Neuropsychology
2007, Vol. 21, No. 1, 45–64


0894-4105/07/$12.00 DOI: 10.1037/0894-4105.21.1.45

# The Latent Structure of Attention-Deficit/Hyperactivity Disorder in a Clinic-Referred Sample

Thomas W. Frazier
The Cleveland Clinic

Eric A. Youngstrom
University of North Carolina at Chapel Hill

Richard I. Naugle
The Cleveland Clinic

The question of whether attention-deficit/hyperactivity disorder (ADHD) represents a continuum of attentional and executive dysfunction or a natural category has yet to be extensively investigated. Subjective report and neuropsychological data from 437 individuals referred for neuropsychological evaluation were analyzed using latent class and taxometric analyses (mean above minus below a cut [MAMBAC], maximum eigenvalue [MAXEIG], and latent mode [LMODE]). Results indicated no significant evidence for a taxonic representation of ADHD across multiple procedures and indicator sets. Similarly, there was no evidence that ADHD subtypes represent a qualitative distinction. These findings may suggest that current diagnostic conceptualizations are inadequate for accurately identifying and characterizing individuals with problems related to attention and executive dysfunction. Alternatively, the null findings may have resulted from inadequate indicator selection. A dimensional model may better facilitate accurate identification of individuals at risk for functional impairment.

*Keywords:* attention-deficit/hyperactivity disorder, taxometrics, latent class

The current diagnostic conceptualization of attention-deficit/hyperactivity disorder (ADHD) requires the presence of at least six of nine symptoms of inattention and/or hyperactivity/impulsivity with onset of symptoms prior to age 7 (American Psychiatric Association, 1994). The diagnostic algorithm results in three possible diagnostic subtypes, inattentive (ADHD–I), hyperactive/impulsive (ADHD–H/I), and combined type (ADHD–C), depending on whether individuals meet criteria in one or both symptom sets. These criteria have changed somewhat from the third to the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–III* and *DSM–IV*; American Psychiatric Association, 1980, 1994), with varying complexity of the diagnostic criteria. However, all of the diagnostic formulations of ADHD have focused on the identification of a categorical syndrome, as have all formulations of mental disorders (Adams & Cassidy, 1993; Carson, 1991). This reflects a general bias of *DSM* formulations toward the medical model of diagnostic classification (Widiger & Clark, 2000). Alternatively, other researchers have discussed the general bias of psychologists toward a dimensional structure for all psychopathology (Dahlstrom, 1995; Meehl, 1995). Increasingly, empirical studies have attempted to test these opposing contentions.

## ADHD Diagnosis

Barkley's (1997, 1998) model of ADHD posits that ADHD results in a core deficit in behavioral inhibition that results from deficits in working memory, regulation of affect, internalization of speech, and reconstitution. Multiple studies have supported this model, with the most consistent findings being deficits in executive cognitive functions (Frazier, Demaree, & Youngstrom, 2004; Woods, Lovejoy, & Ball, 2002). However, because these studies have relied on existing diagnostic criteria, it is not clear whether the identified deficits result from differences between distinct categories or taxa (i.e., ADHD vs. non-ADHD) or are merely descriptions of the extreme ends of a continuum of dysfunction. It should also be noted that Barkley's model is applicable only to the hyperactive/impulsive and combined subtypes, as the inattentive type (or at least a subset of individuals receiving this diagnosis) is now thought by some researchers to represent a different disorder (Milich, Balentine, & Lynam, 2001).

Neuropsychological studies have provided some support for ADHD subtype distinctions, with significant differences reported between ADHD–C and ADHD–I on measures of response control and attention/working memory (Hinshaw, Carte, Sami, Treuting, & Zupan, 2002; Lockwood, Marcotte, & Stern, 2001). However, some studies have failed to find significant neuropsychological differences in spite of robust differences in demographic, psychiatric, social functioning, and school functioning variables (Chhabildas, Pennington, & Willcutt, 2001; Faraone, Biederman, Weber, & Russell, 1998). When significant differences have been found, they have tended to be small effects, resulting in modest individual classification (Hinshaw et al., 2002). Even if significant

Thomas W. Frazier, Department of Psychiatry and Psychology, Section of Neuropsychology, and Department of Pediatrics, Section of Behavioral Medicine, The Cleveland Clinic, Cleveland, Ohio; Eric A. Youngstrom, Department of Psychology, University of North Carolina at Chapel Hill; Richard I. Naugle, Section of Neuropsychology, The Cleveland Clinic.

We would like to thank John Ruscio in assisting us with addressing reviewers' comments concerning the taxometric methods used in this article.

Correspondence concerning this article should be addressed to Thomas W. Frazier, Department of Pediatrics, Section of Behavioral Medicine A120, The Cleveland Clinic, 9500 Euclid Avenue, Cleveland, OH 44195. E-mail: fraziet2@ccf.org

neuropsychological and psychosocial differentiation are assumed, the question remains whether these differences result from descriptions of the ends of a continuum or continua or whether they result from ADHD subtypes representing discrete categories.

Milich and colleagues (2001) argued that ADHD subtypes are distinct, unrelated entities. Their conclusions were based on an extensive review of the history of the diagnosis as well as cluster analytic studies that differentiated between children. A recent study using latent class analysis found strong support for two distinct categories, inattention only and inattention with hyperactivity/impulsivity, consistent with those two of the three *DSM–IV* subtypes (Neuman et al., 1999). However, cluster analysis is not able to accurately determine whether data result from a latent taxonic structure or a latent dimensional structure (Waller & Meehl, 1998), and latent class analysis may be more useful for distinguishing between multiple latent categories, with coherent kinetic methods preferred for examining the taxonic hypothesis (Schmidt, Kotov, & Joiner, 2004).

On the basis of the above cited findings and limitations of previously applied methodology, several researchers have indicated that the actual latent status of ADHD subtypes has yet to be determined (Barkley, 2001; Hinshaw, 2001; Lahey, 2001). In fact, there are multiple possibilities regarding the actual latent structure of ADHD subtypes, including multiple latent categories, multiple dimensions, mixtures of dimensions and categories, and nesting of dimensional structure within latent categories. The primary purpose of the present study was to examine whether any latent categories composed of individuals with attention and/or hyperactivity/impulsivity problems exist. Additionally, we examine whether a subset of individuals with inattention, but not hyperactivity, will emerge as a distinct category relative to individuals with both symptom clusters.

## Importance of Latent Structure Identification

The taxonic distinction is not merely academic. Matching the assessment method to the latent structure of a disorder increases the reliability and validity of assessment (for an extended discussion of this issue, see J. Ruscio & Ruscio, 2004). Accurate identification of the latent structure of a disorder is also likely to facilitate research. Studies can implement a more appropriate design (group comparisons vs. correlational designs or quantitative trait genetic models), maximizing power. Also, researchers can focus studies on the most likely etiological pattern. Some researchers have argued that taxonicity is more suggestive of all-or-none causes, whereas dimensionality is more consistent with multiple additive or graded etiologies (Haslam, 1997). Finally, taxometric analyses can help identify moderators of treatment outcome (see Beauchaine, 2003). For example, if a subset of individuals with significant inattention represents an ADHD taxon, these individuals may differentially benefit from particular treatment options (i.e., medication vs. cognitive–behavioral treatment). Thus, identifying the latent structure of ADHD is likely to significantly enhance the interpretation of past research using group designs and to facilitate future research and clinical work.

## Taxometric Analysis of ADHD

A recent taxometric investigation found that ADHD was best characterized as representing a continuum (Haslam et al., 2006). However, this study examined only one indicator set with three variables, each representing a heterogeneous composite of symptoms from the *DSM–IV* ADHD symptom clusters. Thus, because of the limited indicators included in this study, an ADHD taxon may have been missed.

Beauchaine (2003) suggested that candidate indicators of an ADHD taxon be expanded beyond *DSM* criteria. Cognitive test data are a promising source of indicators. Neuropsychological test scores have been shown to differentiate between ADHD subjects and normal controls (Frazier et al., 2004). The magnitude of effects sizes for neuropsychological tests when comparing ADHD and non-ADHD groups typically ranges from medium to large, with the largest effect sizes occurring for measures of academic skills, sustained attention and impulsivity, and intellectual measures. Additionally, neuropsychological test scores provide direct indices of the two primary symptom clusters, inattention and impulsivity/hyperactivity. Thus, although neuropsychological measures do not directly examine the *DSM* criteria, they do assess the primary symptoms of ADHD, and they directly measure some of the core deficits thought to underlie the condition (Barkley, 1998).

Taxometric analyses have typically focused on subjective report data (see Schmidt et al., 2004, for review). However, research has demonstrated that observers' expectations can induce pseudotaxonicity into rating scale data (Beauchaine & Waters, 2003). Supplementing subjective report data with objective test data should provide an additional consistency test to ensure that taxonicity is not inferred because of rater biases. The present study examined the taxonic conjecture by including indicator sets with subjective and objective test data.

## Hypotheses

With the paucity of previous empirical work regarding the latent structure of ADHD and the large number of possibilities regarding the latent structure of ADHD (dimensional, one latent category, multiple latent categories, one or more latent categories that each represent a continuum, etc.), we found it difficult to generate specific predictions. In the interest of making a strong prediction, and with the hopes of spurring additional research, we expected taxonic results across all of the indicator sets examined. This expectation is based on prior findings of ADHD clusters or classes (for a review, see Milich et al., 2001; Neuman et al., 1999) and the current categorical diagnostic conceptualization. The evidence of specific structural brain abnormalities was also considered in generating predictions (Castellanos et al., 1996, 2001), although it should be recognized that these findings could also be consistent with comparisons of the extreme ends of a continuum. Because previous evidence is not particularly strong and current classification systems are based largely on historical considerations, the above predictions should be viewed as strictly a starting point for research in this area.

No predictions were made regarding the composition of the taxon group(s). Many researchers might argue that the individuals who receive clinical diagnoses of ADHD–I represent a distinct category (Milich et al., 2001); however, we could also argue that individuals with both inattention and hyperactivity/impulsivity represent a distinct group. This is based on the relatively uncontroversial clinical observation that individuals with acquired brain injury often show symptoms of inattention and impulsivity that are

categorically distinct from normal variations of inattention and impulsivity. Thus, the possibility that a developmentally based version of impulsivity/hyperactivity, analogous to the acquired version associated with head injury, should not be dismissed (Amor et al., 2005). For example, a subset of individuals with ADHD–C may show qualitatively distinct inattention and impulsivity resulting from selective dysfunction of ventromedial and/or lateral orbitofrontal regions due to genetic and/or early environmental factors, such as perinatal bradycardia or mild fetal alcohol syndrome.

A secondary purpose of the present study was to examine whether ADHD–I represents a distinct category relative to ADHD–C. We expected that a subset of individuals with the clinical diagnosis of ADHD–I would be distinguishable from individuals with ADHD–C and individuals with ADHD–I who are simply subthreshold for hyperactivity/impulsivity symptoms based on previous suggestions of a latent category of individuals with inattention only (Hinshaw, 2001).

## Method

### *Participants*

Data for the present study were obtained from a deidentified patient registry that has been reviewed and approved by the Institutional Review Board at the Cleveland Clinic Foundation. The database consisted of neuropsychological test data from children, adolescents, and adults referred for neuropsychological assessment at the Cleveland Clinic Foundation Section of Neuropsychology. The patient population of this clinic represents a diverse group of individuals drawn from a large metropolitan area and the surrounding suburbs. Approximately 60% of referrals come from other hospital departments (primarily neurology, psychiatry, and psychology) and 40% of referrals come from community sources (typically pediatricians) or represent self-referrals. Individuals with neurological disorders (e.g., seizure disorders, tumors, head injuries, infectious disorders) and full-scale IQs below 70 were excluded from the sample. These exclusion criteria are consistent with the *DSM–IV* diagnostic guidelines that the symptoms of ADHD cannot be better explained by other preexisting neurodevelopmental disorders or mental retardation (American Psychiatric Association, 1994, pp. 83–85) and also follow recommendations that taxometric analyses not include extreme cases whose performance on indicators would make them outliers (Schmidt, Kotov, & Joiner, 2004).

The final sample consisted of 437 individuals (age: $M = 14.96$ years, $SD = 9.70$, range = 5–66; 61% male; 87% right-handed). This sample represents all of the individuals (children and adults) referred to this clinic for evaluation of attention, mood, and/or behavior problems during a 4-year period (2000–2004). All individuals referred during this period were administered a comprehensive neuropsychological assessment battery. Clinical diagnoses of ADHD were made on the basis of the neuropsychological evaluation using *DSM–IV* criteria and relied heavily on clinical interview, neuropsychological tests of attention and executive function, and, for children, parent report of ADHD symptoms. These diagnoses required evidence of impairment in two settings for both child and adult participants. Typically this involved evidence of impairment at home and school for children and young adults or impairment at home and work for middle-aged adults. No individuals were given clinical diagnoses of ADHD–H/I. This likely reflects the lower prevalence of this subtype in the population, particularly in older samples (Barkley, 1998), as well as the heavy emphasis on the assessment of attention during neuropsychological evaluations. Review of the data indicated that it is likely that if only parent-report of ADHD symptoms had been used to make the diagnosis, approximately 10%–15% of individuals with ADHD–C would have been diagnosed as ADHD–H/I, depending on the cutoff used.

A number of individuals were not diagnosed with ADHD or any other neurodevelopmental disorder ($n = 123$). Approximately 44% of these individuals were diagnosed with a mood, anxiety, or adjustment disorder. A review of these individuals' records indicated that in most cases these problems were not severely affecting the participant's everyday functioning, and in all cases, these problems did not significantly influence the patient's test performance (i.e., none of these cases evidenced a neuropsychological impairment defined as more than 2 *SD*s below the mean). Therefore, these individuals are subsequently referred to as *non-ADHD*, because they did not qualify for a clinical diagnosis of ADHD or any other neurodevelopmental disorder (i.e., learning disability, Asperger syndrome, or pervasive developmental disorder not otherwise specified). Table 1 presents characteristics separately by

Table 1
*Total Sample Characteristics*

| | Non-ADHD ($n = 123$) | | | ADHD inattentive ($n = 86$) | | | ADHD combined ($n = 228$) | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | *M* | *SD* | % | *M* | *SD* | % | *M* | *SD* | % |
| Age (years) | 16.67 | 12.13 | | 13.26 | 7.30 | | 14.69 | 8.89 | |
| Sex (% male) | | | 59 | | | 49 | | | 66 |
| Handedness (% right) | | | 89 | | | 85 | | | 86 |
| Learning disability | | | | | | 28 | | | 28 |
| Developmental disability | | | | | | 4 | | | 11 |
| FSIQ | 103.3 | 13.4 | | 102.0 | 11.0 | | 101.9 | 14.5 | |
| Total Achievement | 103.8 | 13.2 | | 100.8 | 12.0 | | 99.5 | 15.0 | |
| Total CPARS | 57.7 | 10.2 | | 60.4 | 8.6 | | 66.1 | 10.8 | |

*Note.* Test results for Full-Scale IQ (FSIQ) and Total Achievement are presented in standard score units ($M = 100$, $SD = 15$); test results for Total Conners Parent ADHD Rating Scale (CPARS) are presented in *T*-score units ($M = 50$, $SD = 10$). ADHD = attention-deficit/hyperactivity disorder.

clinical diagnostic group. Non-ADHD participants tended to be older than both ADHD groups (non-ADHD: $M = 16.7$, $SD = 12.1$; ADHD–I: $M = 13.3$, $SD = 7.3$; ADHD–C: $M = 14.7$, $SD = 8.9$), non-ADHD versus ADHD–I, $t(207) = 2.33$, $p = .02$; non-ADHD versus ADHD–C, $t(349) = 1.75$, $p = .08$; ADHD–I versus ADHD–C, $t(312) = 1.33$, $p = .19$. Diagnostic groups also significantly differed in sex distributions, $\chi^2(2, N = 437) = 8.10$, $p = .02$. A greater proportion of males were found in the ADHD–C subsample (percentage male: non-ADHD = 59.3, ADHD–I = 48.8, ADHD–C = 66.2), consistent with previous literature (Barkley, 1998). For this reason, age- and/or sex-corrected norms were used for neuropsychological tasks and parent-report measures to minimize the likelihood of identifying age or sex taxa (Schmidt et al., 2004). Inspection of the data also reveals that the sample included a wide range of scores for all measures, including the full range of raw scores on the Conners Parent Rating Scale subscales (Conners, 1997). Thus, although this is a clinical sample, it is likely that the full range of impulsivity/hyperactivity and inattention was sampled.

Indicator sets were constructed using only subjective report, objective test data, or mixed subjective report and objective test indicators. Inclusion of varied indicator sets facilitates examination of the generalizability of results. If only subjective report measures produce taxonic findings, the resulting taxon might have been artificially created owing to expectancy effects (Beauchaine & Waters, 2003). Identification of a similar taxon in subjective report, objective test data, and mixed indicator sets would provide strong support for the taxonic conjecture. Alternatively, consistent findings of dimensional structure across multiple indicator sets provide strong evidence against the taxonic conjecture (Meehl, 2004). Subjective report and objective test data measure different aspects of behavior (Shaywitz, Fletcher, & Shaywitz, 1994, 1995) and thus may be differentially sensitive to an ADHD taxon. Inclusion of both subjective report and neuropsychological test data further enhances the possibility that an ADHD taxon would be detected, if present, owing to specification of a diverse array of indicator sets tapping different aspects of cognition and behavior.

### *Suitability of Indicator Sets to Taxometric Procedures*

The initial search for an ADHD taxon used 11 indicator sets on the basis of previous research suggesting performance decrements on measures of attention, impulsivity, academic tasks, intellectual ability measures, working memory, processing speed, and parent-report measures of ADHD symptoms (Barkley, 1997, 1998; Frazier et al., 2004; Goodyear & Hynd, 1992).

Along with theoretical considerations, indicator sets were also chosen according to estimates of indicator validity. Some research has suggested that indicator validities of $d = 1.2$ or greater produce reliable results and accurate classification; however, other studies have suggested that some taxometric procedures may yield interpretable results with indicator validities greater than $d = 1.0$ (for a review, see Schmidt et al., 2004). With this in mind, we examined two very different estimates of indicator validity. The first involved effect sizes derived from comparisons of individuals diagnosed with ADHD in previous research. Effect sizes ($d$) for group discriminations were based on previous meta-analytic work for the neuropsychological variables (Frazier et al., 2004) and normative data for the parent-report variables (Conners, 1997). The range of effect sizes is presented in parentheses for each indicator set described below. Although the effect sizes based on clinical diagnoses are in some cases significantly lower than 1.2, this is likely due to the misassortment of a few individuals when making diagnoses. Such misassortment is expected even under the most careful of research conditions owing to the imperfection of measurement instruments used in making diagnoses. However, these values still demonstrate medium to large effect sizes for all indicator sets.

The second empirical consideration concerned the estimated latent indicator validity derived under the assumption that the data are taxonic. Latent indicator validity estimates, referred to previously as a priori taxometric power analyses, have been used in previous research to estimate the potential utility of indicator sets (Gibb, Alloy, Abramson, Beevers, & Miller, 2004; A. M. Ruscio, Ruscio, & Keane, 2002). Estimates of average indicator validity were derived using the formula provided in Meehl and Yonce (1996, p. 1146). Nuisance correlations were estimated by computing the average correlation from the upper and lower 20% of the resulting distribution. To provide a conservative estimate of latent indicator validity, moderate base-rate estimates were used based on clinical diagnoses.

A priori taxometric power analyses are particularly useful for ruling out the possibility that dimensional results were obtained owing to low indicator validities of the taxon. Table 2 presents sample sizes, sample correlations, estimated nuisance correlations, average skewness, and latent indicator validity ($d$) for each indicator set. Inspection of this table reveals that all indicator sets produced latent indicator validities greater than $d = 1.0$, with most being greater than 1.25. These results suggest that all indicator sets should produce clear results.

In addition to indicator validity, there are two other important considerations for determining whether an indicator set is suitable for taxometric procedures: the overall sample size and the estimated taxon base rate. Taxometric procedures tend to yield less interpretable results when the overall sample size is too low. Meehl (1995) suggested a minimum overall sample size of 300, although other work has suggested that smaller sample sizes can be used for some taxometric procedures when other data parameters are good (Meehl & Yonce, 1994, 1996). Schmidt and colleagues (2004) argued that having a moderate base rate is important for obtaining clear findings when using small to moderate sample sizes. All but one of our total sample indicator sets had sample sizes greater than 300. The exception was the mixture set; however, this set had nine indicators, and larger numbers of indicators tend to produce more stable findings (Beauchaine & Beauchaine, 2002). Several indicator sets from the ADHD–C only subsample and both indicator sets for the ADHD only subsample had sample sizes below 300. However, all sample sizes were greater than or equal to 245, included highly reliable indicators ($r > .75$) with good score ranges, showed clear separation of taxonic and dimensional results in simulations, and had otherwise favorable data parameters. Therefore, we expected results from these indicator sets to be adequate.

The base rate of ADHD clinical diagnosis in the total research sample ranged from .72 to .73. These moderate base rates should produce clear results with sample sizes greater than 300 when at least three indicators are available (Schmidt et al., 2004). Because the overall sample sizes were lower in the ADHD–C subsample

Table 2
*Indicator Set Characteristics for the Total Sample, ADHD–C Subsample, and ADHD Subsample*

| Sample and indicator set | Non-ADHD (*n*) | ADHD–I (*n*) | ADHD–C (*n*) | Total sample (*N*) | Sample (*r*) | Nuisance (*r*) | Average skew | Latent validity (*d*) |
|---|---|---|---|---|---|---|---|---|
| Total sample | | | | | | | | |
| ACH | 87 | 70 | 158 | 315 | .65 | .27 | 0.03 | 2.33 |
| CPARS | 94 | 72 | 176 | 342 | .44 | .27 | −0.30 | 1.23 |
| CPARS–I | 93 | 72 | 173 | 338 | .80 | .24 | 0.47 | 3.69 |
| CPARS–H/I | 93 | 72 | 173 | 338 | .69 | .09 | −0.74 | 3.09 |
| CPT Factors | 103 | 76 | 205 | 384 | .53 | .11 | −1.04 | 2.13 |
| CPT Validity | 103 | 76 | 205 | 384 | .38 | .25 | 0.18 | 1.03 |
| IQ Indices | 110 | 78 | 206 | 394 | .52 | .04 | 0.02 | 2.23 |
| IQ Subtests | 108 | 77 | 205 | 390 | .36 | .15 | −0.14 | 1.28 |
| NP–PS | 92 | 69 | 186 | 347 | .34 | −.14 | −0.11 | 1.95 |
| Mixture | 69 | 53 | 122 | 244 | .28 | .08 | −0.33 | 1.17 |
| ADHD–C subsample | | | | | | | | |
| ACH | 87 | | 158 | 245 | .66 | .24 | 0.08 | 2.32 |
| CPARS | 94 | | 176 | 270 | .47 | .21 | −0.32 | 1.47 |
| CPARS–I | 93 | | 173 | 266 | .82 | .31 | 0.46 | 3.59 |
| CPARS–H/I | 93 | | 173 | 266 | .71 | .13 | −0.62 | 2.99 |
| CPT Factors | 103 | | 205 | 308 | .50 | .13 | −1.09 | 1.82 |
| CPT Validity | 103 | | 205 | 308 | .46 | .25 | 0.13 | 1.32 |
| IQ Indices | 110 | | 206 | 317 | .47 | .03 | 0.05 | 1.91 |
| IQ Subtests | 108 | | 205 | 313 | .37 | .14 | −0.13 | 1.27 |
| NP–PS | 92 | | 186 | 278 | .33 | −.21 | −0.15 | 1.92 |
| Mixture | 69 | | 122 | 191 | .28 | .07 | −0.35 | 1.12 |
| ADHD subsample | | | | | | | | |
| CPARS–H/I | | 72 | 173 | 245 | .68 | .09 | −0.56 | 2.97 |
| NP–PS | | 69 | 186 | 255 | .35 | .21 | −0.10 | 1.11 |

*Note.* ADHD = attention-deficit/hyperactivity disorder (C = combined type; I = inattentive type); ACH = Achievement tests from the Woodcock–Johnson Tests of Ability (3rd ed.); CPARS = Conners Parent ADHD Rating Scale (I = inattention items; H/I = hyperactivity/impulsivity items); CPT Factors = untransformed factor scores derived from the Conners Continuous Performance Test (CPT); CPT Validity = Conners CPT variables with the highest validity for separation of clinical groups; IQ Indices = index scores from the Wechsler Adult Intelligence Scale—III (WAIS–III) and Wechsler Intelligence Scale for Children—III (WISC–III); IQ Subtests = IQ subtests from the WAIS–III and WISC–III; NP–PS = neuropsychological measures of processing speed; Mixture = selected measures with the highest clinical validity from the other indicator sets.

but the base rates of ADHD were closer to .50 (ADHD–C sample base-rate range = .64–.67), results for these analyses were expected to be comparable to those for the total sample.

Each of the indicator sets was examined in both the total sample and a truncated sample with ADHD–I participants removed (ADHD–C subsample). The more distilled sample was included for two reasons. This sample concentrates on the more extreme group, posited to show all of the deficits associated with ADHD, and thus provides an additional consistency test, because similar results in both samples would support either dimensional or taxonic conclusions. Also, some researchers have suggested that ADHD–I and ADHD–C may be separate, unrelated disorders (Milich et al., 2001). Examining a mixed sample may lower the base rate of any potential taxa and potentially change patterns of nuisance covariance, thus increasing the risk of spurious null findings. Individuals with ADHD–C represented the larger proportion of individuals in the total sample, and therefore examination of these individuals independently was methodologically more tenable than examining the ADHD–I subsample independently. All indicator sets were standardized into a *z*-score metric using the sample mean and were transformed so that higher scores indicated better performance or fewer reported symptoms. When possible, all analyses were rerun using the raw scores. Highly similar results were obtained; therefore, only results using the standard scores are presented.

*Achievement (d = 0.70–1.01).* The basic reading, math calculation, written expression, and academic fluency indices from the Woodcock–Johnson Tests of Ability—Third Edition were used as separate variables in one indicator set (Woodcock, McGrew, & Mather, 2001). Previously reported internal consistency reliabilities of these indices ranged from .91 to .95 (McGrew & Woodcock, 2001).

*Conners Parent ADHD Rating Scale (CPARS)—Revised (d = 1.17–3.38).* The Attention/Cognitive Symptoms, Hyperactivity/Impulsivity, and Oppositional Behavior scales from the CPARS–Revised: Short Form were examined in one indicator set (Conners, 1997). Previously reported internal consistency reliabilities for these scales ranged from .86 to .94 (Conners, 1997).

We also examined two indicator sets composed of more homogeneous subsets of CPARS items. These sets were included because taxometric procedures may perform poorly if only one or two indicators from the more heterogeneous indicator sets described above have low validity for separating the taxon from the complement group (Meehl, 1995). The subsets were composed of only inattention items (CPARS–I) and only hyperactivity/impulsivity items (CPARS–H/I). Because individual items have limited range, items were pooled into item parcels by taking the mean of the items (four items per parcel for inattentive items and three items per parcel for hyperactive/impulsive items). This procedure resulted in two sets of three indicators per set. Item parcels

facilitate sufficient range for each indicator variable (13-point for inattentive items; 10-point for hyperactive/impulsive items). The internal consistency reliability of these parcels was quite good for three- to four-item scales (for CPARS–I, $r = .75–.86$; for CPARS–H/I, $r = .64–.78$). Previous empirical work has used a similar procedure for creating input indicators (A. M. Ruscio et al., 2002).

*Conners Continuous Performance Test (CPT; Conners, 2000a; d = 0.48–0.91).* Two sets of indicators consisted exclusively of measures from the Conners CPT. The first indicator set was derived by averaging individual measures included in the Attention, Impulsivity, and Vigilance factors of CPT output. Only omissions, commissions, hit reaction time, hit reaction time standard error, variability, and hit reaction time interstimulus interval were included in the Attention factor average, as the other individual measures were not available. The Impulsivity and Vigilance factors were computed by averaging each of the individual measures included in these sections of the CPT output. This indicator set is referred to subsequently as *CPT Factors*. All of the indicators included in this factor had substantial significant negative skew (Attention = –0.96, Impulsivity = –1.86, and Vigilance = –0.31; average skew = –1.04). Previously reported split-half reliability coefficients for a subset of the individual measures composing these factors ranged from .66 to .95 (Conners, 2000b).

The second indicator set consisted of three individual CPT scores, Commissions, Variability, and Detectability (CPT Validity). These measures were chosen because they discriminated ADHD subjects from non-ADHD subjects better than any of the other individual CPT measures in the present sample.

Comparison of results from CPT Factors and CPT Validity ensured that a taxon was not spuriously identified owing to the effect of skew on the shape of taxometric graphs. CPT Factors had an average skewness of –1.04, whereas CPT Validity had an average skewness of 0.18. Negative skew tends to result in a left-to-right downward tilt to taxometric graphs, whereas positive skew tends to result in a left-to-right upward tilt. Thus, analyses using CPT Factors might tend to spuriously identify a low base-rate taxon when the data are actually consistent with a dimensional structure, whereas CPT Validity might tend to spuriously identify a high base-rate taxon when the data are actually consistent with a dimensional structure. Alternatively, if both sets of indicators were to produce a taxon with a similar base rate, this would provide strong support for the existence of the taxon.

*Wechsler Intelligence Scales (IQ; d = 0.41–0.85).* Two sets of indicators were derived from either the Wechsler Adult Intelligence Scale—Third Edition (WAIS–III; Wechsler, 1997a) or the Wechsler Intelligence Scale for Children—Third Edition (WISC–III; Wechsler, 1991). The first set consisted of four index scores provided by each intellectual measure (IQ Indices). Although the index scores are not identical across measures, they are quite similar in composition, are highly correlated, and often include the same subtests. For example, the Working Memory index on the WAIS–III consists of Digit Span, Arithmetic, and Letter–Number Sequencing, whereas the WISC–III Freedom From Distractibility index consists of only Digit Span and Arithmetic. The second indicator set consisted of the IQ subtests thought to be most sensitive to the attention and working memory deficits observed in ADHD (IQ Subtests). This set included the Digit Span, Arithmetic, Digit Symbol (or Coding on the WISC–III), and Symbol Search subtests (Kaufman, 1994). Approximately 27% of individuals included in the IQ Indices set received the WAIS–III and 73% the WISC–III, and these percentages did not significantly differ as a function of ADHD clinical diagnosis, $\chi^2(2, N = 394) = 1.62$, $p = .44$. Similar results were obtained for the IQ Subtests indicator set (26% received the WAIS–III, and 74% the WISC–III), $\chi^2(2, N = 390) = 1.46$, $p = .48$. Previously reported reliabilities for the indices and subtests used in these indicator sets ranged from .70 to .96 (Wechsler, 1991, 1997b).

*Neuropsychological measures of processing speed (NP–PS; d = 0.48–0.85).* Neuropsychological measures of processing speed were included because of suggestions that individuals who exhibit symptoms of sluggish cognitive tempo may identify a distinct subset of individuals diagnosed as ADHD–I (Beauchaine, 2003; Carlson & Mann, 2002). In this context, measures of processing speed were conceptualized as behavioral indicators of sluggish cognitive tempo (Sullwold, 2004). These measures included the Coding and Symbol Search subtests of the Wechsler scales and hit reaction time from the Conners CPT.

*Mixture.* The final indicator set consisted of measures chosen from all of the previous indicator sets. In general, measures were chosen to have little overlap with other measures outside of their original indicator set and to have high validity for the separation of ADHD from control participants. This was done to examine the generalizability of any identified taxa, to evaluate an indicator set including variables with the highest estimated validity, and to include an indicator set with a larger number of indicators than other sets. Having a large number of indicators may produce clearer results (Schmidt et al., 2004). The indicators included CPARS attention/cognitive symptoms, CPARS hyperactivity/impulsivity, Achievement—math calculation, Achievement—written expression, CPT Attention, CPT Impulsivity, CPT Vigilance, IQ—Working Memory/Freedom From Distractibility, and IQ—Processing Speed.

To examine whether ADHD–I is a distinct subtype relative to ADHD–C, all control cases were removed from the sample (ADHD subsample) and two indicator sets were examined, CPARS–H/I and NP–PS. The CPARS–H/I set was included on the basis of findings of previous latent class analyses suggesting that the distinction between ADHD subgroups occurs for hyperactive/impulsive items (Neuman et al., 1999). The NP–PS set was included to examine the hypothesis that a subset of individuals with ADHD–I represents a distinct category of individuals with sluggish cognitive tempo (Beauchaine, 2003; Carlson & Mann, 2002).

### Analytic Strategy

Taxometric methodologists have recommended using multiple taxometric procedures as well as other techniques for identifying latent classes (Schmidt et al., 2004). In particular, latent class analysis (LCA) can be a useful first step in the search for latent classes/taxa (Solomon, Haaga, & Arnow, 2001). LCA is particularly well suited for discriminating between two or more latent classes, whereas taxometric procedures are best able to discriminate between one or two latent classes. The presence of more than two classes in a data set may cause taxometric procedures to perform poorly, as each latent class may have a low base rate as well as potentially distinctive profiles on the indicators. Therefore, we began by subjecting indicator sets to LCA. Classification error rate, the Bayesian information criterion (BIC), and the Akaike

information criterion (AIC) were examined to determine the presence of two or more latent classes (McCutcheon, 2002). If no strong evidence of more than two latent classes emerged in LCA results, indicator sets were deemed suitable for taxometric procedures.

Taxometric analyses were performed using program code developed by J. Ruscio (2004b) for the R programming environment. Three taxometric procedures were used in the present study: mean above minus below a cut (MAMBAC), maximum eigenvalue (MAXEIG), and latent mode (LMODE). Examination of three procedures serves as an additional consistency test, particularly because these are three relatively dissimilar methods (Schmidt et al., 2004). Several statistics were used to examine taxometric results: the variability of the base rate identified in MAXEIG analyses, the weighted-fit *d* statistic provided in J. Ruscio's (2004a) implementation of MAXEIG and MAMBAC, the nose count test based on raters' evaluations of taxometric graphs (for an extended discussion of this test, see Schmidt et al., 2004), and the convergence of base rates across taxometric procedures. To calculate the weighted-fit *d* index, 20 simulation samples were requested (10 taxonic, 10 dimensional) for each analysis. This index, and the simulation procedure more generally, were recently used to evaluate the latent structure of psychopathic personality (Marcus, John, & Edens, 2004), and a recent Monte Carlo study demonstrated that a closely related index, the comparative curve fit index, performs quite well at similar sample sizes and data parameters to the present study (J. Ruscio, Ruscio, & Meron, in press).

For the nose count test, eight raters unfamiliar with taxometric procedures rated the individual output indicator graphs resulting from each analysis, the average graph derived from the mean of all indicator graphs, and an overlay graph that presents the average graph superimposed on the average results ($\pm 1$ *SD*) derived from 10 simulated taxonic and 10 simulated dimensional data sets with roughly equivalent characteristics to the research data. For each taxometric procedure, raters were first presented with classic taxonic and dimensional graphs and given a verbal description of the general shape of taxonic and dimensional plots (i.e., MAMBAC taxonic plots are generally convex and shaped like a hill; dimensional plots are generally concave and shaped like a valley). Then, raters were shown a page with eight suboptimal simulated data sets (four taxonic, four dimensional). These data sets were simulated to represent the presumed ends of the distribution of analytic parameters for the research indicator sets. Thus, four simulated data sets (two taxonic, two dimensional) were based on $N = 300$, three indicators, base rate of taxon = .70, sample $r = .40$, nuisance $r = .20$, and mild skew; the other four simulated data sets (two dimensional, two taxonic) were based on $N = 380$, four indicators, base rate of taxon = .50, sample $r = .50$, nuisance $r = .10$, and no skew. Raters were asked to place a number (1 = taxonic, 2 = dimensional, or 3 = unspecified) over each average simulation graph. These example simulated data sets were used exclusively to educate raters about the shapes of taxonic and dimensional graphs under optimal and suboptimal analytic conditions and are distinct from simulated data used to examine the performance of decision rules. These simulated data sets are also different from the 20 simulated data sets (10 taxonic, 10 dimensional) generated as part of the analyses of research data for the computation of the fit *d* statistic and overlay graphs.

After rating example simulation data, raters were shown the individual and average indicator graphs and asked to make the same judgments they had made for the simulated graphs. Finally, they were shown the overlay graphs and asked to choose which graph contained lines that most closely matched or to indicate that they were not able to choose. For all decisions, raters were encouraged to make a choice between dimensional and taxonic and to reserve the unspecified category for instances in which they were unable to make a choice. Ratings were done separately for all MAMBAC graphs, followed by all MAXEIG graphs, and finally for all LMODE graphs. LMODE produces only individuals graphs and overlay graphs, because each indicator set generates only one individual graph, obviating the need for averaging.

Convergence of base-rate estimates should be observed when results have consistently identified a taxon. Although there are no guidelines for interpreting the absolute magnitude of convergence, widely divergent results do not support the taxonic conjecture. Therefore, this decision rule was examined only in situations where other rules suggested possible taxonicity.

A simulation study was undertaken to determine whether decision rules described above would provide accurate results for each indicator set and taxometric analysis.[1] For each indicator set, two data sets were simulated (one dimensional, one taxonic). Each pair of simulated data was generated using the specifications of one of the eight original indicator sets from the total sample. The CPARS–I and CPARS–H/I indicator sets were not included in this simulation study as their characteristics are largely overlapping with the CPARS indicator set. Specifications included sample size, number of indicators, taxon base rate, indicator validity, average interindicator total sample correlation, nuisance correlations, and skewness. The taxon base rate, indicator validity, and nuisance correlations were derived from preliminary analyses of the total sample data. Skewness values provided to the simulation program were estimated using guidelines presented in the manual (J. Ruscio, 2004a) and based on the actual skewness values of the indicator sets. Simulated data were analyzed using each taxometric procedure to determine whether any procedures performed poorly under conditions similar to those seen in actual data. Each of the decision rules was evaluated, as decision rules may provide variable results across taxometric procedures and indicator sets. The nose count test was implemented by having four raters unfamiliar with taxometric procedures evaluate individual, average, and overlay taxometric graphs. The reliability of ratings was determined using the kappa coefficient to determine whether raters could reliably distinguish between taxonic and dimensional graphs.

## Results

### *LCA*

Successive latent class models were tested with solutions ranging from one to six classes for each indicator set. Six was chosen as an upper limit to ensure that models considered the possibility of there being even more classes than identified in previous work (Milich et al., 2001). Classification error rates tended to increase from one- to six-class solutions, with the biggest rise occurring

[1] CPARS–I, CPARS–H/I, and NP–PS indicator sets were not included in the simulation study.

from the one- to two-class solution. The exception to this pattern occurred for the IQ Subtests indicator set, where a local minimum was reached at the four-class step, and the NP–PS set, where the minimum was reached at the four- and five-class steps. AIC values tended to decline from one- to six-class solutions, with the exception of the IQ Indices and NP–PS indicator sets, where a local minimum was reached at the five-class step. BIC values showed local minima for all indicator sets. However, the number of classes identified varied across indicator sets from two to five, and in several cases the minimum was extremely close to values for adjacent solutions.

To further examine the possibility of two or more latent classes, LCA classifications were compared with clinical diagnostic classifications for each indicator set. In each set, the two-class classification was compared with a variable coded 1 = non-ADHD, 2 = any ADHD subtype. The three-class classification was compared with a variable coded as 1 = non-ADHD control, 2 = ADHD–I, 3 = ADHD–C. The four-class solution was compared with a variable coded 1 = non-ADHD control, 2 = ADHD–I, 3 = ADHD–C, 4 = any ADHD subtype with a learning disability diagnosis. Agreement of LCA results and clinical diagnoses was evaluated by computing the percentage of cases for which both variables agreed. In every case, the tables were structured to yield the highest possible level of agreement. Agreement was poor for all classifications, with two-class solutions showing the highest agreement (average percentage agreement: two-class solution = 55.2, three-class solution = 40.6, four-class solution = 33.2).

Overall, these results indicate no clear or consistent winner in terms of the number of latent classes. Given the findings for the AIC and classification errors criteria, it is possible that these results indicate a large number of latent classes (i.e., six or more). However, this would be inconsistent with previous theory regarding the number of ADHD subtypes (Milich et al., 2001). Additionally, because these results do not suggest a definitive number of latent classes and LCA is not useful for distinguishing between one or two latent classes (i.e., dimensional vs. categorical representations), results may be consistent with a latent dimension. Finally, the lack of a definitive class solution indicates that taxometric procedures should not have been substantially influenced by the presence of multiple, low base rate taxa. Therefore, all indicator sets were submitted to taxometric procedures.

### *Simulation Analyses*

Table 3 presents the results of the simulation study. Inspection of the bottom portion of the table reveals low reliability for ratings of individual and average MAMBAC and MAXEIG graphs (MAMBAC individual mean κ = .32, range = .22–.45; MAMBAC average mean κ = .36, range = .31–.46; MAXEIG

Table 3
*Simulation Study Results*

| | MAMBAC | | | | | | MAXEIG | | | | | | LMODE | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rating % | | | Base rate | | | Rating % | | | Base rate | | | Rating % | | | |
| Indicator set | Tax | Dim | Unsp | *M* | *SD* | Fit (*d*) | Tax | Dim | Unsp | *M* | *SD* | Fit (*d*) | Tax | Dim | Unsp | Base rate (*M*) |
| ACH dimensional | 0 | 100 | 0 | .53 | .04 | 5.3 | 0 | 100 | 0 | .52 | .27 | 2.9 | 0 | 100 | 0 | .50 |
| ACH taxonic | 100 | 0 | 0 | .39 | .03 | −6.5 | 100 | 0 | 0 | .27 | .03 | −2.6 | 100 | 0 | 0 | .60 |
| CPARS dimensional | 0 | 100 | 0 | .54 | .03 | 11.2 | 0 | 100 | 0 | .19 | .11 | 3.0 | 0 | 100 | 0 | .56 |
| CPARS taxonic | 100 | 0 | 0 | .44 | .06 | −2.3 | 100 | 0 | 0 | .12 | .06 | −2.3 | 100 | 0 | 0 | .40 |
| CPT Factors dimensional | 0 | 100 | 0 | .41 | .02 | 6.9 | 25 | 75 | 0 | .19 | .07 | −0.4 | 0 | 100 | 0 | .53 |
| CPT Factors taxonic | 100 | 0 | 0 | .39 | .07 | 0.8 | 100 | 0 | 0 | .07 | .01 | −3.4 | 50 | 50 | 0 | .64 |
| CPT Validity dimensional | 0 | 100 | 0 | .63 | .03 | 16.0 | 0 | 100 | 0 | .66 | .21 | 1.4 | 0 | 100 | 0 | .47 |
| CPT Validity taxonic | 100 | 0 | 0 | .50 | .05 | 3.5 | 100 | 0 | 0 | .57 | .05 | −2.3 | 100 | 0 | 0 | .31 |
| IQ Indices dimensional | 0 | 100 | 0 | .53 | .05 | 6.6 | 0 | 100 | 0 | .56 | .25 | 4.1 | 0 | 100 | 0 | .50 |
| IQ Indices taxonic | 100 | 0 | 0 | .40 | .04 | −1.5 | 100 | 0 | 0 | .44 | .08 | −1.5 | 100 | 0 | 0 | .47 |
| IQ Subtests dimensional | 0 | 100 | 0 | .51 | .02 | 5.4 | 0 | 100 | 0 | .31 | .08 | 2.6 | 0 | 100 | 0 | .50 |
| IQ Subtests taxonic | 75 | 25 | 0 | .43 | .04 | 1.4 | 100 | 0 | 0 | .38 | .05 | −2.6 | 100 | 0 | 0 | .48 |
| Mixture dimensional | 0 | 100 | 0 | .53 | .04 | 1.2 | 0 | 100 | 0 | .53 | .18 | 2.5 | 0 | 100 | 0 | .50 |
| Mixture taxonic | 25 | 75 | 0 | .46 | .08 | 1.1 | 100 | 0 | 0 | .29 | .08 | −1.1 | 0 | 100 | 0 | .42 |

| | Reliability of ratings | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Individual graphs | | | Average graphs | | Overlay graphs | | |
| | MAMBAC | MAXEIG | LMODE | MAMBAC | MAXEIG | MAMBAC | MAXEIG | LMODE |
| Avg. κ | .32 | .57 | 1.00 | .46 | .63 | .93 | .99 | .92 |

*Note.* Fit *d* = weighted fit *d* statistic; negative values favor taxonic results, and positive values favor dimensional results. Rating % refers to the percentage of overlay graphs rated as taxonic (tax), dimensional (dim), or unspecified (unsp). MAMBAC = mean above minus below a cut; MAXEIG = maximum eigenvalue; LMODE = latent mode; ACH = Achievement tests from the Woodcock–Johnson Tests of Ability (3rd ed.); CPARS = Conners Parent ADHD Rating Scale; CPT Factors = untransformed factor scores derived from the Conners Continuous Performance Test (CPT); CPT Validity = Conners CPT variables with the highest validity for separation of clinical groups; IQ Indices = index scores from the Wechsler Adult Intelligence Scale—III (WAIS–III) and Wechsler Intelligence Scale for Children—III (WISC–III); IQ Subtests = IQ subtests from the WAIS–III and WISC–III; Mixture = selected measures with the highest clinical validity from the other indicator sets.

individual mean κ = .57, range = .30–.88; MAXEIG average mean κ = .63, range = .23–.86). However, ratings of overlay MAMBAC and MAXEIG graphs had excellent reliability (MAMBAC mean κ = .93, range = .71–1.00; MAXEIG mean κ = .99, range = .88–1.00). Both individual and overlay LMODE graphs produced excellent reliability (individual mean κ = 1.00, range = 1 to 1; overlay mean κ = .92, range = .73–1.00). On the basis of the higher reliability of ratings for MAMBAC and MAXEIG overlay graphs and to reduce redundancy, only results from these ratings are presented.

Results of the simulation study indicated that ratings produced clear and interpretable results for most indicator sets and taxometric procedures (i.e., 87.5% of analyses rated with 100% agreement/confidence). The exceptions occurred for ratings of MAMBAC analyses on the Mixture indicator set and LMODE analyses on the CPT Factors and Mixture indicator sets. The standard deviation of base-rate estimates derived from MAXEIG analyses produced good separation between dimensional and taxonic data for all indicator sets except IQ Subtests. The weighted-fit *d* statistic performed well for most analyses. The exceptions occurred for MAMBAC analyses on the CPT Factors, CPT Validity, IQ Subtests, and Mixture indicator sets and MAXEIG analyses on the CPT Factors indicator set. For these analyses, the weighted-fit *d* statistic typically showed good separation, but both values—simulated dimensional and taxonic—were in the same direction. Decision rules for this subset of analyses should be interpreted with caution on the basis of their poor performance in the simulation study.

### *Taxometric Analyses*

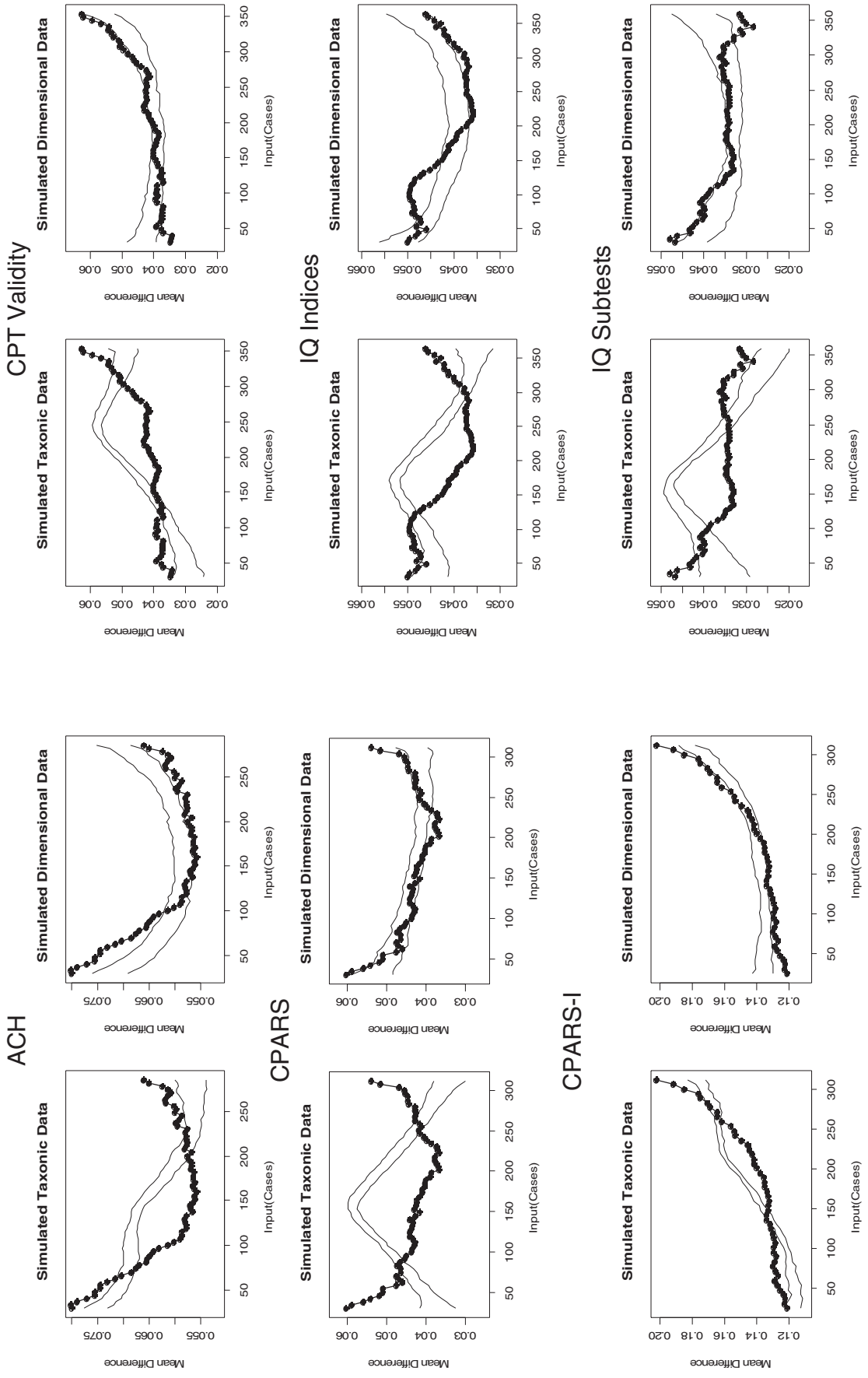
Table 4 presents classification decisions (taxonic vs. dimensional) and base rates, separately for each sample, indicator set, and taxometric analysis (for an extended discussion of the three taxometric procedures used in the present study, see Cole, 2004; Schmidt et al., 2004).

*MAMBAC.* MAMBAC was performed using all indicators as separate output variables. To increase statistical power, a composite input indicator was created by summing all of the remaining indicators. MAMBAC plots that show a convex shape suggest taxonic structure, and plots with a concave shape suggest dimensional structure. Figure 1 presents MAMBAC overlay graphs for

Table 4
*Taxometric Statistics, Separately for Each Sample, Taxometric Procedure, and Indicator Set*

| | MAMBAC | | | | | | MAXEIG | | | | | | LMODE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rating % | | | Base rate | | | Rating % | | | Base rate | | | Rating % | | | | |
| Indicator set | Tax | Dim | Unsp | *M* | *SD* | Fit (*d*) | Tax | Dim | Unsp | *M* | *SD* | Fit (*d*) | Tax | Dim | Unsp | Base rate *M* Range | |
| Total sample | | | | | | | | | | | | | | | | | |
| ACH | 0 | 100 | 0 | .44 | .04 | 3.3 | 0 | 87.5 | 12.5 | .28 | .24 | 1.8 | 0 | 100 | 0 | .50 | .22 |
| CPARS | 0 | 100 | 0 | .47 | .06 | 11.7 | 0 | 100 | 0 | .28 | .14 | 5.5 | 75 | 25 | 0 | .63 | .35 |
| CPARS–I | 0 | 100 | 0 | .62 | .02 | 3.3 | 0 | 75 | 25 | .82 | .06 | 1.7 | 0 | 100 | 0 | .48 | .34 |
| CPARS–H/I | 0 | 100 | 0 | .40 | .04 | 4.5 | 0 | 100 | 0 | .16 | .04 | 1.6 | 0 | 100 | 0 | .56 | .40 |
| CPT Factors | 0 | 100 | 0 | .33 | .01 | 4.8 | 62.5 | 0 | 37.5 | .15 | .09 | 0.2 | 0 | 100 | 0 | .61 | .46 |
| CPT Validity | 0 | 100 | 0 | .61 | .02 | 8.4 | 87.5 | 12.5 | 0 | .55 | .18 | −1.4 | 0 | 87.5 | 12.5 | .49 | .12 |
| IQ Indices | 0 | 100 | 0 | .44 | .02 | 5.4 | 0 | 87.5 | 12.5 | .33 | .12 | 2.2 | 0 | 100 | 0 | .53 | .20 |
| IQ Subtests | 0 | 100 | 0 | .43 | .04 | 6.1 | 0 | 87.5 | 12.5 | .40 | .20 | 0.5 | 0 | 100 | 0 | .49 | .09 |
| NP–PS | 0 | 100 | 0 | .43 | .07 | 7.7 | 0 | 100 | 0 | .47 | .26 | 2.0 | 12.5 | 87.5 | 0 | .50 | .07 |
| Mixture | 0 | 100 | 0 | .39 | .07 | 2.9 | 12.5 | 87.5 | 0 | .30 | .16 | 1.6 | 37.5 | 62.5 | 0 | .60 | .30 |
| ADHD–C subsample | | | | | | | | | | | | | | | | | |
| ACH | 0 | 100 | 0 | .44 | .03 | 3.0 | 0 | 87.5 | 12.5 | .28 | .25 | 2.2 | 0 | 100 | 0 | .50 | .22 |
| CPARS | 0 | 100 | 0 | .52 | .04 | 9.0 | 0 | 87.5 | 12.5 | .46 | .35 | 4.1 | 50 | 50 | 0 | .53 | .07 |
| CPARS–I | 0 | 100 | 0 | .61 | .02 | 2.6 | 37.5 | 37.5 | 25 | .16 | .05 | −0.1 | 0 | 100 | 0 | .48 | .45 |
| CPARS–H/I | 0 | 100 | 0 | .39 | .04 | 8.9 | 0 | 100 | 0 | .18 | .06 | 1.1 | 12.5 | 87.5 | 0 | .56 | .38 |
| CPT Factors | 0 | 100 | 0 | .36 | .02 | 7.6 | 0 | 100 | 0 | .17 | .09 | 2.3 | 0 | 100 | 0 | .61 | .44 |
| CPT Validity | 0 | 100 | 0 | .59 | .02 | 5.9 | 0 | 100 | 0 | .73 | .11 | −2.9 | 0 | 100 | 0 | .46 | .27 |
| IQ Indices | 25 | 75 | 0 | .44 | .03 | 2.6 | 0 | 100 | 0 | .31 | .09 | 2.8 | 0 | 100 | 0 | .53 | .22 |
| IQ Subtests | 0 | 100 | 0 | .58 | .05 | 4.1 | 12.5 | 75 | 12.5 | .41 | .20 | 1.9 | 0 | 100 | 0 | .50 | .17 |
| NP–PS | 0 | 100 | 0 | .49 | .04 | 7.9 | 0 | 100 | 0 | .47 | .22 | 1.6 | 12.5 | 87.5 | 0 | .50 | .03 |
| Mixture | 0 | 100 | 0 | .43 | .04 | 3.3 | 37.5 | 62.5 | 0 | .42 | .21 | −0.3 | 37.5 | 62.5 | 0 | .46 | .04 |
| ADHD subsample | | | | | | | | | | | | | | | | | |
| CPARS–H/I | 0 | 100 | 0 | .37 | .02 | 2.6 | 12.5 | 87.5 | 0 | .18 | .06 | 1.1 | 0 | 100 | 0 | .53 | .35 |
| NP–PS | 0 | 100 | 0 | .51 | .03 | 6.8 | 0 | 100 | 0 | .61 | .17 | 2.0 | 0 | 100 | 0 | .49 | .12 |

*Note.* The final column refers to the range of base rates from the three taxometric procedures for each indicator set. MAMBAC = mean above minus below a cut; MAXEIG = maximum eigenvalue; LMODE = latent mode; tax = taxonic, dim = dimensional, unsp = unspecified (could not be classified as either taxonic or dimensional); fit = weighted-fit *d* statistic (negative values favor taxonic results, positive values favor dimensional results); ACH = Achievement tests from the Woodcock–Johnson Tests of Ability (3rd ed.); CPARS = Conners Parent ADHD Rating Scale (I = inattention items; H/I = hyperactivity/impulsivity items); CPT Factors = untransformed factor scores derived from the Conners Continuous Performance Test (CPT); CPT Validity = Conners CPT variables with the highest validity for separation of clinical groups; IQ Indices = index scores from the Wechsler Adult Intelligence Scale—III (WAIS–III) and Wechsler Intelligence Scale for Children—III (WISC–III); IQ Subtests = IQ subtests from the WAIS–III and WISC–III; NP–PS = neuropsychological measures of processing speed; Mixture = selected measures with the highest clinical validity from the other indicator sets; ADHD = attention-deficit/hyperactivity disorder (C = combined type).


CPT Validity
IQ Indices
IQ Subtests
ACH
CPARS
CPARS-I
Simulated Taxonic Data
Simulated Dimensional Data
Mean Difference
Input(Cases)

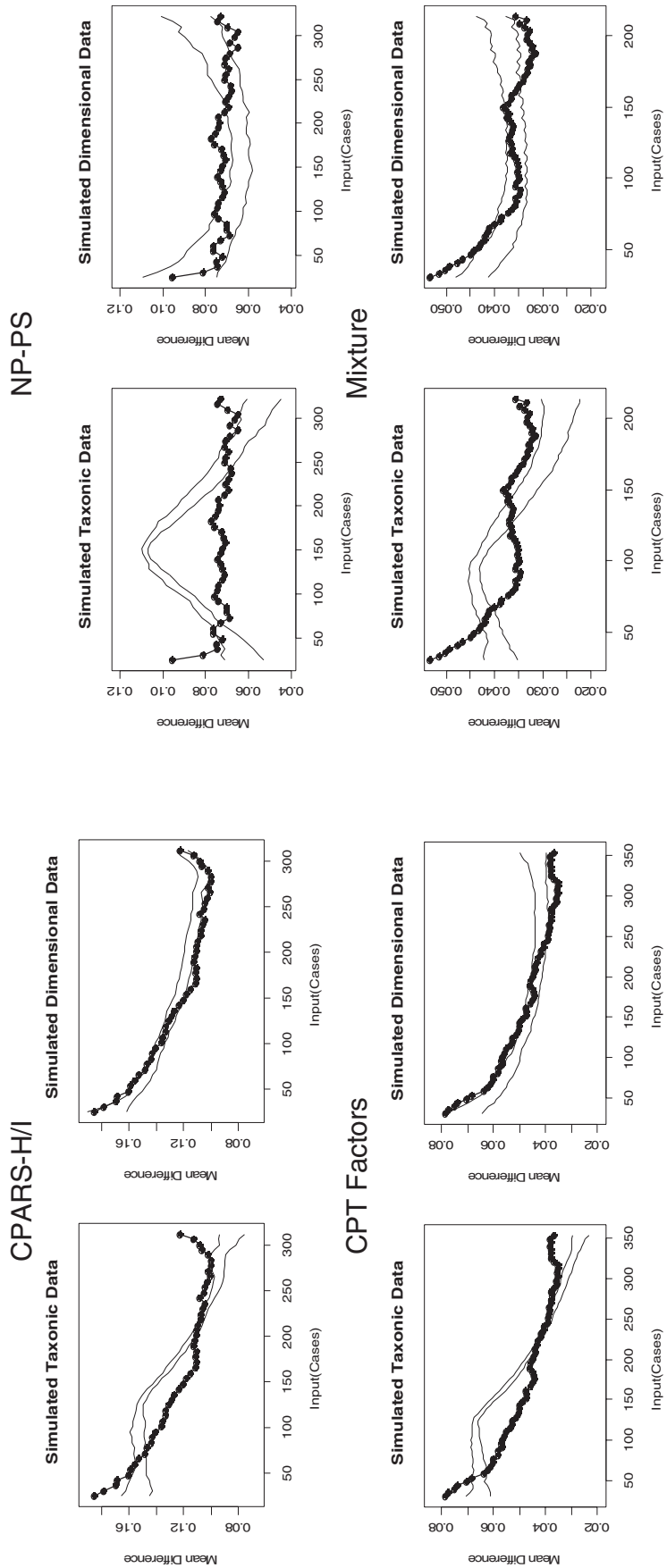



*Figure 1.* Mean above minus below a cut (MAMBAC) graphs for each indicator set in the total sample, including overlays with simulated taxonic data (left) and dimensional data (right). ACH = Achievement; CPARS = Conners Parent ADHD Rating Scale (I = inattention items; H/I = hyperactivity/impulsivity items); CPT = Conners Continuous Performance Test; NP–PS = neuropsychological measures of processing speed.

each indicator set in the total sample, and Table 4 presents the classification ratings, base-rate estimates, and fit indices. The analysis for each indicator set results in two overlay graphs. In the first graph, research data, averaged from each individual graph, are overlaying simulated taxonic data (left), and in the second graph research data are overlaying dimensional data (right). Decision rules and graphs indicated strong support for dimensional latent structure across all indicator sets. Inspection of Figure 1 indicates that many graphs showed a clear concave appearance. In situations where the appearance was not clearly concave, the match of research data to simulated dimensional data was dramatically better (weighted-fit $d = 3.3$ to 11.7, where highly positive $d$ values indicate dimensional structure). Similar results were obtained for ADHD–C and ADHD subsamples. Ratings and fit $d$ values indicated strong support for dimensional structure across all indicator sets.

*MAXEIG.* MAXEIG was performed using one input indicator and all remaining indicators as output variables. Analyses were permuted using each potential indicator as the input variable. Overlapping windows (90% overlap) were used for each slice of the input indicator. MAXEIG plots should show a convex shape (peaked) if the data are taxonic, whereas irregular or flat plots should be seen if the data are dimensional. Figure 2 presents MAXEIG overlay graphs for each indicator set in the total sample. The majority of decision rules and graphs indicated dimensional structure. The exceptions were some indications of taxonicity in the CPT Factors and CPT Validity indicator sets. For the CPT Validity indicator set, 87.5% of graphs were rated as taxonic and the weighted-fit $d$ statistic was negative (–1.4). However, the variability of base-rate estimates was quite high ($SD = .18$) relative to the variability of base-rate estimates from simulated taxonic data ($SD = .05$) and was more consistent with the variability of base-rate estimates from simulated dimensional data ($SD = .21$). Similarly, for the CPT Factors indicator set, 62% of ratings indicated taxonic structure, whereas 38% were unspecified. However, the variability of base-rate estimates was more consistent with dimensional structure (research data = .09, simulated taxonic = .01, and simulated dimensional = .07). All other indicator sets produced fairly clear evidence of dimensionality, including positive weighted-fit $d$ values ($d = 0.5$ to 5.5), a greater percentage of graphs rated as dimensional (87% to 100%), and highly variable base-rate estimates within an indicator set (*SD*s ranging from .12 to .24).

Analyses of the ADHD–C and ADHD subsamples produced a somewhat different pattern of results that indicated greater support of dimensional findings. All but one of the indicator sets yielded a greater percentage of dimensional ratings (62%–100%) than taxonic ratings (0%–38%). The exception was CPARS–I, where ratings (37.5% taxonic, 37.5% dimensional, 25% unspecified) and the weighted-fit $d$ were equivocal. The variability of base-rate estimates was also highly variable within indicator sets (*SD*s ranging from .05 to .35). The CPT Validity indicator set was the only data set to produce a highly negative weighted-fit $d$ statistic ($d = –2.9$), whereas the CPARS–I and Mixture indicator sets produced negative but small weighted-fit $d$ values (CPARS–I $d = –0.1$; Mixture $d = –0.3$). All other indicator sets yielded highly positive weighted-fit $d$ statistics ($d$s = 1.1 to 4.1).

*LMODE.* LMODE analyses examined all of the indicators in each indicator set simultaneously. Each LMODE analysis graphs the distribution of scores on the first principal component. Figure 3 presents LMODE overlay graphs with research data overlaying simulated dimensional and taxonic data. Research data are presented using the heavier line and remain consistent between both overlay graphs; simulation data are presented using two lighter lines and vary between graphs. Because LMODE generates a single base-rate estimate per indicator set, it is not meaningful to estimate the mean or standard deviation of the base-rate estimates. Results indicated moderate evidence for taxonic structure for the CPARS indicator set but for no other indicator sets. Inspection of the CPARS overlay graph reveals two small, but separate, bimodal peaks in the distribution. Raters were instructed to rate any graph containing two separate peaks with an observable downturn between peaks as taxonic. However, researchers have debated the significance of small peaks (Schmidt et al., 2004). Similarly, overlay graphs for the Mixture indicator set revealed two small peaks, possibly representing separate modes. However, raters were not able to adequately distinguish between simulated dimensional and taxonic data for this indicator set. Therefore, these peaks may be consistent with either taxonic or dimensional findings. Ratings and graphs for all other indicator sets were strongly consistent with dimensional structure. Results for the ADHD–C and ADHD subsamples were also suggestive of dimensionality, except that ratings for the CPARS indicator set were now equivocal (50% taxonic, 50% dimensional).

### *Consistency of Base-Rate Estimates*

The consistency of base-rate estimates for each analysis in which other decision rules suggested possible taxonicity was examined to confirm or reject this possibility. Table 4 presents the range of average base-rate estimates across analytic procedures, separately for each indicator set. In the total sample, only CPT Validity (range = .12), IQ Subtests (range = .09), and NP–PS (range = .07) showed modest consistency between base rates, whereas CPT Factors estimates were inconsistent (range = .46). MAXEIG and MAMBAC analyses using negatively skewed CPT indicators tended to produce low base-rate estimates (.15 to .36). Alternatively, analyses using positively skewed indicators produced higher base-rate estimates (.55–.73). This possibility was anticipated on the basis of simulation results examining the effects of skewness and further supports the dimensional hypothesis. Truly taxonic findings should have yielded consistent base-rate estimates, especially for highly similar indicator sets such as those based on CPT variables.

In the ADHD–C subsample, CPARS (range = .07) and Mixture (range = .04) showed good consistency between average base rates. However, these indicator sets showed little consistency in the total sample (CPARS, range = .35; Mixture, range = .30), and the other decision rules were either equivocal or indicated dimensionality. Thus, the evidence for taxonicity of these indicator sets is weak.

To further examine the possibility that CPARS, CPT Validity, CPT Factors, and Mixture indicator set analyses had identified an ADHD taxon, we examined agreement between LMODE, MAXEIG, and LCA classifications and ADHD diagnoses (1 = ADHD any type, 2 = no ADHD). Results indicated poor agree-

ment between these classifications and clinical diagnoses (CPARS, $\kappa$ = .13–.14; CPT Validity, $\kappa$ = .02–.11; CPT Factors, $\kappa$ = .04–.14; Mixture, $\kappa$ = .11–.18).

## Discussion

The primary goal of the present study was to investigate whether there was a category of cases showing changes in behavior or performance on neuropsychological tasks that would be consistent with a diagnosis of ADHD, versus behavioral and performance indicators ranging along a continuum. The investigation used a mix of subjective and performance measures that were chosen because of their demonstrated association with clinical diagnoses of ADHD. These multiple indicators of potential ADHD status were evaluated using four different statistical methods (MAMBAC, MAXEIG, LMODE, and LCA with continuous indicators) and multiple interpretive guidelines, with data drawn from a single clinical infrastructure with a high base rate of clinically diagnosed ADHD. The present findings, based on taxometric procedures, indicated that the core symptoms of ADHD may be best represented by a dimensional, not categorical, latent structure, consistent with the results of a previous taxometric study (Haslam et al., 2006). In the few cases where possible taxonic results were suggested, these findings were not confirmed by other taxometric procedures, decision rules, or results in a more distilled ADHD–C subsample. However, we should note that null findings may have occurred as a result of inadequate indicator selection, a significant problem for many taxometric investigations in which dimensional results are suggested.

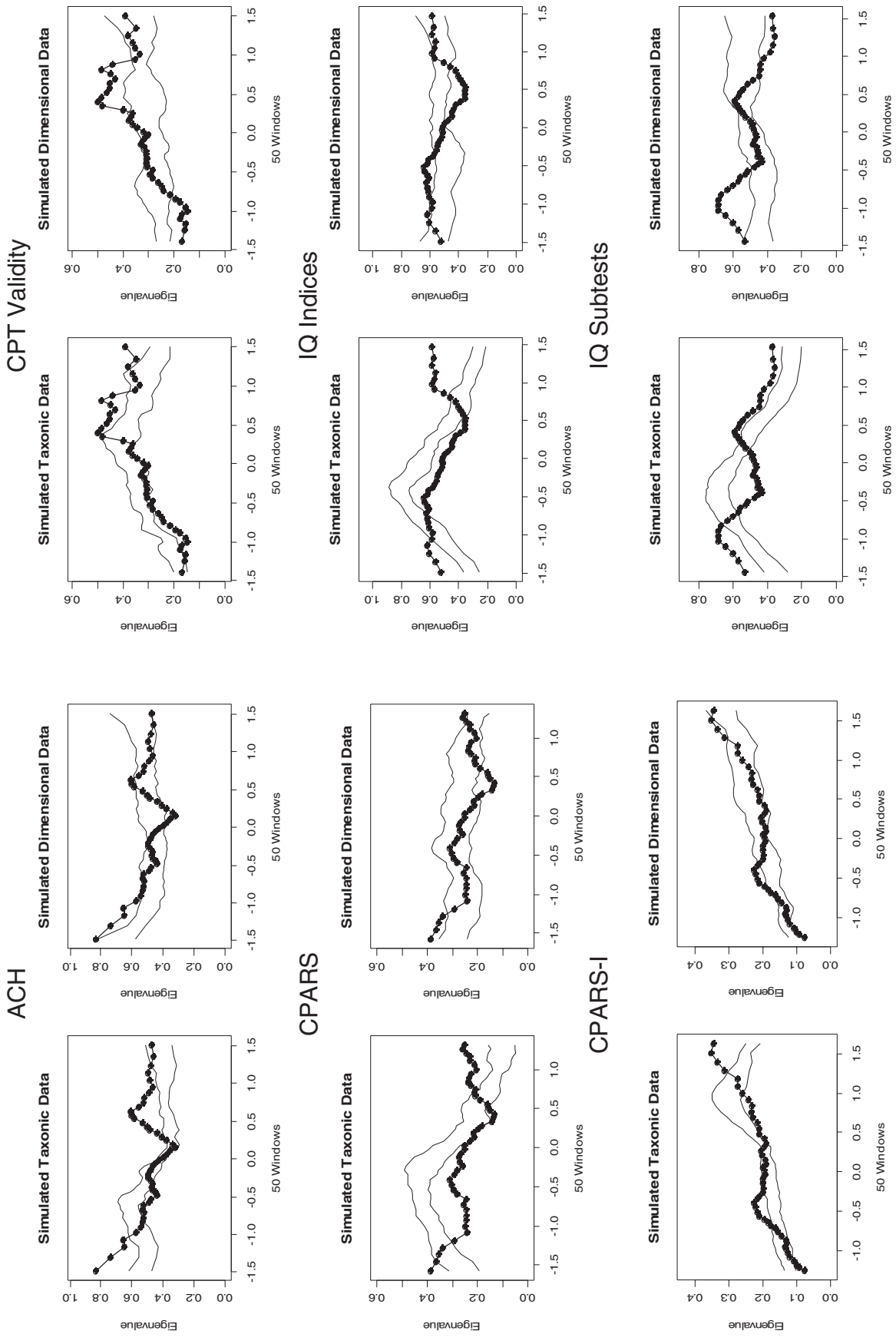
Similarly, there was no evidence for a categorical distinction among ADHD subtypes. This latter finding suggests that ADHD subtypes may simply represent different ways of parsing the dimensions of inattention and hyperactivity/impulsivity. However, the latter conclusion regarding a categorical distinction between ADHD subtypes is limited by the fact that only two indicator sets were examined. It may be that other indicators will demonstrate a categorical distinction among ADHD subtypes. Sluggish cognitive tempo has been suggested as a potential indicator of the categorical distinction between ADHD subtypes (Beauchaine, 2003). Included in the present research was an indicator set composed of processing speed measures, presumably tapping the cognitive aspects of this construct. However, future research is needed with subjective report and possibly other behavioral indicators of this construct in order to further evaluate the possibility of a distinct subset of ADHD–I individuals. Future studies should also focus on indicators with substantial validity for separating ADHD from non-ADHD cases, as this will ensure that an ADHD taxon is not missed owing to inadequate indicator validity.

The absence of taxonic findings in a taxometric analysis should not be mistaken for strong evidence of dimensionality (Beauchaine, 2003). The present results, as with any taxometric analysis, may have been due to misspecification or poor validity of indicators. The indicators were chosen specifically because they have demonstrated a strong relationship to clinical and research diagnoses of ADHD (Frazier et al., 2004). Furthermore, results from preliminary power analyses suggest that poor validity was not a significant problem. However, future research using other candidate indicators, including specific *DSM–IV* symptom criteria, is needed to address the problem of misspecification. It is possible, albeit unlikely, that there exists a group of individuals who differ qualitatively in terms of their *DSM–IV* symptomatology without showing related discontinuities on other behavioral or neuropsychological indicators. With this caveat in mind, failure to reject the null hypothesis of dimensionality has important implications for assessment and diagnosis.

Treating a dimensional, continuous construct as a categorical one hinders accurate assessment. Dichotomization of the continuum may reduce the reliability and validity of assessment (MacCallum, Zhang, Preacher, & Rucker, 2002). This is especially true if the dichotomy is implemented arbitrarily without reference to research suggesting the optimal cut point on the latent trait(s). Because *DSM* revisions have assumed a categorical diagnostic approach, no research has examined the number or composition of latent factors needed to accurately capture individuals at risk for functional impairments resulting from problems with attention, impulsivity, hyperactivity, behavioral dysregulation, and other constructs associated with ADHD. Without identification of the number and composition of these factors, the optimal cut point for initiating intervention remains unknown.

The categorical approach also results in significant loss of information because individuals with only mild, above threshold symptoms are labeled similarly to individuals with severe symptoms or impairment. Individuals with severe symptoms may show a greater range of impairment and may require different types of interventions or additional intervention strategies relative to individuals with fewer symptoms. Similarly, categorical diagnoses place individuals with mild, subthreshold symptoms into the non-ADHD category. It is possible that some of these individuals demonstrate functional impairment and would benefit from treatment. For example, in the PTSD literature, it has been noted that individuals with subthreshold symptoms often demonstrate significant impairment (Stein, Walker, Hazen, & Forde, 1997), and high degrees of impairment have also been documented for subsyndromal bipolar disorder (Lewinsohn, Seeley, & Klein, 2003). Unfortunately, these individuals are often denied services because they do not qualify for the PTSD or bipolar I diagnoses. It is possible that the rising and relatively high diagnostic prevalence of ADHD in the United States reflects an attempt to avoid the possibility of children not receiving services. However, the more sensitive threshold used to make categorical ADHD diagnoses may be generating a number of false positives (i.e., individuals without significant functional impairment or who do not need services). Future research should attempt to identify the number and content of dimensions needed to capture individuals at risk for functional impairments due to problems with attention and hyperactivity/impulsivity. It may be that only one dimension (ADHD symptom severity) consisting of problems with attention is necessary to accurately identify individuals needing intervention and disability services. If the structure of ADHD is truly dimensional, accurate identification and measurement of dimensions will lead to a reduction in diagnostic errors. Future studies should also distinguish identified dimensions from symptoms of other disorders, to limit diagnostic redundancy (Campbell & Fiske, 1959). Finally, future research should focus on other variables that, in combination with ADHD symptom severity, may predict disability, functional impairment, or treatment response.

If dimensional results are further replicated, then the present findings also have significant methodological implications for past


ACH
CPARS
CPARS-I
IQ Indices
IQ Subtests
CPT Validity
Simulated Taxonic Data
Simulated Dimensional Data
Eigenvalue
50 Windows

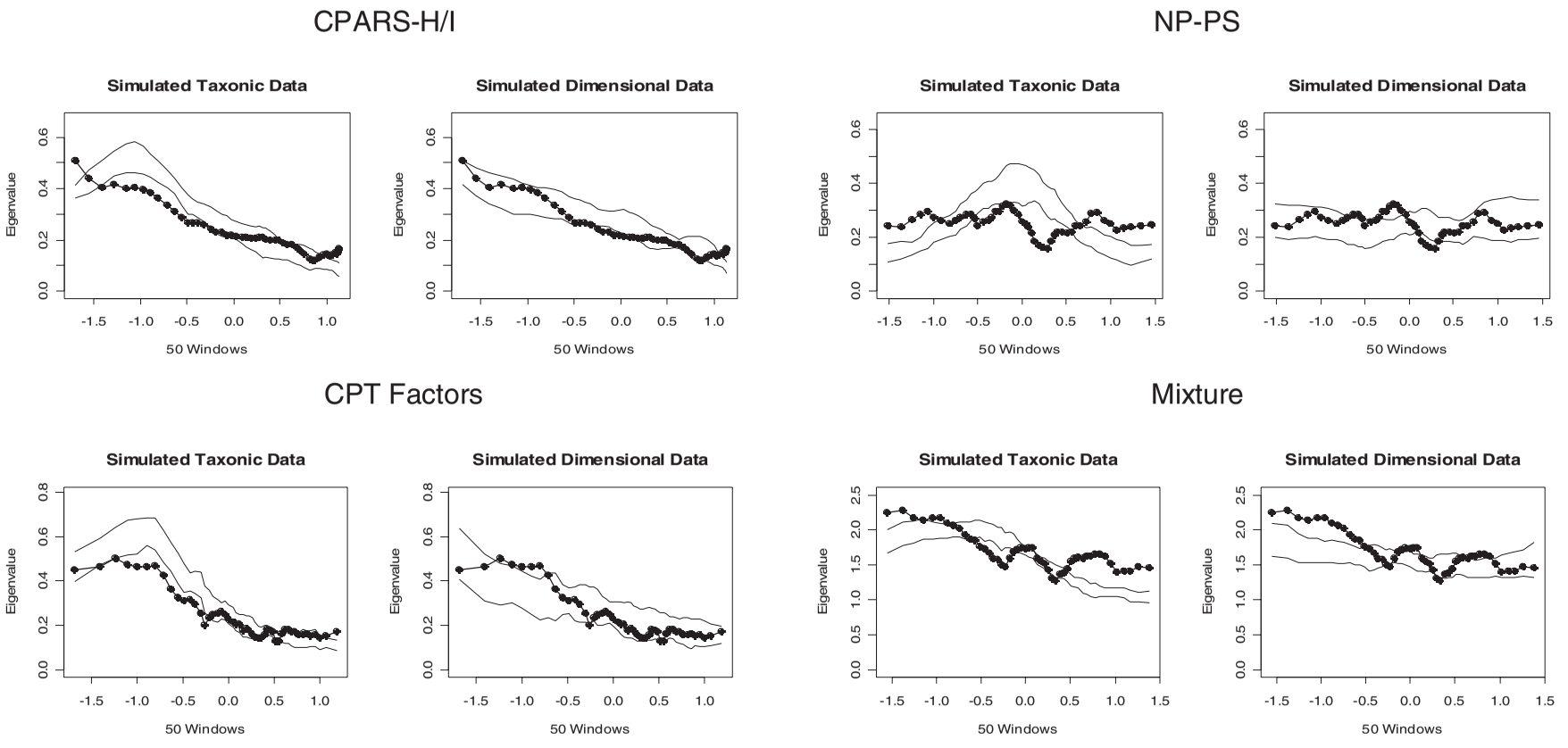
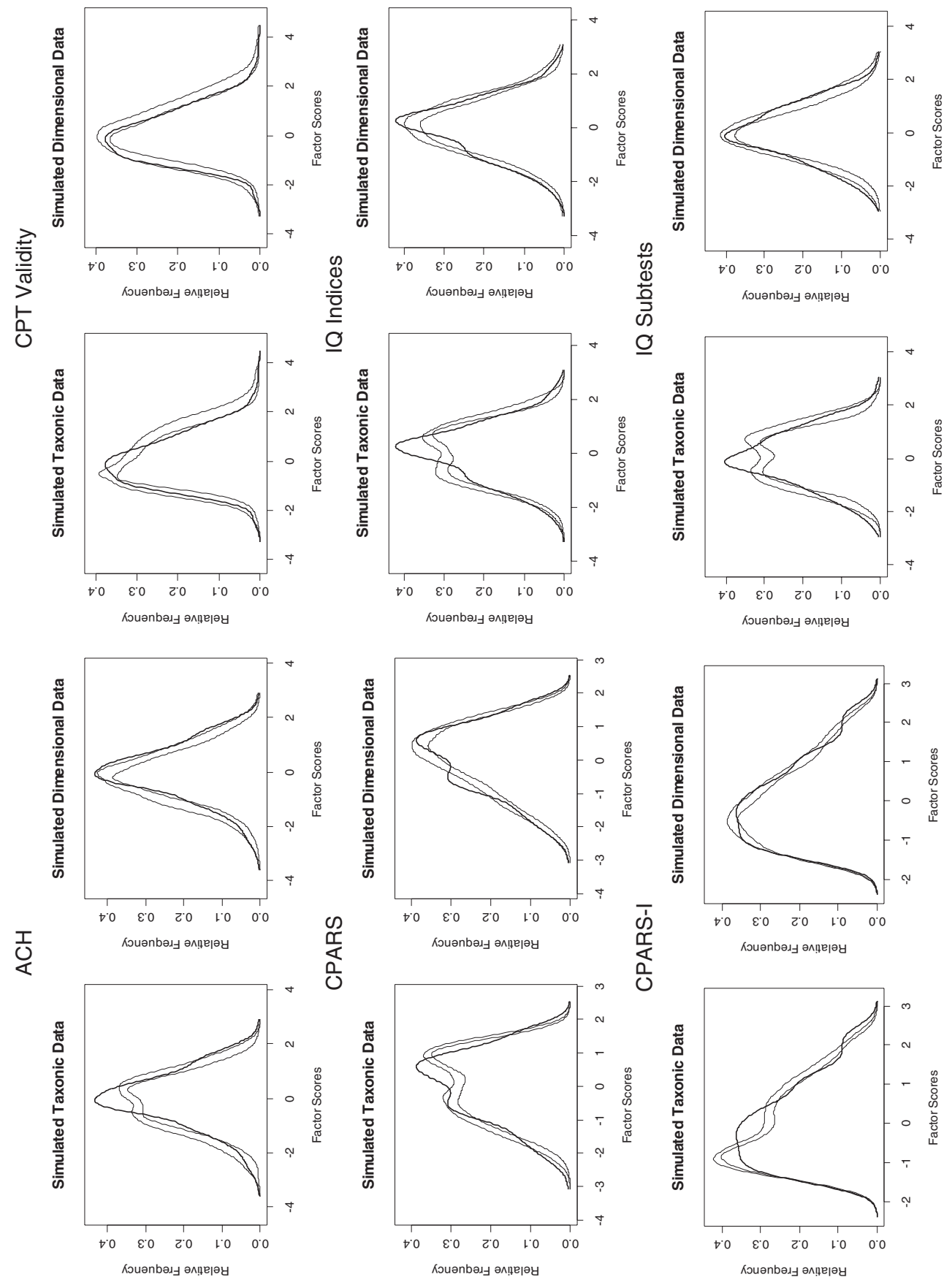



*Figure 2.* Maximum eigenvalue (MAXEIG) graphs for each indicator set in the total sample, including overlays with simulated taxonic data (left) and dimensional data (right). ACH = Achievement; CPARS = Conners Parent ADHD Rating Scale (I = inattention items; H/I = hyperactivity/impulsivity items); CPT = Conners Continuous Performance Test; NP–PS = neuropsychological measures of processing speed.


ACH
CPT Validity
CPARS
IQ Indices
CPARS-I
IQ Subtests
Simulated Taxonic Data
Simulated Dimensional Data
Relative Frequency
Factor Scores

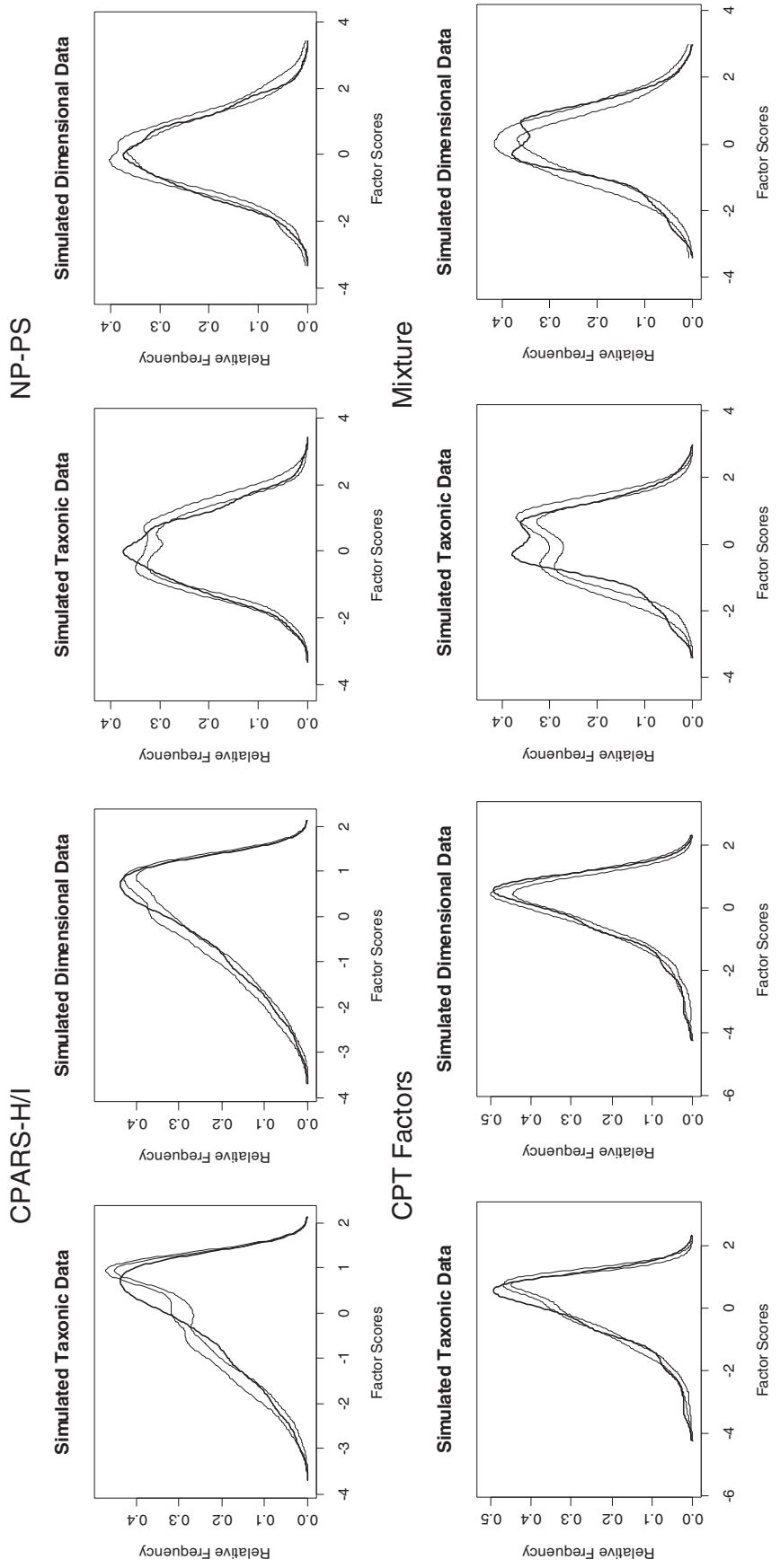



*Figure 3.* Latent mode (LMODE) graphs for each indicator set in the total sample, including overlays with simulated taxonic data (left) and dimensional data (right). ACH = Achievement; CPARS = Conners Parent ADHD Rating Scale (I = inattention items; H/I = hyperactivity/impulsivity items); CPT = Conners Continuous Performance Test; NP–PS = neuropsychological measures of processing speed.

and future ADHD research. Previous research has focused on group comparisons using *DSM* criteria. The present results indicate that previous studies should be interpreted as comparing ends of a continuum rather than distinct, qualitatively different entities. This may help to account for the inconsistent nature of findings in many areas: Some studies may show large effects due to comparisons of the extreme ends of the latent trait or traits, whereas other studies may show modest effects due to the inclusion of individuals falling in the middle of the continuum.

The present findings suggest that studies should shift from extreme group comparisons to correlational designs. These types of research designs can more accurately examine the relationship between ADHD symptoms and social, educational, or occupational impairment or treatment response. In particular, studies can examine whether the relationship between symptoms is linear or nonlinear and whether some dimensions of behavior (i.e., hyperactivity/impulsivity) show different relationships with impairment or treatment. The present findings also imply that future studies using a dimensional measurement approach are likely to have increased statistical power as a result of increased reliability and variability of scores.

The findings in this study, if supported in future studies, have clear implications regarding investigations of the etiology of ADHD. Previous work has suggested multiple cognitive, emotional, and behavioral deficits in ADHD. This has led some researchers to suggest a core deficit that results in these observable problems (Barkley, 1997; Quay, 1997). However, the multiplicity of identified deficits is unlikely to be completely explained by a single etiology, and recent genetic studies have failed to find a major or dominant genetic link (see Shastry, 2004, for a review). The present results are consistent with genetic findings and suggest that researchers should begin to focus on identifying the likely multifactorial etiology of ADHD and continue using methods that treat ADHD symptoms as polygenetic quantitative traits (Hudziak, 2001; Hudziak, Wadsworth, Heath, & Achenbach, 1999).

The notion of multiple consistency tests and replication serves as the foundation for taxometric procedures. The present study design was rich in consistency tests. These included the use of eight different indicator sets incorporating a range of methods, such as academic achievement data, behavioral ratings, and performance on both general neuropsychological tasks (e.g., Wechsler scales) and tasks more specifically implicated in ADHD (e.g., indices derived from continuous performance tests), as well as the use of three mathematically distinct taxometric procedures and a variety of different decision rules. The design involved 48 subsets of analyses, with many permutations of MAXEIG and MAMBAC analyses and decision rules underlying each indicator subset.

The choice of indicators was guided both by theory (Barkley, 1997; Quay, 1997) and by prior research into the effect sizes of group differences between ADHD and non-ADHD cases (Frazier et al., 2004), and the present analyses confirmed the high level of indicator validity before proceeding with the taxometric analyses. These selection procedures reduce the possibility that the taxometric models were misspecified, inadvertently including a set of indicators not relevant to the construct of ADHD. The investigation used a large, clinically referred sample that was likely to have a moderate base rate of ADHD. The investigation also used simulation results for comparison with research data to enhance the interpretability of findings. Use of these consistency tests partially offsets the concern of accepting the null hypothesis of dimensionality, because multiple analyses converged on dimensional results.

An additional strength of the present study was its use of an alternative statistical procedure, latent class analyses, for identifying distinct groups of individuals. LCA provides a safeguard for identifying whether there are more than two taxa mixed into the same data set. Other taxometric methods provide rigorous tests of the one-group (dimensional) versus two-group models but could produce ambiguous or misleading results if the data contain a mixture of more than two taxa.

There were several limitations to this study. The sample sizes for most of the total sample analyses were only slightly larger than those proposed by other researchers and developers of taxometric procedures (Meehl, 1995; Schmidt et al., 2004), and the sample sizes for three of the combined type sample analyses were below recommended thresholds. The indicator sets used in the present research did not include all 18 symptoms of ADHD as delineated in *DSM–IV*. Thus, as stated previously, a taxon may have been missed owing to indicator misspecification. However, we believe that the wide range of indicators used and previous research demonstrating substantial validity of these indicators reduces the likelihood of misspecification. Beauchaine (2003) recommended other candidate indicators for the identification of an ADHD taxon or differentiation of ADHD subtaxa, including indicators of sluggish cognitive tempo and urinary 3-methoxy-4-hydroxyphenylglycol, a norepinephrine metabolite. Some research has also suggested that frontal electroencephalogram activity may differentiate between individuals with and without ADHD (Monastra et al., 1999). Future research should examine these candidate *DSM–IV* and physiological indicators, ideally in tandem, to determine whether the present findings suggestive of dimensionality generalize across other important methods for assessing ADHD symptoms and correlates.

The present study also used a clinic-referred sample, which permitted a larger base rate of the clinical diagnosis of ADHD, thereby circumventing the problem of having too few taxon members for identification. However, samples that are more representative of the population permit greater generalization. Specifically, very large epidemiologic samples ($>$ 3,000 cases) sufficient to ensure adequate detection of very low base rates are needed. Unfortunately, obtaining these samples is extremely costly and typically prohibits the administration of multiple cognitive measures, thereby limiting indicator selection. Ultimately, determination of the latent structure of ADHD will rely on replication in both epidemiologic samples, in which a few carefully chosen subjective report indicators are collected, and clinical samples, in which large numbers of subjective report and objective indicators can provide comprehensive coverage of possible ADHD markers.

An additional limitation of the present study was that data from the WAIS–III and WISC–III were treated as a single indicator set. However, we do not believe this was a substantial problem, for two reasons. First, the indicators derived from both tests are highly similar. Second, because of the Flynn effect, the mixture of different versions is likely to have biased the present findings toward identification of a pseudotaxon and not a dimensional solution. Therefore, mixing these tests is unlikely to have negatively impacted the present results.

Overall, the present findings suggest that a dimensional approach to assessment and interpretation may be more appropriate,

at least in a clinical sample. Even if an ADHD taxon were to emerge using a different indicator set or sample, the present findings clearly challenge the widespread clinical practices of using cognitive ability tests, including specific subtest data (Kaufman, 1994), as well as continuous performance tests as methods for identifying a category of youths at risk for ADHD. Several studies have questioned the value of cognitive ability tests as diagnostic measures for ADHD (Watkins, Kush, & Glutting, 1997a, 1997b). The present findings suggest that the failure of these tests could be attributed to their not being valid indicators of ADHD status, despite clinical and theoretical arguments to the contrary. Alternatively, the present findings are also consistent with the possibility that there is no naturally occurring taxon of ADHD, and the growing body of evidence is more supportive of dimensional models of individual differences in attention and motor activity.

## References


Adams, H. E., & Cassidy, J. F. (1993). The classification of abnormal behavior. In P. B. Sutker & H. E. Adams (Eds.), *Comprehensive handbook of psychopathology* (2nd ed., pp. 3–25). New York: Plenum Press.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

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

Amor, L. B., Grizenko, N., Schwartz, G., Lageix, P., Baron, C., Ter-Stepanian, M., et al. (2005). Perinatal complications in children with attention-deficit hyperactivity disorder and their unaffected siblings. *Journal of Psychiatry & Neuroscience, 30,* 120–126.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unified theory of ADHD. *Psychological Bulletin, 121,* 65–94.

Barkley, R. A. (1998). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (2nd ed.). New York: Guilford Press.

Barkley, R. A. (2001). The inattentive type of ADHD as a distinct disorder: What remains to be done. *Clinical Psychology Science and Practice, 8,* 489–493.

Beauchaine, T. P. (2003). Taxometrics and developmental psychopathology. *Development and Psychopathology, 15,* 501–527.

Beauchaine, T. P., & Beauchaine, R. J. (2002). A comparison of maximum covariance and k-means cluster analysis in classifying cases into known taxon groups. *Psychological Methods, 7,* 245–261.

Beauchaine, T. P., & Waters, E. (2003). Pseudotaxonicity in MAMBAC and MAXCOV analyses of rating-scale data: Turning continua into classes by manipulating observer's expectations. *Psychological Methods, 8,* 3–15.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56,* 81–104.

Carlson, C. L., & Mann, M. (2002). Sluggish cognitive tempo predicts a different pattern of impairment in the attention deficit hyperactivity disorder, inattentive type. *Journal of Clinical Child and Adolescent Psychology, 31,* 123–129.

Carson, R. C. (1991). Dilemmas in the pathway of *DSM–IV*. *Journal of Abnormal Psychology, 100,* 302–307.

Castellanos, F. X., Giedd, J. N., Berquin, P. C., Walter, J. M., Sharp, W., Tran, T., et al. (2001). Quantitative brain magnetic resonance imaging in girls with attention-deficit/hyperactivity disorder. *Archives of General Psychiatry, 58,* 289–295.

Castellanos, F. X., Giedd, J. N., March, W. L., Hamburger, S. D., Vaituzis, A. C., Dickstein, D. P., et al. (1996). Quantitative brain magnetic resonance imaging in attention-deficit. *Archives of General Psychiatry, 53,* 607–616.

Chhabildas, N., Pennington, B. F., & Willcutt, E. G. (2001). A comparison of neuropsychological profiles of the *DSM–IV* subtypes of ADHD. *Journal of Abnormal Child Psychology, 29,* 529–540.

Cole, D. A. (2004). Taxometrics in psychopathology research: An introduction to some of the procedures and related methodological issues. *Journal of Abnormal Psychology, 113,* 3–9.

Conners, C. K. (1997). *Conners' Rating Scales—Revised.* North Tonawanda, NY: Multi-Health Systems.

Conners, C. K. (2000a). *Conners' CPT II: Continuous Performance Test II. Computer program for Windows.* North Tonawanda, NY: Multi-Health Systems.

Conners, C. K. (2000b). *Conners' CPT II: Continuous Performance Test II. Technical guide and software manual.* North Tonawanda, NY: Multi-Health Systems.

Dahlstrom, W. G. (1995). Pigeons, people, and pigeon-holes. *Journal of Personality Assessment, 64,* 2–20.

Faraone, S. V., Biederman, J., Weber, W., & Russell, R. L. (1998). Psychiatric, neuropsychological, and psychosocial features of *DSM–IV* subtypes of attention-deficit/hyperactivity disorder: Results from a clinically referred sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 37,* 185–193.

Frazier, T. W., Demaree, H. A., & Youngstrom, E. A. (2004). A meta-analysis of intellectual and neuropsychological test performance in attention-deficit/hyperactivity disorder. *Neuropsychology, 18,* 543–555.

Gibb, B. E., Alloy, L. B., Abramson, L. Y., Beevers, C. G., & Miller, I. W. (2004). Cognitive vulnerability to depression: A taxometric analysis. *Journal of Abnormal Psychology, 113,* 81–89.

Goodyear, P., & Hynd, G. (1992). Attention deficit disorder with (ADD–H) and without (ADD–W) hyperactivity: Behavioral and neuropsychological differentiation. *Journal of Child Clinical Psychology, 21,* 273–305.

Haslam, N. (1997). Evidence that male sexual orientation is a matter of degree. *Journal of Personality and Social Psychology, 73,* 862–870.

Haslam, N., Williams, B., Prior, M., Haslam, R., Graetz, B., & Sawyer, M. (2006). The latent structure of attention-deficit/hyperactivity disorder: A taxometric analysis. *Australian and New Zealand Journal of Psychiatry, 40,* 639–647.

Hinshaw, S. P. (2001). Is the inattentive subtype of ADHD a separate disorder? *Clinical Psychology: Science and Practice, 8,* 498–501.

Hinshaw, S. P., Carte, E. T., Sami, N., Treuting, J. J., & Zupan, B. A. (2002). Preadolescent girls with attention-deficit/hyperactivity disorder: II. Neuropsychological performance in relation to subtypes and individual classification. *Journal of Consulting and Clinical Psychology, 70,* 1099–1111.

Hudziak, J. J. (2001). The role of phenotypes (diagnoses) in genetic studies of attention-deficit/hyperactivity disorder and related child psychopathology. *Child and Adolescent Psychiatric Clinics of North America, 10,* 279–297.

Hudziak, J. J., Wadsworth, M. E., Heath, A. C., & Achenbach, T. M. (1999). Latent class analysis of Child Behavior Checklist attention problems. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 985–991.

Kaufman, A. S. (1994). *Intelligent testing with the WISC–III.* New York: Wiley.

Lahey, B. B. (2001). Should the combined and predominantly inattentive types of ADHD be considered distinct and unrelated disorders? Not now, at least. *Clinical Psychology: Science and Practice, 8,* 494–497.

Lewinsohn, P. M., Seeley, J. R., & Klein, D. N. (2003). Bipolar disorder in adolescents: Epidemiology and suicidal behavior. In B. Geller & M. P. DelBello (Eds.), *Bipolar disorder in childhood and early adolescence* (pp. 7–24). New York: Guilford Press.

Lockwood, K. A., Marcotte, A. C., & Stern, C. (2001). Differentiation of attention-deficit/hyperactivity disorder subtypes: Application of a neuropsychological model of attention. *Journal of Clinical and Experimental Neuropsychology, 23,* 317–330.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7,* 19–40.

Marcus, D. K., John, S. L., & Edens, J. F. (2004). A taxometric analysis of psychopathic personality. *Journal of Abnormal Psychology, 113,* 626–635.

McCutcheon, A. L. (2002). Basic concepts and procedures in single- and multiple-group latent class analysis. In J. A. Hagenaars & A. L. McCutcheon (Eds.), *Applied latent class analysis* (pp. 56–85). Cambridge, England: Cambridge University Press.

McGrew, K. S., & Woodcock, R. W. (2001). *Woodcock–Johnson III: Technical manual.* Itasca, IL: Riverside.

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist, 50,* 266–275.

Meehl, P. E. (2004). What's in a taxon? *Journal of Abnormal Psychology, 113,* 39–43.

Meehl, P. E., & Yonce, L. J. (1994). Taxometric analysis: I. Detecting taxonicity with two quantitative indicators using means above and below a sliding cut (MAMBAC procedure). *Psychological Reports, 74,* 1059–1274.

Meehl, P. E., & Yonce, L. J. (1996). Taxometric analysis: II. Detecting taxonicity using covariance of two quantitative indicators in successive intervals of a third indicator (MAXCOV procedure). *Psychological Reports, 78,* 1091–1227.

Milich, R., Balentine, A. C., & Lynam, D. R. (2001). ADHD combined type and ADHD predominantly inattentive type are distinct and unrelated disorders. *Clinical Psychology: Science and Practice, 8,* 463–488.

Monastra, V. J., Lubar, J. F., Linden, M., VanDeusen, P., Green, G., Wing, W., et al. (1999). Assessing attention deficit disorder via quantitative electroencephalography: An initial validation study. *Neuropsychology, 13,* 424–433.

Neuman, R. J., Todd, R. D., Heath, A. C., Reich, W., Hudziak, J. J., Bucholz, K. K., et al. (1999). Evaluation of ADHD typology in three contrasting samples: A latent class approach. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 25–36.

Quay, H. C. (1997). Inhibition and attention deficit hyperactivity disorder. *Journal of Abnormal Child Psychology, 25,* 7–13.

Ruscio, A. M., Ruscio, J., & Keane, T. M. (2002). The latent structure of posttraumatic stress disorder: A taxometric investigation of reactions to extreme stress. *Journal of Abnormal Psychology, 111,* 290–301.

Ruscio, J. (2004a). *Taxometric program documentation, R language.* Retrieved December 30, 2004, from the Elizabethtown College Web site: http://www.etown.edu/psychology/Faculty/Ruscio.htm

Ruscio, J. (2004b). *Taxometric programs in the R language.* Retrieved December 30, 2004, from the Elizabethtown College Web site: http://www.etown.edu/psychology/Faculty/Ruscio.htm

Ruscio, J., & Ruscio, A. M. (2004). Clarifying boundary issues in psychopathology: The role of taxometrics in a comprehensive program of structural research. *Journal of Abnormal Psychology, 113,* 24–38.

Ruscio, J., Ruscio, A. M., & Meron, M. (in press). Applying the bootstrap to taxometric analysis: Generating empirical sampling distributions to help interpret results. *Multivariate Behavioral Research.*

Schmidt, N. B., Kotov, R., & Joiner, T. E. (2004). *Taxometrics: Toward a new diagnostic scheme for psychopathology.* Washington, DC: American Psychological Association.

Shastry, B. S. (2004). Molecular genetics of attention-deficit hyperactivity disorder (ADHD): An update. *Neurochemistry International, 44,* 469–474.

Shaywitz, B. A., Fletcher, J. M., & Shaywitz, S. E. (1994). Issues in the definition and classification of attention deficit disorder. *Topics in Language Disorders, 14,* 1–25.

Shaywitz, B. A., Fletcher, J. M., & Shaywitz, S. E. (1995). Defining and classifying learning disabilities and attention-deficit/hyperactivity disorder. *Journal of Child Neurology, 10,* 50–57.

Solomon, A., Haaga, D. A. F., & Arnow, B. A. (2001). Is clinical depression distinct from subthreshold depressive symptoms? A review of the continuity issue in depression research. *Journal of Nervous and Mental Disease, 189,* 498–506.

Stein, M. B., Walker, J. R., Hazen, A. L., & Forde, D. R. (1997). Full and partial posttraumatic stress disorder: Findings from a community survey. *American Journal of Psychiatry, 154,* 1114–1119.

Sullwold, A. C. (2004). Relationship of sluggish cognitive tempo to symptoms of attention-deficit/hyperactivity disorder in clinically referred college students. *Dissertation Abstracts International, 64,* 5805B.

Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua.* Thousand Oaks, CA: Sage.

Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997a). Discriminant and predictive validity of the WISC–III acid profile among children with learning disabilities. *Psychology in the Schools, 34,* 309–319.

Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997b). Prevalence and diagnostic utility of the WISC–III SCAD profile among children with disabilities. *School Psychology Quarterly, 12,* 235–248.

Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children—Third edition.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997a). *Manual for the Wechsler Adult Intelligence Scale—Third edition.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997b). *WAIS–III WMS–III technical manual.* San Antonio, TX: Psychological Corporation.

Widiger, T. A., & Clark, L. A. (2000). Toward *DSM–V* and the classification of psychopathology. *Psychological Bulletin, 126,* 946–963.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III Tests of Ability.* Itasca, IL: Riverside.

Woods, S. P., Lovejoy, D. W., & Ball, J. D. (2002). Neuropsychological characteristics of adults with ADHD: A comprehensive review of initial studies. *Clinical Neuropsychologist, 16,* 12–34.

Received April 28, 2005
Revision received September 6, 2006
Accepted September 7, 2006 ■