.59	1.00					
RLLB2	−0.57	−0.09	0.41	0.38	0.42	0.55	0.59	0.56	0.73	1.00				
RLLB3	−0.55	−0.08	0.37	0.39	0.43	0.52	0.60	0.59	0.65	0.71	1.00			
RLHB1	−0.50	−0.17	0.47	0.43	0.43	0.58	0.61	0.61	0.62	0.67	0.64	1.00		
RLHB2	−0.49	−0.19	0.46	0.43	0.44	0.58	0.61	0.61	0.62	0.65	0.68	0.85	1.00	
RLHB3	−0.54	−0.18	0.46	0.45	0.47	0.60	0.64	0.64	0.62	0.65	0.69	0.81	0.88	1.00
SD	2.41	11.79	7.14	7.95	8.37	6.16	7.09	6.98	9.10	8.66	8.61	7.88	8.11	8.22

Note: Age = Age measured in whole years; IQ = Composite IQ score from the Kaufman Brief Intelligence Test; ALB1–ALB3 = AX target paradigm, low density condition, blocks 1–3; AHB1–AHB3 = AX target paradigm, high density condition, blocks 1–3; RLLB1–RLLB3 = Repeated letter paradigm, low density condition, blocks 1–3; RLHB1–RLHB3 = Repeated letter paradigm, high density condition, blocks 1–3; SD = standard deviation.

Table 3 Correlations among age, IQ, and normalized commission error rates

	Age	IQ	ALB1	ALB2	ALB3	AHB1	AHB2	AHB3	RLLB1	RLLB2	RLLB3	RLHB1	RLHB2	RLHB3
Age	1.00													
IQ	−0.15	1.00												
ALB1	−0.43	−0.11	1.00											
ALB2	−0.38	−0.07	0.64	1.00										
ALB3	−0.33	−0.05	0.64	0.68	1.00									
AHB1	−0.50	−0.15	0.57	0.48	0.49	1.00								
AHB2	−0.47	−0.18	0.53	0.45	0.49	0.70	1.00							
AHB3	−0.45	−0.12	0.52	0.42	0.46	0.67	0.76	1.00						
RLLB1	−0.47	−0.06	0.49	0.46	0.49	0.47	0.46	0.44	1.00					
RLLB2	−0.41	−0.09	0.40	0.37	0.37	0.43	0.44	0.38	0.52	1.00				
RLLB3	−0.38	−0.05	0.46	0.44	0.44	0.44	0.43	0.42	0.54	0.62	1.00			
RLHB1	−0.53	−0.18	0.49	0.46	0.45	0.59	0.61	0.57	0.48	0.53	0.53	1.00		
RLHB2	−0.50	−0.05	0.48	0.41	0.42	0.53	0.58	0.57	0.47	0.47	0.52	0.66	1.00	
RLHB3	−0.46	−0.15	0.47	0.41	0.43	0.53	0.57	0.54	0.47	0.55	0.60	0.66	0.69	
SD	2.41	11.79	6.94	7.09	7.20	9.16	9.09	8.96	7.16	7.30	7.78	9.22	8.99	9.11

Note: Age = Age measured in whole years; IQ = Composite IQ score from the Kaufman Brief Intelligence Test; ALB1–ALB3 = AX target paradigm, low density condition, blocks 1–3; AHB1–AHB3 = AX target paradigm, high density condition, blocks 1–3; RLLB1–RLLB3 = Repeated letter paradigm, low density condition, blocks 1–3; RLHB1–RLHB3 = Repeated letter paradigm, high density condition, blocks 1–3; SD = standard deviation.

effects table of Figure 2). All the alternative models exhibited deterioration in fit relative to the unconstrained case. This indicates that the influence of age varied with increasing density for the repeated letter task but not for the AX CPT. As such, separate age-based omission error norms are required for the repeated letter tasks whereas omission error rates for the AX version can be judged against a common age-based standard.

OE IQ effects. IQ scores were significantly related to performance after accounting for the influence of age. This is reflected in the very poor fit of the model in which regression weights linking IQ to omission error scores were constrained to values of zero (see model IX in the IQ table in Figure 2). All other models in the series fit the data adequately, including the most parsimonious case (i.e., Model VIII) which restricted all IQ-related regression weights to a single common value. The latter showed no significant deterioration in fit relative to the completely unconstrained model (see chi-square difference value for model VIII in IQ effects table of Figure 2) and none of the less parsimonious alternatives provided any substantial improvement in fit (see CFI_{inc} values for models II–VII in IQ effects table of Figure 2). This pattern of results indicates that IQ exerted a main effect on omission error rates after controlling for effects of age (i.e., made equal predictive contributions to performance across all four tasks). This suggests that omission error norms for all tasks require additive adjustments of identical magnitude based on individual differences in IQ over and above adjustments based on age.

Fit of final OE model and parameter estimates. Fit statistics and parameter estimates for the model incorporating both age and IQ effects are shown in the middle left portion of Figure 2. The model demonstrated excellent fit to the observed data (see GFI, RMSEA, and CFI values in Figure 2). The generality of the model was evaluated by assessing its fit among various subsamples of children. Identical results were obtained when the model was fit to data obtained from boys alone, girls alone, children with IQ scores falling in the average range, and children with IQ scores in the above-average range. The interaction of gender and IQ was examined to assess whether model fit varied as a multiplicative function of both gender and intelligence. The interaction was non-significant.

Finally, age was correlated with the crystallized (i.e., Vocabulary subtest) but not the fluid (i.e., Matrices subtest) component of intelligence. Consequently, the model was examined using the former measure as the estimate of IQ rather than the composite score to ensure that obtained parameter estimates were not unduly influenced by the inclusion of age-correlated verbal scores. No differences in model fit or parameter estimates were observed. Collectively, these data suggest that gender differences were not evident in our data, and parameter estimates were not unduly influenced by the somewhat elevated IQ scores of the children studied or the inclusion of age-related verbal items in the composite measure of intelligence.

Examination of the raw regression weights in Figure 2 show that a one-year difference in age was associated with a decrease in omission error *t*-score of 1.42 points for both low and high density AX CPTs (see non-parenthetical values above arrows a1 and a2 in Figure 2). The standardized regression weights indicate that a pair of children separated by one standard deviation in age differed by approximately two-thirds of a standard deviation in omission error rates on this instrument (see parenthetical values above arrows a1 and a2 in Figure 2). In contrast, age was more strongly related to performance on the repeated letter CPT at low target density in comparison to the high target probability condition. A difference of one year was associated with a 2.17 point reduction in omission error *t*-score in the low density condition, whereas the same age discrepancy was associated with smaller performance gains (1.70 points) at high density (see non-parenthetical values above arrows a3 and a4 in Figure 2). Standardized regression weights show that a standard deviation in age was associated with decreases of .71 and .60 standard deviations in omission error rates for low- and high-density conditions, respectively (see parenthetical values above arrows a3 and a4 in Figure 2).

After controlling for the effects of age, predictive contributions made by IQ scores were statistically significant but small. An increase of ten IQ points was associated with a 1.4 point decrease in omission error *t*-score for all tasks (see non-parenthetical values above arrows b1 to b4 in Figure 2). The standardized weights suggest that a standard deviation in IQ was associated with decreases of .22 to .32 standard deviations in omission error score (see parenthetical values above arrows b1 to b4 in Figure 2). The joint contributions of age and intelligence accounted for 37% to 50% of the variation in omission error rates across CPTs (see R^2 values shown within the circles labeled AXL, AXH, RLL, and RLH in Figure 2). These values are 5% to 10% higher than those that emerged when age was used as the sole predictor of performance.

CE age effects. Results of analyses pertaining to the influence of age on commission error rates are shown in Figure 3. Models III, VI, and VII fit the data as well as the completely unconstrained case (see chi-square difference values for models III, VI, and VII in age effects table in Figure 3), and all yielded comparable improvements in fit relative to the most constrained case (i.e., model VIII, see CFI_{inc} values for models III, VI, and VII in age effects table of Figure 3). Model III is the most parsimonious of these alternatives. Consequently, model III was accepted as the best description of the data among the tested alternatives. It expresses a two-way interaction of age with target density by restricting age effects to be equal within density levels for both target paradigms. Thus, separate age-based commission error norms appear to be necessary for low and high-density tasks.

CE IQ effects. The poor fit of model IX indicates that IQ scores were significantly related to performance after accounting for the influence of age (see chi-square difference value for model IX in IQ effects table of Figure 3). Models III, V, VI, and VII fit the data as well as the completely unconstrained case (see Chi-Square difference values for models III, V, VI, and VII in IQ table of Figure 3). Model III is the most parsimonious of these alternatives, and models V, VI, and VII failed to yield any substantial improvement in fit over model III using the unconstrained model as a comparative standard (see CFI_{inc} values for models III, V, VI, and VII in IQ table of Figure 3). Consequently, model III was accepted as the best description of the data among the tested alternatives. It restricts IQ effects to be equal within density levels for both target paradigms. This suggests that additive IQ-based adjustments over and above corrections for individual differences in age are necessary for developmentally appropriate interpretation of commission error rates, and the size of the IQ-based corrections must be larger for high density conditions than for low density conditions for both AX and repeated letter CPTs.

Fit of final CE model and parameter estimates. Fit statistics and parameter estimates for the model incorporating both age and IQ effects on commission error scores are shown in the middle left portion of Figure 3. The relationships posited by this model were highly consistent with observed patterns of covariation among commission error rates, subject attributes and task variables (see GFI, RMSEA, and CFI values in Figure 3). Model fit and parameter estimates were evaluated in subsamples of children to ensure generality. Identical model fit and parameter estimates were obtained for boys, girls, children with IQ scores in the average range, and children with IQ scores in the above-average range. Gender and IQ did not interact as predictors of commission error rates. Model fit and parameter estimates did not differ when an estimate of fluid IQ was substituted for the composite intelligence score. These findings are consistent with those obtained using the OE data. They indicate that gender was unrelated to CE rates, and parameter estimates were not unduly influenced by the above-average IQ scores of the children studied, or the inclusion of age-related verbal data in the score used to estimate IQ.

A one-year discrepancy in age was associated with decreases in commission error *t*-scores of 1.21 and 1.98 points for low and high density tasks, respectively (see non-parenthetical values on arrows a1 and a3, and a2 and a4, respectively, in Figure 3). An increase of one standard deviation in age corresponded to decreases of approximately .5 and .6 standard deviations in commission error rates for low and high density conditions (see parenthetical values on arrows a1 and a3, and a2 and a4, respectively, in Figure 3).

Values of regression weights linking IQ to commission error scores were lower than those associated with age. After controlling for the influences of age, ten point discrepancies in IQ were associated with differences of .84 and 1.8 *t*-score points for low and high density conditions, respectively (see non-parenthetical values on arrows b1 and b3, and b2 and b4, respectively, in Figure 3). The standardized estimates indicate that one standard deviation in IQ was associated with differences of approximately .17 and .27 standard deviations in commission error rates for low and high density tasks, respectively (see parenthetical values above arrows b1 and b3, and b2 and b4, respectively, in Figure 3). The joint contributions of age and intelligence accounted for 26% to 43% of the variation in commission error rates across CPTs (see R^2 values in circles labeled AXL, AXH, RLL, and RLH in Figure 3). Inclusion of IQ in the model added 2% to 9% to R^2 values that emerged when age was used as the sole predictor of performance.

Discussion

Effects of task variables on performance

Changes of target density exerted clearly different effects on performance across AX and repeated letter CPTs for both error types in accord with the formulation derived from Cowan's embedded process model of working memory. Specifically, both OE and CE rates increased as functions of elevations in density for the AX CPT. In contrast, OE rates declined at high density for the repeated letter model and CE rates increased by a much smaller margin at high target density than was observed for the AX task.

Target paradigm and density also interacted as influences on the magnitude and form of performance decrements. Omission error rates increased at different rates across tasks, and these increases reached statistical significance for three of the four instruments under study. Temporal changes in commission error rates were comparable in form for all four tasks (i.e., all trends were linear) but differed in magnitude as a joint function of target paradigm and target probability.

The data provide strong psychometric evidence that scores may not generalize adequately across CPTs differing in design, and vigilance decrements may vary widely in both form and magnitude across combinations of task variables and error type. This strongly implies that CPT scores cannot be interpreted independently of task structure.

These findings have significant implications for the study of attention and its relationship to child psychopathology. The present study asserts and provides preliminary evidence for a relationship between frequency of demand for shifts of attentional focus and CPT performance. Attentional phenomena are generally attributed to executive control processes within working memory frameworks described by both Cowan (Cowan, 1999) and Baddeley (Baddeley & Logie, 1999). Thus, if further evidence accrues in support of this hypothesis then CPTs may be legitimately viewed as measures of executive control processes as defined within these models of working memory.

In addition, global performance deficits on CPTs have consistently distinguished adults with schizophrenia (Cornblatt & Keilp, 1994) from healthy control cases. Similarly, global OE and CE rates discriminate children diagnosed with ADHD from their typically developing peers (Losier et al., 1996), although distinctions between children with ADHD and other psychiatric disorders are often difficult to demonstrate (Barkley, 1997). Since the present study suggests that the capacity to shift focus of attention contributes to CPT performance, this facet of working memory function may help to account for observed differences between psychiatric patient groups and normal controls.

Finally, the present study examined performance within a group of children across multiple CPTs comprised of theoretically guided combinations of task variables, whereas the preponderance of investigations of CPT performance in children have utilized CPTs defined by single, fixed sets of task parameters (Losier et al., 1996). The results show that this approach can yield conceptually and psychometrically illuminating data that cannot be obtained using single tasks. Although only a few other studies have reported experimental analyses of task variables to date, the cumulative results suggest that further efforts in this direction are strongly warranted (Borgaro et al., 2003; Chee et al., 1989; J.P. Leung et al., 2000; Rose et al., 2001).

Relationships of subject attributes to CPT performance

Influence of development. The evidence clearly showed that both omission and commission error rates were related to age. This is hardly surprising in view of the frequency with which relationships between age and CPT performance have been reported. It is also consistent with findings suggesting that the magnitude of age effects on performance can vary across levels of singly manipulated task variables (Lin et al., 1999). The data reported here are nevertheless unique. They show that the four tasks were not equally age dependent, and the level of age dependence was controlled by different combinations of task variables across error types. Specifically, age-related decreases in OE rates were equal across density conditions for the AX CPT, whereas OE rates declined at a higher rate as a function of increasing age at low compared to high target density for the repeated letter model. This pattern of age-dependencies has clear implications for stratification of norms among the instruments under study. Specifically, separate age-based norms are necessary to appropriate interpretation of OE rates derived from low and high density versions of the repeated letter paradigm. In contrast, a single set of OE norms would suffice for the two AX CPTs.

CE rates were primarily controlled by density levels rather than target paradigm or interactions of density with target paradigm (i.e., higher CE rates were uniformly observed under higher levels of target density). Furthermore, CE rates declined as a function of age in all conditions, but the relationship was stronger at high than at lows target density irrespective of target paradigm. Thus, a single set of age-stratified norms would suffice for the two CE rates obtained under low density conditions. Similarly, CE rates obtained under high density conditions could be referenced against a common data base for purposes of interpretation.

Intelligence and CPT performance. After controlling for individual differences in age, small but significant relationships of equal magnitude were observed between intelligence and omission error rates. In contrast, intelligence was more strongly related to commission error rates in high than low target density conditions for both target paradigms. Thus, IQ-dependence did not provide a means of distinguishing tasks from each other when performance was indexed on the basis of omission errors, but did discriminate between density levels when CEs were used as the dependent measure.

These results are in broad accord with Ballard's (Ballard, 1996) conclusion that CPT error rates are typically correlated to at least a modest degree with measures of intelligence. They are unique, however, in showing that statements about the relationship of intelligence to CPT performance must be conditioned on target density level, age, and intelligence. This has important implications for clinical interpretation of performance indices, especially when multiple, different CPTs are administered. For example, a pair of children within the same age range but who differ with respect to intelligence may also differ in omission error rate. Since the difference in OE rate is confounded by a discrepancy in broader cognitive skill, interpretation of the difference in CPT performance solely in terms of attention would be inappropriate. Correction of scores for both age and intelligence is likely to moderate or eliminate interpretive risks of this nature.

Gender and CPT performance. No evidence emerged in the present investigation to suggest that gender contributed to prediction of CPT scores singly or in combination with any task or other subject variable. This finding is consistent with numerous other studies that have failed to find gender differences on CPTs using alphanumeric stimuli (Edley & Knopf, 1987; Lam & Beale, 1991; McKay, Halperin, Schwartz, & Sharma, 1994; Swanson & Cooney, 1989). Studies of adolescents and young adults, however, suggest that the likelihood of gender-related variation in CPT performance may be controlled by level of demand for visual-spatial (i.e., non-verbal) analysis of stimuli (Dittmar et al., 1993; Lin et al., 1999; Prinzel & Freeman, 1997). This issue has not been examined systematically among children. Consequently, experimental analysis of gender differences across visual-spatial and alphanumeric target stimuli among children would fill a gap in the existing literature and further clarify conditions under which it may be necessary to consider gender in the interpretation of CPT scores.

Conclusions

Target paradigm and target density interacted as influences on both OE and CE rates in accord with predictions based on a contemporary model of working memory. Functional distinctions among the CPTs under study were accompanied by variable patterns of relationship to subject attributes relevant to construction of norms. Collectively, these data show that conceptually and developmentally appropriate interpretation of CPT scores must take task structure and error type into account. Contemporary models of working memory offer a useful theoretical frame of reference for conceptualizing CPT performance in children. Studies examining variations in performance across multiple CPTs designed to impose varying levels of challenge on working memory resources are likely to yield theoretically and clinically illuminating data that cannot be obtained from tasks based on fixed sets of task parameters.

Correspondence to

Colin B. Denney, Department of Psychological Sciences, Purdue University, 703 3rd Street, #1130A, West Lafayette, IN 47907, USA; Email: cbdphd@psych.purdue.edu

References

Achenbach, T.M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont Department of Psychiatry.

Aylward, G.P., Gordon, M., & Verhulst, S.J. (1997). Relationships between continuous performance task scores and other cognitive measures: Causality or commonality. *Assessment*, *4*, 325–336.

Baddeley, A.D. (1997). Attention and the control of memory. In A.D. Baddeley (Ed.), *Human memory: Theory and practice* (rev. edn, pp. 85–102). Upper Saddle River, NJ: Pearson Education.

Baddeley, A.D., & Logie, R.H. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). New York: Cambridge University Press.

Balakrishnan, J.D. (1998). Measures and interpretations of vigilance performance: Evidence against the detection criterion. *Human Factors*, *40*, 601–623.

Ballard, J.C. (1996). Computerized assessment of sustained attention: A review of factors affecting vigilance performance. *Journal of Clinical and Experimental Neuropsychology*, *18*, 843–863.

Banaschewski, T., Brandeis, D., Heinrich, H., Albrecht, B., Brunner, E., & Rothenberger, A. (2003). Association of ADHD and conduct disorder-brain electrical evidence for the existence of a distinct subtype. *Journal of Child Psychology and Psychiatry*, *44*, 356–376.

Barkley, R.A. (1997). Attention-deficit/hyperactivity disorder. In E.J. Mash & L.G. Terdal (Eds.), *Assessment of childhood disorders* (3rd edn, pp. 71–129). New York, NY: Guilford Press.

Barkley, R.A., Grodzinsky, G., & DuPaul, G.J. (1992). Frontal lobe functions in attention deficit disorder with and without hyperactivity: A review and research report. *Journal of Abnormal Child Psychology*, *20*, 163–188.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.

Biederman, J., Faraone, S.V., Doyle, A., Lehman, B.K., Kraus, I., Perrin, J., et al. (1993). Convergence of the child behavior checklist with structured interview-based psychiatric diagnosis of ADHD children with and without comorbidity. *Journal of Child Psychology and Psychiatry*, *34*, 1241–1251.

Borgaro, S., Pogge, D.L., Deluca, V.A., Bilginer, L., Stokes, J., & Harvey, P.D. (2003). Convergence of different versions of the continuous performance test: Clinical and scientific implications. *Journal of Clinical and Experimental Neuropsychology*, *25*, 283–292.

Browne, M.W., & Cudeck, C.R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Chee, P., Logan, G., Schachar, R., Lindsay, P., & Wachsmuth, R. (1989). Effects of event rate and display time on sustained attention in hyperactive, normal, and control children. *Journal of Abnormal Child Psychology*, *17*, 371–391.

Chen, W.J., Faraone, S.V., Biederman, J., & Tsuang, M.T. (1994). Diagnostic accuracy of the Child Behavior Checklist for attention-deficit hyperactivity disorder: A receiver-operating characteristic analysis. *Journal of Consulting and Clinical Psychology*, *62*, 1017–1025.

Conners, C., Epstein, J.N., Angold, A., & Klaric, J. (2003). Continuous performance test performance in a normative epidemiological sample. *Journal of Abnormal Child Psychology*, *31*, 555–562.

Corkum, P.V., & Siegel, L.S. (1993). Is the continuous performance task a valuable research tool for use with children with attention-deficit hyperactivity disorder? *Journal of Child Psychiatry and Psychology, 34,* 1217–1239.

Cornblatt, B.A., & Keilp, J.G. (1994). Impaired attention, genetics, and the pathophysiology of schizophrenia. *Schizophrenia Bulletin*, *20*, 31–46.

Cornblatt, B.A., Lenzenweger, M.F., & Erlenmeyer-Kimling, L. (1989). The continuous performance test, identical pairs version: II. Contrasting attentional profiles in schizophrenic and depressed patients. *Psychiatry Research*, *29*, 65–85.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, *104*, 163–191.

Cowan, N. (1997). *Attention and memory: An integrated framework*. New York: Oxford University Press.

Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). New York: Cambridge University Press.

Davies, D.R., & Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press.

Dittmar, M.L., Warm, J.S., Dember, W.N., & Ricks, D.F. (1993). Sex differences in vigilance performance and perceived workload. *Journal of General Psychology*, *120*, 309–322.

Edley, R.S., & Knopf, I.J. (1987). Sustained attention as a predictor of low academic readiness in a preschool population. *Journal of Psychoeducational Assessment*, *4*, 340–352.

Eghbalieh, B., Crinella, F.M., Hunt, L.E., & Swanson, J.M. (2000). Prenatal stimulant exposure affects alerting and executive control in six- and seven-year-olds. *Journal of Attention Disorders*, *4*, 5–13.

Elvevag, B., Weinberger, D.R., & Suter, J.C. (2000). Continuous performance test and schizophrenia: A test of stimulus–response compatibility, working memory, response readiness, or none of the above? *American Journal of Psychiatry*, *157*, 772–780.

Epstein, J.N., Conners, C.K., Sitarenios, G., & Erhardt, D. (1998). Continuous performance test results of adults with attention deficit hyperactivity disorder. *The Clinical Neuropsychologist*, *12*, 155–168.

Fried, P.A., & Watkinson, B. (2001). Differential effects on facets of attention in adolescents prenatally exposed to cigarettes and marijuana. *Neurotoxicology and Teratology*, *23*, 421–430.

Grandjean, P., Weihe, P., Burse, V.W., Needham, L.L., Storr-Hansen, E., Heinzow, B. et al. (2001). Neurobehavioral deficits associated with PCB in 7-year-old children prenatally exposed to seafood neurotoxicants. *Neurotoxicology and Teratology*, *23*, 305–317.

Halperin, J.M., Newcorn, J.H., Matier, K., Sharma, V., McKay, K., & Schwartz, S. (1993). Discriminant validity of attention-deficit hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*, 1038–1043.

Hatta, T. (1993). Differences in the vigilance performance of children as related to gender and achievement in abacus learning. *Psychologia*, *36*, 47–52.

Horn, W.F., Wagner, A.E., & Ialongo, N. (1989). Sex differences in school-aged children with pervasive attention deficit hyperactivity disorder. *Journal of Abnormal Child Psychology*, *17*, 109–125.

Ifuku, G.K. (1996). *State of Hawaii data book*. Honolulu, Hawaii: Department of Business, Economic Development, and Tourism.

Kaufman, A.S., & Kaufman, N.L. (1990). *Manual for the Kaufman brief intelligence test (K-BIT)*. Circle Pines, MN: American Guidance Service.

Kearney, K.A., Hopkins, R.H., Mauss, A.L., & Weisheit, R.A. (1983). Sample bias resulting from a requirement for written parental consent. *Public Opinion Quarterly*, *47*, 96–102.

Kline, R.B. (1998). *Principles and practice of structural equation modeling*. New York: The Guilford Press.

Lam, C.M., & Beale, I.L. (1991). Relations among sustained attention, reading performance, and teacher's ratings of behavior problems. *Remedial and Special Education*, *12*, 40–46.

Leung, J.P., Leung, P.W., & Tang, C.S. (2000). A vigilance study of ADHD and control children: Event rate and extra-task stimulation. *Journal of Developmental and Physical Disabilities*, *12*, 187–201.

Leung, P.W., & Luk, S.L. (1988). Differences in attention control between 'clinic-observable' and reported hyperactivity: A preliminary report. *Child Care Health and Development*, *14*, 199–211.

Lin, C., Hsiao, C.K., & Chen, W.J. (1999). Development of sustained attention assessed using the continuous performance test among children 6–15 years of age. *Journal of Abnormal Child Psychology*, *27*, 403–412.

Lindsay, R.L., Tomazic, T., Levine, M.D., & Accardo, P.J. (2001). Attentional function as measured by a continuous performance task in children with dyscalculia. *Journal of Developmental and Behavioral Pediatrics*, *22*, 287–296.

Linn, R.L., & Slinde, J.A. (1977). The determination of the significance of change between pre- and post-testing periods. *Review of Educational Research*, *47*, 121–150.

Liu, S.K., Chiu, C.H., Chang, C., Hwang, T.J., Hwu, H.G., & Chen, J. (2002). Deficits in sustained attention in schizophrenia and affective disorder: Stable versus state-dependent markers. *American Journal of Psychiatry*, *159*, 975–982.

Loehlin, J.C. (1992). *Latent variable models* (2nd edn). Hillsdale, NJ: Erlbaum.

Loge, D.V., Staton, R.D., & Beatty, W.W. (1990). Performance of children with ADHD on tests sensitive to frontal lobe dysfunction. *Journal of the American Academy of Child and Adolescent Psychiatry*, *29*, 540–545.

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

Lovett, M.C., Reder, L.M., & Lebiere, C. (1999). Modeling working memory in a unified architecture: An ACT-R perspective. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 135–182). New York: Cambridge University Press.

Maruyama, G.M. (1998). *Basics of structural equation modeling*. Thousand Oaks, CA: Sage Publications, Inc.

Matier-Sharma, K., Perachio, N., Newcorn, J.H., Sharma, V., & Halperin, J.M. (1995). Differential diagnosis of ADHD: Are objective measures of attention, impulsivity, and activity level helpful? *Child Neuropsychology, 1,* 118–127.

McKay, K.E., Halperin, J.M., Schwartz, S.T., & Sharma, V. (1994). Developmental analysis of three aspects of information processing: Sustained attention, selective attention, and response organization. *Developmental Neuropsychology*, *10*, 121–132.

Morris, R.D. (1996). Relationships and distinctions among the concepts of attention, memory, and executive function: A developmental perspective. In G.R. Lyon & N.A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 3–10). Baltimore: Paul H. Brookes Publishing Co.

Okazaki, S., Maekawa, H., & Futakami, S. (2001). Control of response inhibition in children with ADHD on the continuous performance test under various intervals. *Japanese Journal of Special Education*, *38*, 97–103.

Pascualvaca, D.M., Anthony, B.J., Arnold, L.E., Rebok, G.W., Ahearn, M.B., Kellam, S.G., et al. (1997). Attention performance in an epidemiological sample of urban children: The role of gender and verbal intelligence. *Child Neuropsychology*, *3*, 13–27.

Porrino, L.J., Rapoport, J.L., Behar, D., Sceery, W., Ismond, D.R., & Bunney, W.E. (1983). A naturalistic assessment of the motor activity of hyperactive boys. *Archives of General Psychiatry*, *40*, 681–687.

Prinzel, L.J., & Freeman, F.G. (1997). Task-specific sex differences in vigilance performance: Subjective workload and boredom. *Perceptual and Motor Skills*, *85*, 1195–1202.

Riccio, C.A., Reynolds, C.R., Lowe, P., & Moore, J.J. (2002). The continuous performance test: A window on the neural substrates for attention. *Archives of Clinical Neuropsychology*, *17*, 235–272.

Rose, C.L., Murphy, L.B., Schickedantz, B., & Tucci, J. (2001). The effects of event rate and signal probability on children's vigilance. *Journal of Clinical and Experimental Neuropsychology*, *23*, 215–224.

Schachar, B., Logan, G., Wachsmuth, R., & Chajczdy, D. (1988). Attaining and maintaining preparation: A comparison of attention in hyperactive, normal, and disturbed control children. *Journal of Abnormal Child Psychology*, *16*, 361–378.

See, J.E., Warm, J.S., Dember, W.N., & Howe, S.R. (1997). Vigilance and signal detection theory: An empirical evaluation of five measures of response bias. *Human Factors*, *39*, 14–29.

Seidel, W.T., & Joschko, M. (1990). Evidence of difficulties in sustained attention in children with ADDH. *Journal of Abnormal Child Psychology*, *18*, 217–229.

Sergeant, J.A., & van der Meere, J. (1990). Convergence of approaches in localizing the hyperactivity deficit. In B.B. Lahey & A.E. Kazdin (eds.), *Advances in clinical child psychology* (vol. 13, pp. 207–246). New York: Plenum Press.

Swanson, H.L., & Cooney, J.B. (1989). Relationship between intelligence and vigilance in children. *Journal of School Psychology*, *27*, 141–153.

Till, C., Koren, G., & Rovet, J.F. (2001). Prenatal exposure to organic solvents and child neurobehavioral performance. *Neurotoxicology and Teratology*, *23*, 235–245.

United States Bureau of the Census. (1990). *1990 Census of population characteristics: Hawaii* (vol. 1990 CP-1-13, June 1992). Washington, DC: US Bureau of the Census.

Walkowiak, J., Altmann, L., Kramer, U., Sveinsson, K., Turfeld, M., Weishoff-Houben, M., et al. (1998). Cognitive and sensorimotor functions in 6-year-old children in relation to lead and mercury levels: Adjustment for intelligence and contrast sensitivity in computerized testing. *Neurotoxicology and Teratology*, *20*, 511–521.

Wilkinson, L., Blank, G., & Gruber, C. (1996). Transformations. In *Desktop data analysis with SYSTAT* (pp. 695–713). Upper Saddle River, NJ: Prentice-Hall.

Manuscript accepted 15 March 2004