Development and Psychopathology, **15** (2003), 501–527
Copyright © 2003 Cambridge University Press
Printed in the United States of America
DOI: 10.1017.S0954579403000270

Taxometrics and developmental psychopathology

THEODORE P. BEAUCHAINE
University of Washington

Abstract
Developmental psychopathologists have criticized categorical classification systems for their inability to account for within-group heterogeneity in biological, etiological, developmental, and cultural influences on behavior. Dichotomizing continuous scores of symptom severity is also inadvisable statistically. Perhaps because of a resulting wariness of categorizing, few explorations into the ontological status of traits or disorders as dimensional versus discrete have been conducted. It is argued here that the limitations of categorizing have little to do with the ontological status of traits and that developmental psychopathologists should be concerned with identifying discrete behavioral syndromes. Common taxometric methods for resolving discrete traits are described, and questions of concern to developmental psychopathologists are outlined that can be addressed through taxometrics studies. These include (a) identifying children who are at risk for future psychopathology, (b) identifying discrete subtypes within current diagnostic classes, (c) locating sensitive periods in the development of discrete pathological traits, (d) discovering moderators of treatment outcome, and (e) elucidating mechanisms of equifinality and multifinality. Although most behavioral traits probably are distributed continuously, identifying those that are discrete will advance the science of developmental psychopathology. Disorders for which taxometric analyses might be applied include anxiety, attention deficit hyperactivity disorder, autism spectrum disorders, conduct problems, depression, and schizophrenia.

There are many reasons to prefer continuous models of psychopathology over categorical models. Some of these are philosophical, whereas others are methodological and pragmatic. Philosophical arguments address one of the most fundamental yet elusive questions facing behavioral scientists, namely, what constitutes a disorder? Do mental disorders reflect failures of biological systems to perform naturally selected functions (e.g., Wakefield, 1992, 1993, 1997, 1999), or are they defined by somewhat arbitrary distinctions derived from social values (e.g., Lilienfeld & Marino, 1995, 1999)? Should a set of symptoms be considered a disorder when induced by a high risk environment, or is evidence of independent internal mechanisms necessary (Wakefield, Pottick, & Kirk, 2002)? Can environmental risk and internal mechanisms even be considered as separate causal agents (e.g., Bremner & Vermetten, 2001)? Could all of the 365 categories in the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IV;* American Psychiatric Association, 1994) possibly be distinct (e.g., Houts, 2002; Kendell, 1989)? Does the *DSM-IV* framework pathologize normal behavior (e.g., Richters & Cicchetti, 1993)?

Each of these questions instantiates the difficulties faced in deciding which behaviors or behavioral syndromes to define as disordered. In fact, whether one endorses or rejects the

This article was written in honor of Paul Everett Meehl, who died in February 2003. May his passing reinvigorate the ideals of empirical rigor and scientific realism that he so dutifully and tirelessly advanced. Rarely is a field so indebted to the work of one individual. Preparation of this article was supported in part by Grant MH63699 from the National Institute of Mental Health. The author expresses thanks to Geraldine Dawson, Liliana J. Lengua, and M. Jamila Reid for their helpful contributions to this work.

Address correspondence and reprint requests to: Theodore P. Beauchaine, Department of Psychology, University of Washington, Box 351525, Seattle, WA 98195-1525; E-mail: tbeaucha@u.washington.edu.

enterprise of diagnosis altogether may follow from his or her answers to these and similar questions. Such answers are likely to depend upon what philosophers of science have termed *essentialist* versus *nominalist* views of human behavior (Flanagan & Blashfield, 2002). Essentialists argue that mental disorders reflect objective underlying causal realities that are independent of human values. Wakefield's (1992, 1993, 1997, 1999) harmful dysfunction analysis reflects an essentialist philosophy, as disturbances in behavior are assumed to be caused by biological dysfunctions that impact evolutionary fitness. Essentialist views of behavior, particularly those that invoke evolutionary accounts of adaptation and maladaptation, have been criticized as being teleological because the link between current adjustment and past evolutionary fitness is often tenuous (e.g., Beauchaine, 1999; Richters & Hinshaw, 1999).

In contrast, nominalists argue that psychiatric disorders reflect deviations from socially constructed prescriptions for behavior and that there are no objective means of demarcating normality from abnormality. This postmodern philosophy is exemplified in the writings of Szasz (2000) and Lilienfeld and Marino (1995, 1999), who suggest that most psychiatric disorders are characterized by unclear boundaries and a lack of defining features. As such, nominalists are likely to interpret behaviors as falling along a continuum of social acceptability, to consider diagnostic cutoffs as arbitrary, and to be wary of diagnosis altogether. Nominalist philosophies are sometimes preferred over essentialist philosophies because they carry no assumptions about unobservable phylogenic mechanisms of behavior. This preference may be prudent when clear links between symptoms and biological dysfunction are lacking.

Although some developmental psychopathologists have weighed in on this philosophical debate, most discussions within the field regarding the merits of dimensional versus discrete models of psychopathology have occurred at an entirely different level. These discussions have focused on the methodological and practical constraints imposed by a categorical diagnostic system and the utility of dimensional assessment approaches toward capturing important developmental and contextual information about children's adjustment (see Cummings, Davies, & Campbell, 2000; Hinshaw & Park, 1999). Although space limitations preclude a full review of the limitations of categorical diagnosis, criticisms include assertions that the *DSM* framework (a) places the locus of disorder within the individual, ignoring contextual information about familial, neighborhood, and other influences; (b) assumes biological bases for psychiatric disorders despite claims of being descriptive and atheoretical; (c) is devoid of developmental guidelines for assessment and diagnosis; (d) fails to account for cultural differences in the expression of maladaptation and distress; (e) is limited in its clinical application because of symptom heterogeneity within diagnostic categories and homogeneity across diagnostic categories; and (f) assumes that behavioral syndromes are categorical entities with discrete etiologies (Cantwell, 1996; Clark, Watson, & Reynolds, 1995; Cummings et al., 2000; Hinshaw, Lahey, & Hart, 1993; Jensen & Hoagwood, 1997; Sonuga–Barke, 1998; Sroufe, 1997). Finally, dichotomizing continuous scores of symptom severity, a necessary consequence of categorizing, results in reduced reliability and statistical power and may produce misleading outcomes in research assessing the precursors, correlates, and sequelae of psychopathology (see MacCallum, Zhang, Preacher, & Rucker, 2002).

For these reasons, many developmental psychopathologists prefer to use empirically derived instruments rather than categorical diagnoses in assessment. Such instruments, of which the Child Behavior Checklist (Achenbach, 1991) is an example, evaluate behavioral traits across multiple continuous dimensions. Symptoms are assessed using factor-analytically derived subscales, providing for comparisons of children's scores to age- and gender-matched norms. Empirical assessment instruments carry no assumptions about biological substrates, environmental influences, or etiological origins of a given behavioral profile. Moreover, symptom overlap among subscales is considered to be clinically relevant information rather than nuisance contamination across diagnostic categories (see Cummings et al., 2000).

These characteristics are appealing because developmental psychopathologists are typi-

cally more interested in the processes through which maladaptation emerges than in the descriptive aspects of behavior (Cicchetti, 1993; Rutter & Sroufe, 2000; Sroufe & Rutter, 1984).[1] Equifinality, for example, suggests that a given disorder can be the end state of numerous developmental pathways, and multifinality suggests that children in similar high risk situations can diverge toward quite disparate end states, only some of which will be disordered (Cicchetti & Rogosch, 1996). Although both categorical and empirically derived instruments are descriptive, empirical assessments are much more flexible for examining diverse developmental trajectories. This is because longitudinal evaluations of behavioral functioning that are indexed to age-matched norms provide profiles of emerging and diminishing symptoms over time. Such profiles are invaluable toward conceptualizing psychopathology (Cicchetti & Rogosch, 1996; Kagan, 1997) and include developmental information that cannot be extracted from categorical classifications. Indeed, continuous scales provide for the assessment and tracking of symptoms that may be subthreshold for a given diagnostic category (Hinshaw et al., 1993).

Evaluating the Ontological Status of Traits

Clear articulation by developmental psychopathologists of the limitations of categorical diagnostic systems represents a major contribution to the study of human behavior. The Western intellectual tradition is replete with instances in which the establishment of rigid and artificial boundaries has resulted in stereotypes, prejudiced policies, and impediments to scientific progress (Waters & Beauchaine, 2003). Yet we should not confuse preferences for continua that are based on philosophical, methodological, or pragmatic concerns with the ontological status of behavioral traits. Whether a particular trait or disorder represents a discrete entity is an empirical question that cannot be settled through methodological convention or philosophical debate (see Meehl, 1992, 1995; Sonuga–Barke, 1998). Based on a growing number of formal taxometrics investigations, evidence from the adult psychopathology literature suggests that at least some traits and disorders are distributed as discrete classes, including endogenous depression (Ambrosini, Bennett, Cleland, & Haslam, 2002; Beach & Amir, 2003; Haslam & Beck, 1994), schizotypy (Blanchard, Gangestad, Brown, & Horan, 2000; Golden & Meehl, 1979; Korfine & Lenzenweger, 1995; Lenzenweger, 1999; Lenzenweger & Korfine, 1992; Tyrka, Cannon, Haslam, Mednick, Schulsinger, Schulsinger, Parnas, 1995; Tyrka, Haslam, & Cannon, 1995), dissociative experiences (Waller, Putnam, & Carlson, 1996; Waller & Ross, 1997), psychopathy (Harris, Rice, & Quinsey, 1994), and Type A behavior patterns (Strube, 1989). In contrast, very few taxometrics investigations have appeared in the child psychopathology or developmental psychopathology literatures (for exceptions see Erlenmeyer–Kimling, Golden, & Cornblatt, 1989; Fraley & Spieker, 2003; Skilling, Quinsey, & Craig, 2001; Woodward, Lenzenweger, Kagan, Snidman, & Arcus, 2000). It is therefore unclear when in development these discrete behavioral traits emerge. Perhaps this lack of taxometrics research within developmental psychopathology reflects the aforementioned preference for continuous models of behavioral functioning and an associated wariness of categorical diagnostic systems. Yet there are a number of reasons why inquiries into the ontological status of behavioral traits and disorders should be of *central* interest to developmental psychopathologists. Moreover, the limitations of categorizing have very little to do with the

1. As an interesting aside, psychology is not the only discipline to wrestle with the importance of description versus process as topics of inquiry. Several months before his untimely death in May 2002, the renowned paleontologist Stephen Jay Gould delivered a series of public lectures on evolutionary theory. At one of these lectures Gould fielded the question, "Which is your favorite dinosaur?" Gould's response was incisive: "I don't like any of them," he stated. In elaborating, Gould noted that a preoccupation with the descriptive features of dinosaurs has distracted many paleontologists from asking questions that are of considerably more scientific value. Gould then explained that as much importance has been traditionally placed on description as on the mechanisms of evolutionary change that produce phenotypic variability, even though questions concerning the latter are of far greater significance.

ontological status of a trait or disorder because individual differences in symptoms are nearly always observed and merit scientific scrutiny, regardless of whether the trait is discrete or continuous. Thus, our choice of categorical versus empirical assessment does not depend on the outcome of taxometrics investigations. In the sections to follow, these points will be elaborated. Moreover, it will be suggested that theory-driven taxometrics research holds the potential to address long-standing yet unanswered questions of unique concern to developmental psychopathologists and can enrich our understanding of emerging disorders in a way that cannot be achieved by conducting similar studies with adults.

Why Look for Typologies?

Before continuing, it may be useful to address a question that often emerges in discussions about taxometrics research, namely, "who cares?" What does it matter whether a trait or a disorder reflects a distinction in kind or a difference in degree? The simplest answer to this question is that identifying taxa enables us to establish nonarbitrary cutoffs that distinguish between those with and without a trait or disorder (Beauchaine & Waters, 2003; Meehl, 1995). In Plato's words, knowing that a trait is distributed discretely allows us to "carve nature at its joints." Moreover, evidence of taxonicity offers strong support for the construct validity of a trait or diagnostic entity, particularly when variables from multiple levels of analysis (e.g., physiological, behavior–observational, self-report) are used as markers of the putative taxon (Beauchaine & Beauchaine, 2002; Meehl, 1995). Note that this also addresses the philosophical question of whether a trait or diagnostic category marks a discrete entity. Such questions are much more appropriate at the level of disorder than at the level of a diagnostic system (see Flanagan & Blashfield, 2002). Arguments against the validity of the *DSM* are often applied to the entire classification system, which cannot possibly be valid or invalid in a binary sense. Rather, some diagnostic categories are of higher construct validity than others.

Yet "merely" carving nature at its joints may be unsatisfying for those with predominantly applied interests, who might still ask, "so what?" Are there any reasons to suspect that prevention or intervention strategies might differ based on knowing that a high risk trait is distributed discretely? Meehl (1992, 1995) argued that this is precisely the case and that therapeutic strategies might differ substantially based on knowing if a person belongs to a high risk taxon group. Consider the diathesis for schizophrenia. As noted above, several authors have confirmed that schizotypy, or a constellation of observable symptoms that appears to mark a genetic liability for schizophrenia spectrum disorders, is distributed as a discrete class, with a base rate somewhere around 5% (Blanchard et al., 2000; Golden & Meehl, 1979; Korfine & Lenzenweger, 1995; Lenzenweger, 1999; Lenzenweger & Korfine, 1992; Tyrka, Cannon, et al., 1995). Note that this base rate is considerably higher than the 1.1% estimate of the prevalence of schizophrenia in the general population (Regier, Narrow, Rae, Manderscheid, Locke, & Goodwin, 1993). Of course, this is expected for a genetic trait that is not fully penetrant. In other words, the genetic liability may be a necessary but insufficient condition for developing schizophrenia. Indeed, although the outcome and course of childhood-onset schizophrenia has been understudied (see Dulmus & Smyth, 2000), the expression of the disorder is clearly sensitive to contextual influences. Expressed emotion is observed at high levels in families of children with schizophrenia spectrum disorders (Hamilton, Asarnow, & Tompson, 1999) and exerts a strong influence on both course and prognosis (Falloon, Boyd, McGill, Williamson, Razani, Moss, Gilderman, & Simpson, 1985; Hogarty, Anderson, Teiss, Kornblith, Greenwald, Ulrich, & Carter, 1991). The implication is that premorbid identification of children with a genetic diathesis for schizophrenia might facilitate targeted family interventions that improve outcome and course and perhaps delay or prevent the onset of the disorder (see Cornblatt, Obuchowski, Roberts, Pollack, & Erlenmeyer–Kimling, 1999). Taxometric analyses of cognitive and neuromotor performance variables suggest that schizotypy can be identified in children and adolescents,

with a base rate similar to that found in adults (Erlenmeyer–Kimling et al., 1989). Thus, both applied and basic interests are served by identifying taxonic traits and by distinguishing between taxon and nontaxon group members. More specific advantages of taxometrics investigations will be presented in later sections.

Traditional Approaches to Searching for Boundaries

Given that identifying taxa is of both basic and applied import, how do we go about doing it? Before addressing this question, two issues must be considered. First, what precisely is meant by the term taxon, and second, what are common misconceptions about the requirements of testing taxonic hypotheses?

The term taxon has multiple uses across scientific disciplines, and it is therefore difficult to arrive at a universal definition (Meehl, 1999). In the strongest sense, taxa represent entities that are clearly different in kind by some functional criterion, such as separate species as defined by reproductive isolation (Mayr, 1942). As is often the case in the behavioral sciences, however, the term has been used more loosely to refer to almost any method of classifying, be it categorical or empirical (Achenbach, 1993). In taxometrics research, a taxon is typically defined by a boundary that separates taxon group members from nontaxon group members, that is, from evidence that the two groups emerge from separate or discrete distributions. However, because distinct distributions are likely to overlap considerably, given both within-groups variability and measurement error, traditional statistical methods are of little use in detecting taxa. Rather, formal taxometric methods are required. Unfortunately, a number of ineffective strategies for identifying taxa continue to appear in the psychology and psychiatry literatures. Some consideration of these strategies is therefore warranted.

Bimodality

Early efforts to differentiate types from continua involved plotting the univariate distributions of traits or symptoms and inspecting the admixed distribution for evidence of bimodality (Kendell, 1989). However, two discrete distributions often appear to be unimodal when mixed, even at quite large effect sizes (Grayson, 1987; Murphy, 1964; Waller & Meehl, 1998). This is illustrated in Figure 1, where two distributions with a mean separation of 2 standard deviations are mixed. Inspection of the admixed distribution reveals no evidence of bimodality. Indeed, the combined distribution is near normal with a kurtosis value (−.52) that is acceptable for almost all statistical tests that assume normality. It is worth emphasizing the magnitude of the effect represented in Figure 1. Recall that Cohen (1988) defined a large effect as .8 standard deviation units of separation between means (Cohen's *d*). Thus, the effect size illustrated in Figure 1 is 2.5 times larger than Cohen's definition of a large effect. Clearly, bimodality is a very weak criterion for inferring taxonicity.

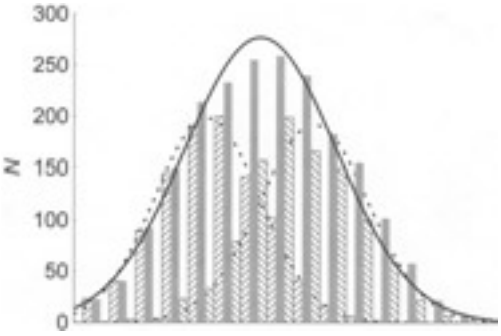

Figure 1. An admixture of two discrete normal distributions of $n = 1000$, each indicated by hatched bars and dashed lines. The effect size separating the distributions is 2 standard deviations. Note that despite this large effect, there is no evidence of bimodality in the combined distribution, indicated by filled solid bars. Note also that the combined distribution is near normal, as indicated by the solid curve.

Figure 1 also illustrates the fallibility of the commonly held belief that discrete traits and disorders are marked by clear and distinct boundaries (see also Lilienfeld & Marino, 1995; Meehl, 1995; Sonuga–Barke, 1998). There are several reasons why this is unlikely to be the case. First, the mechanisms responsible for placing individuals into taxon and nontaxon groups are almost always latent (unobservable) constructs that are assessed by imperfect manifest

indicators. Even genetic disorders of high penetrance are not manifested equivalently across individuals. Huntington's chorea, for example, is a progressive and degenerative disorder of the nervous system that results in mood lability, emotional instability, and motor control abnormalities. Although the disorder is transmitted by an allelic variant of a single dominant gene on chromosome 4, its course is variable across individuals, with age of onset dispersed across a wide range (Brooks, Murphy, Janota, & Lishman, 1987). Thus, the latent cause of the disorder is necessarily taxonic (either one has the genetic variant or one does not), but the manifest indicators are variable (see also Meehl, 1995, 2001).

Second, indicators of taxon group membership may be distributed normally among nontaxon group members. Using procedures to be described below, Woodward et al. (2000) reported evidence that extreme behavioral reactivity among infants is distributed discretely. About 10% of 4-month-olds were classified into a highly reactive group based on excessive arching, crying, hyperextension, and leg movements in response to a series of stimuli. Moreover, group membership was predictive of extreme behavioral inhibition 4 years later. Yet it would be naive to expect a clear boundary between taxon and nontaxon group members for specific indicators of behavioral reactivity or fear. As an adaptive emotion, fear is observed in the behavioral repertoires of all infants, with individual variability that characterizes most evolved traits. Thus, infants in the Woodward et al. sample who were nontaxon group members exhibited a wide range of behavioral reactivity.

Third, most behavioral indicators used in psychological and psychiatric research are measured with imperfect precision and reliability, contributing to distributional overlap even when true taxa exist. This final point will be reconsidered in later sections of this article that outline strategies for future taxometrics work.

Cluster analysis

Following recognition that bimodality and clear boundaries are unlikely to be observed in the presence of discretely distributed traits, several alternative methods aimed at testing taxonic conjectures have been developed. Undoubtedly the most popular of these in the psychological literature is the set of over 300 algorithms collectively referred to as *cluster analysis* (Blashfield & Aldenderfer, 1988). These methods divide data sets into multiple partitions, either by maximizing between groups variance or by minimizing within-groups variance. Although details about specific cluster analytic methods cannot be presented here because of space constraints, there is ample evidence to suggest that they do not provide strong tests of taxonic structure because they always yield subgroups that differ significantly on the indicator variables, regardless of whether true taxa exist (see Blashfield & Aldenderfer, 1988; Klein & Riso, 1993). Moreover, even when known taxa are embedded in simulated data sets, clustering algorithms perform poorly when there is overlap among groups and often fail to identify the correct number of clusters (Atlas & Overall, 1994; Krieger & Green, 1999). This state of affairs exists despite significant efforts over the last two decades to develop accurate "stopping rules." Thus, cutoffs for group membership that are derived from cluster analyses are just as likely to be arbitrarily located along a continuous dimension as they are to distinguish between true taxa. Because clustering algorithms are structure *imposing* rather than structure *seeking*, these methods carry the real risk of identifying false joints in nature and provide very weak tests of typological models (Beauchaine & Beauchaine, 2002; Meehl, 1979).

Other approaches

Several additional techniques have appeared in the psychology and psychiatry literatures in efforts to identify homogeneous subgroups of individuals within larger samples. These include mixture analysis (e.g., Fleiss, 1972), latent class analysis (e.g., Lazarsfeld & Henry, 1968), and more recently developed methods for identifying groups through analyses of latent growth trajectories (e.g., Nagin, 1999). The operating characteristics of both mixture analysis and latent class analysis have been explored by a number of researchers, as sum-

marized by Klein and Riso (1993). In brief, these techniques suffer from some of the same limitations as cluster analysis. Both divide most data sets into partitions that are consistent with a discrete latent class interpretation, yet there is no way to determine whether the identified classes are truly discrete. This is also the case when identifying groups using latent growth trajectories (Nagin, 1999; Nagin & Tremblay, 2001). Thus, although these techniques may have heuristic utility, none can determine the ontological status of a trait or a disorder as discrete versus continuous.

Schizotaxia, Schizotypy, Schizophrenia Revisited

In his 1994 address for a Distinguished Professional Contribution award from the American Psychological Association (APA), Meehl (1995) summarized the results of a now 30-year effort to develop techniques that could distinguish between types and continua. The resulting algorithms, referred to collectively as coherent cut kinetics (CCKs), were devised with the intent of identifying latent taxa when they exist and rejecting the taxon hypothesis when they do not, thereby addressing an inherent limitation of alternative methods of classifying. Meehl's (1995) paper outlined the two most commonly used CCK procedures, Mean Above Minus Below A Cut (MAMBAC; Meehl & Yonce, 1994), and Maximum Covariance (MAXCOV; Meehl & Yonce, 1996; Waller & Meehl, 1998)–HITMAX. These techniques will be described in some detail below; but before doing so, a description of the context in which they were developed will be presented. This is important because the sustained effort of Meehl and his colleagues to address one of the most fundamental problems in research on psychopathology exemplifies the ingenuity, tenacity, and rigor required when tackling difficult scientific questions (Waters & Beauchaine, 2003).

Meehl (1962) was one of the first proponents of a diathesis–stress model of psychopathology. He believed that a genetic diathesis for schizophrenia, which he labeled *schizotaxia*, was transmitted through a single gene but resulted in schizophrenia for only a subset of afflicted people. He also believed that a core set of observable phenotypic indicators, expressed as *schizotypy*, marked the genetic variant, regardless of whether a person developed schizophrenia. Through careful observation that included extensive clinical work with schizophrenia probands and their family members, Meehl (1962) identified four schizotypic characteristics as putative markers of the schizotaxic genotype: *anhedonia*, or a limited capacity to experience pleasure[2]; *interpersonal aversiveness*, or social fear, distrust, and anticipation of rejection; *ambivalence*, or seemingly concurrent motivation toward interpersonal approach and withdrawal; and *cognitive slippage*, or somewhat loose control of associations. Although unknown to Meehl at the time, we might add to this list smooth pursuit and saccade eye tracking abnormalities, which are also observed in the relatives of schizophrenia probands (e.g., Curtis, Calkins, Grove, Feil, & Iacono, 2001), and have been proposed to result from a single dominant gene (e.g., Avila, McMahon, Elliott, & Thaker, 2002; Levy, Holzman, Matthysse, & Mendell, 1993).

It bears repeating that Meehl expected an unknown fraction of schizotypes to decompensate into diagnosable schizophrenia. Moreover, he recognized that schizotypy would be difficult to identify with specificity because the distribution of schizotypic characteristics among genetic positives was likely to overlap considerably with the distribution of similar traits in the general population. Thus, new methods would be required to disentangle the admixed distributions of schizotypes and genetic normals. This is a classic case of a well-articulated theory serving to motivate and inform the development of a new methodological approach, and more importantly, a new way of

2. In more recent writings, Meehl (1975, 1989, 1990) recanted the assertion that anhedonia marks a genetic liability for schizophrenia, suggesting instead that hedonic capacity is a dimensionally distributed trait. However, Blanchard et al. (2000) reported that anhedonia was distributed discretely among a large sample of college students, with a base rate similar to that found in taxometrics investigations using other schizotypy markers. Meehl's original position may have therefore been correct.

thinking about psychopathology. Although a single gene locus has never been identified for schizophrenia, the schizotypic traits outlined by Meehl have been shown in repeated taxometrics investigations to mark a group that is distributed as a discrete latent class (Blanchard et al., 2000; Erlenmeyer–Kimling et al., 1989; Golden & Meehl, 1979; Korfine & Lenzenweger, 1995; Lenzenweger, 1999; Lenzenweger & Korfine, 1992; Tyrka, Cannon, et al., 1995; Tyrka et al., 1995). It is to describing the most common CCK methods that this article now turns.

CCK Algorithms

MAXSLOPE procedure

Meehl and colleagues' CCK algorithms include several interrelated procedures that identify taxonic structure by progressing along successive cuts of an indicator variable and examining the statistical behavior of related variables in contiguous regions of the cut (Meehl, 1999; Meehl & Yonce, 1994, 1996).[3] A description of MAXSLOPE (Grove & Meehl, 1993) provides an intuitively appealing illustration of the CCK approach. Consider two hypothetical groups, those with and without a genetic diathesis for schizophrenia. Assume that reliable measures of anhedonia and social anxiety are obtained and that, for illustrative purposes, the groups are roughly equal in size. In MAXSLOPE, the indicators are plotted against one another and a smoothed regression line is fitted to the scatterplot (see Figure 2). The slope of the regression line (dy/dx) is then calculated at successive points of the x variable (social anxiety). If two discrete groups are present and the effect size is adequate, then a discontinuity is observed in the regression slope, which is maximized at the level of social anxiety (x) that best discriminates between groups. In other words, the correlation between social anxiety (x) and anhedonia (y), represented by the regression slope, is highest at the point of greatest admixture. The only requirement is that anhedonia and social anxiety are correlated less within groups than between groups. If only one group is present, then no discontinuity in slope is observed.

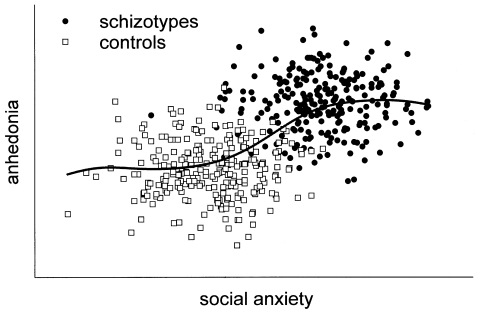

Figure 2. An illustration of the MAXSLOPE procedure using a hypothetical bivariate distribution of anhedonia and social anxiety. Note that the slope of the smoothed regression line is relatively flat within groups, but it steepens where schizotypes (n = 250) and controls (n = 250) are mixed. The maximum slope is obtained at the value of social anxiety that best discriminates between groups.

MAXCOV procedure

Although easier to explain, MAXSLOPE is less commonly used than MAXCOV, a related procedure that operates on triads of indicators rather than variable pairs (see Meehl & Yonce, 1996; Waller & Meehl, 1998). MAXCOV examines the covariance, or unstandardized correlation of two indicators across successive intervals of a third indicator, and a smoothed plot is fitted through the resulting function. To illustrate, assume that in addition to having reliable indicators of social anxiety (x) and anhedonia (y), we collect eye-tracking abnormality data (z) from groups with and without a genetic diathesis for schizophrenia. A hypothetical scatterplot of social anxiety and anhedonia is presented in the top left portion of Figure 3, with their covariance across intervals of eye tracking abnormality appearing below. Similar to the MAXSLOPE example, if two discrete groups are present and the effect size is adequate, then a discontinuity is observed in the covariance of variables x and y (anhedonia and social anxiety), which is maximized at the level of variable z (eye tracking) that best differentiates between groups. This

3. In total, there are 13 CCK procedures (Meehl, 1999). This article emphasizes the most commonly used and extensively tested of these algorithms.

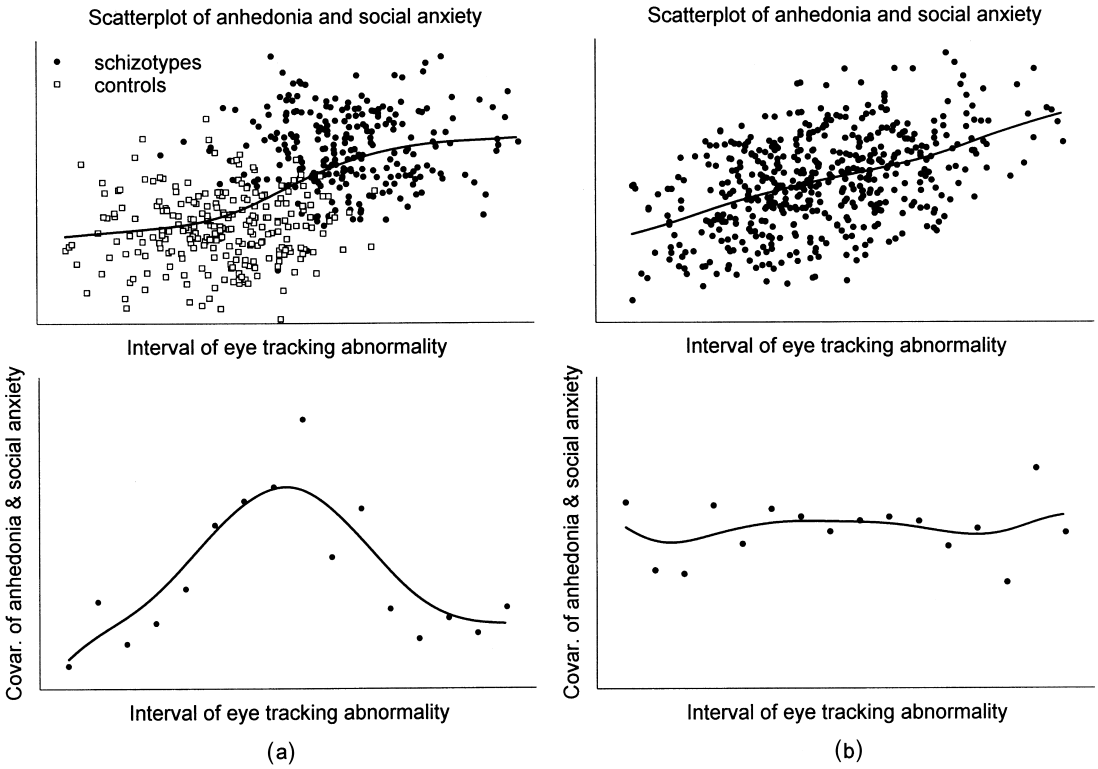

Figure 3. An illustration of the MAXCOV procedure using hypothetical distributions of anhedonia and social anxiety across intervals of eye tracking abnormality. (a) In the taxonic case the greatest covariance of anhedonia and social anxiety is observed within the interval of eye tracking abnormality that best discriminates between groups of schizotypes ($n = 250$) and controls ($n = 250$). (b) In the continuous case there is no peak in the covariance function ($n = 500$). Note that the overall correlation between anhedonia and social anxiety is .50 in both samples.

is referred to as the HITMAX value, because group membership assignments are most accurate when it is used as a demarcating boundary. As illustrated in the bottom left portion of Figure 3, the covariance function exhibits a marked peak at the HITMAX value. At lower taxon base rates, the location of this peak migrates toward the right (see Meehl, 1995; Meehl & Yonce, 1996). In contrast, given continuous data the covariance function of x and y across intervals of z is relatively flat, as shown in the bottom right portion of Figure 3.

For each trivariate combination of indicators, MAXCOV yields estimates of the taxon base rate, the sample sizes of the taxon and nontaxon groups, the false positive and false negative rates of group membership, and the HITMAX value. Moreover, MAXCOV can be used to assign individual cases or observations to the taxon and nontaxon groups using Bayes' theorem (see Beauchaine & Beauchaine, 2002; Meehl & Yonce, 1996; Waller & Meehl, 1998). This versatility has led to MAXCOV being the most commonly used CCK procedure.

MAMBAC procedure

When testing taxonic hypotheses, MAXCOV is often used in conjunction with yet another CCK algorithm, MAMBAC (Meehl & Yonce, 1994). In this procedure, indicators are analyzed in variable pairs. One variable is first sorted in ascending order, which also sorts the other variable with some degree of efficiency if both are valid markers of the analyzed trait. Continuing with the previous example, sort-

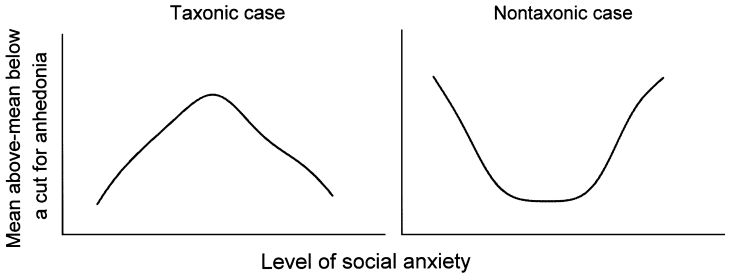

Figure 4. An illustration of the MAMBAC procedure using hypothetical distributions of anhedonia and social anxiety. In the taxonic case a peak in the MAMBAC plot is observed at the level of social anxiety that best discriminates between groups of schizotypes (n = 500) and controls (n = 500). In the nontaxonic (continuous) case the MAMBAC plot is U shaped (n = 1000).

ing on social anxiety (x) will also sort on anhedonia (y) because the two variables are correlated. Next a sliding cut is moved across all values of social anxiety, and the mean of anhedonia is calculated both above and below the cut. At each point the mean above the cut is subtracted from the mean below the cut, and the resulting function is plotted. If a taxon group is present, a marked peak is produced in the MAMBAC function at the value of social anxiety (x) that best discriminates between those in the taxon group and those in the nontaxon group (Meehl, 1995; Meehl & Yonce, 1994). In contrast, continuously distributed data produce a U-shaped function (see Figure 4).

Evaluating Taxonic Hypotheses

In the 30 years since Meehl (1973) introduced the first of his CCK algorithms, much has been learned about conducting taxometrics investigations. Based on an increasing number of studies appearing in the adult literature, several theoretical papers by Meehl and others, and Monte Carlo simulations examining the efficiencies of CCK procedures, it is now possible to offer a number of recommendations for future taxometrics work. Above all, taxometrics investigations require that indicators be selected with great attention to validity and measurement precision, the latter of which is rarely considered in psychological research. Moreover, because CCKs require larger samples than are available in most studies, taxometric questions have often been pursued in an ad hoc fashion with large data sets of convenience. As such, many of these investigations have been atheoretical, using imprecise rating scale measures that fail to tap putative mechanisms of taxonicity directly. Under these conditions, there is a greater chance of detecting taxa that are spurious and a lesser chance of detecting taxa that truly exist. Furthermore, null findings have frequently been reported as supporting continuous models of psychopathology, a logical inconsistency that would probably not emerge through the editorial process, given more commonly used statistical methods. In the sections to follow, each of these issues will be elaborated and general recommendations for future taxometrics research will be provided. Next, implications for addressing questions of specific interest to developmental psychopathologists will be outlined.

Selecting valid indicators

Indicator validity refers to the degree to which a variable marks the latent construct it is purported to measure. In the case of schizotypy, the results of numerous studies have confirmed that anhedonia, perceptual aberration, and social withdrawal mark a discretely distributed group of individuals who are at elevated risk for schizophrenia, thereby establishing the validity of each variable as an indicator of the schizotypy construct (Blanchard et al., 2000; Erlenmeyer–Kimling et al., 1989; Golden & Meehl, 1979; Korfine & Lenzenweger, 1995; Lenzenweger, 1999; Lenzenweger & Korfine, 1992; Tyrka, Cannon, et al., 1995). However, if the validity of indicators is contingent

upon their performance in previous taxometrics investigations, how does one choose markers of a hypothesized latent taxon that has yet to be identified? As outlined by Meehl (1995), the search for taxonic entities is a bootstrapping endeavor because there are almost never gold standards or litmus tests for confirming the validity of symptoms as markers of a discrete diagnostic entity. If there were such gold standards, then there would be no need to conduct taxometrics investigations in the first place.

Given that the validity of an indicator cannot be known a priori, it is important that candidates be selected based on strong theory (Meehl, 1999; Waters & Beauchaine, 2003). This way, if no taxon is identified, it is unlikely to be a function of indicator invalidity. For example, suppose I suspect that a subset of boys with conduct disorder (CD) are at elevated genetic risk for developing psychopathy. This is similar to Meehl's (1962) schizotaxia example, because both genetic and environmental factors contribute to the psychopathy phenotype (see Lykken, 1995), yet the specific loci of the genotype have not been identified. In order for CCKs to provide a reasonable test of the taxonic hypothesis, it is imperative that I select indicators that are reasonably *specific* to the psychopathy construct. Thus, I should avoid using broadband externalizing symptoms such as disobedience, oppositionality, or impulsivity as indicators, because such symptoms are observed across a wide range of children and adolescents and are only sometimes indicative of psychopathology (see, e.g., Hinshaw, 2003). These symptoms are therefore unlikely to mark genetic risk for psychopathy with specificity. Given that psychopathy may be distributed as a discrete class in adulthood (Harris et al., 1994), it would be best to select variables that distinguish between externalizing adults with and without psychopathy and to consider variables that mark developmental precursors of psychopathic traits. Candidates might include manipulativeness, callousness, lack of empathy, and certain physiological markers of underarousal. Each of these has been tied more specifically to either psychopathy or severe conduct problems than to broadband externalizing symptoms (e.g., Barry, Frick, DeShazo, McCoy, Ellis, & Loney, 2000; Beauchaine, 2001; Beauchaine, Katkin, Strassberg, & Snarr, 2001; Frick, Bodin, & Barry, 2000).

Increasing measurement precision

Although validity is often conflated with measurement precision, the two are not equivalent. Precision refers specifically to our ability to measure a construct without error. Many psychological constructs are low in measurement precision, yet have adequate validity. This is particularly true for constructs assessed via self- and other-report. For example, whereas depression may be assessed with reasonable validity using scores from the Children's Depression Inventory (CDI; Kovacs, 1992), there is considerable measurement error in any given score and it is inappropriate to make equivalency statements across scores. Thus, although significant impairment might be inferred in a child with a CDI score of 30 (validity), it cannot be assumed that two different children with scores of 30 suffer from equivalent levels of depression (precision). Indeed, Likert scales reflect an ordinal level of measurement and do not carry the precision of interval or ratio scales (Stevens, 1951).

This discussion is important because high measurement precision is *essential* for taxometrics investigations. Extensive Monte Carlo simulations suggest that MAXCOV, the most commonly used CCK, is effective in detecting latent taxa only at effect sizes (d) of 1.2 or larger (Beauchaine & Beauchaine, 2002; Meehl, 1995). To place this in context, a typical effect size in psychological research is around .6, or half as large (Cohen, 1988). Thus, given an average effect size, existing taxa will go undetected, even if valid indicators are used. CCKs simply do not have the discriminating power to detect small and medium effects.

Several strategies can be used to increase measurement precision. Factor analysis of individual items into common components is one such strategy. The effect is to reduce error associated with single items by isolating shared variance across items (see Nunnally & Bernstein, 1994). Using factor scores as indicators should therefore enhance the sensitivity of CCKs

to underlying taxa. In fact, there are taxometrics procedures designed specifically to deal with factor scores (Waller & Meehl, 1998). Unfortunately, many taxometric investigations conducted to date have used individual items from factor analytically derived scales. Although this strategy boosts the number of indicators available for analysis, it may hamstring efforts to identify latent taxa due to low measurement precision.

A second strategy for reducing measurement error is to assess growth in the same construct across time. In doing so, each additional assessment point provides enhanced precision (Rogosa, Brandt, & Zimowski, 1982). For example, gender differences in the growth trajectories of depressive symptoms emerge *before* mean differences in depression can be detected (Cole, Tram, Martin, Hoffman, Ruiz, Jacquez, & Maschman, 2002). This increased precision suggests that longitudinal growth trajectories may be useful in taxometrics investigations. To date, however, only one longitudinal taxometric study has been conducted (Tyrka, Haslam, et al., 1995), and growth trajectories were not used as indicators. This point should be of particular interest to developmental psychopathologists, given their concern with symptom emergence over time.

Finally, precision can be increased by using indicators that carry inherently less measurement error than self- and other-report. Such data sources might include behavior–observational measures, biological or physiological markers, or etiological information such as age of onset, among others. When collected carefully, these indicators are more precise by nature than data derived from Likert scales.

Using indicators from multiple levels of analysis

There are several additional reasons to include variables other than those collected via rating scales in taxometrics research. First, a much stronger case for the validity of a disorder as a discrete class can be made when indicators are drawn from multiple levels of analysis, including behavioral symptoms, biological signs, and etiological markers (Feighner, Robins, Guze, Woodruff, Winokur, & Munoz, 1972; Robins & Guze, 1970; see also Cicchetti & Dawson, 2002). Second, rating scales are known to elicit systematic response tendencies, including both halo effects (e.g., Saal, Downey, & Lahey, 1980) and positive/negative response biases (e.g., Macmillan & Creelman, 1990; Rajendar, 1996), either of which might produce spurious taxonic structure. Third, human beings are prone toward categorical thinking (see Malt, 1993; Rosch & Lloyd, 1978; Smith, 1995), which includes a natural tendency to classify based on past experiences and preexisting beliefs (see Cantor & Genero, 1986; Cantor & Mischel, 1979; Flanagan & Blashfield, 2002; Semin & Rosch, 1981). When making categorical decisions, raters also become more confident with increasing experience (Dawes, Faust, & Meehl, 1989), and their beliefs become more divergent over time, as assessed by Likert measures (Simon, Pham, Le, & Holyoak, 2001). Indeed, as much as 50% of the variance in Likert scales may be attributable to rater bias (Hoyt & Kerns, 1999). Thus, an exclusive reliance on rating scale data may be problematic when performing taxometrics research, particularly when evidence suggests that raters hold categorical beliefs about the nature of a construct. Attachment classifications, for example, are assigned based on expert ratings of children's behavior during separation from and reunion with their mothers (Ainsworth, Blehar, Waters, & Wall, 1978; Sroufe & Waters, 1977). Because many researchers endorse categorical models of attachment (see Fraley & Waller, 1998), subtle biases could impact their ratings, producing latent distributions that appear to be discrete. In fact, we have demonstrated that manipulating the cognitive sets of raters as categorical versus dimensional can produce spurious taxa when only Likert scale measures are used (Beauchaine & Waters, 2003). Because no single indicator can drive taxonic findings, this does not preclude the use of rating scale data in taxometrics investigations. However, it does suggest that variables from other levels of analysis should be included.

Proving the null

In reporting the results of taxometrics analyses, it has become common to interpret nega-

tive findings as supporting a continuous model of the construct being assessed. Because CCK procedures are structure seeking, however, there is a logical inconsistency in this practice. The null hypothesis in taxometrics research is that the analyzed trait or disorder is distributed along a continuum, and the alternative hypothesis is that it is distributed categorically. To conclude that a construct is continuously distributed based on negative findings is therefore tantamount to proving the null. Thus, it is not the case that taxometrics analyses can indicate the presence of a continuous distribution, as some have suggested (e.g., Widiger, 2001). It should be emphasized that this is not an arcane point born of adherence to statistical dogma. As outlined above, taxometrics analyses require high measurement precision and indicator validity to detect fairly large effects by psychological standards. Thus, CCKs cannot distinguish types from continua for traits or disorders with no associated indicators of large effect size. Although there are a number of reasons to suspect that many if not most traits are distributed continuously (Klein & Riso, 1993), negative results from taxometrics analyses do not provide strong support for the continuity of a trait any more than positive results from cluster analyses provide strong support for typologies.

Reducing candidate indicators and examining multiple consistency tests

Provided that the construct being assessed is discrete, some markers are typically more precise than others, and therefore do a better job of differentiating taxon group members from nontaxon group members, even if great care is taken in selecting indicators. Because of this, Meehl has suggested that candidate variables be screened first using MAMBAC, and that only those that appear to mark a latent taxon be subjected to subsequent MAXCOV analyses (Meehl, 1995, 2001). Typically, this is the strategy used to filter an initial group of candidate indicators into a smaller number that are more efficient at differentiating between groups (e.g., Waller et al., 1996). This process is necessary because if even a minority of indicators are inefficient discriminators, the results from multiple MAXCOV runs will be inconsistent and existing latent taxa will go undetected (e.g., Beauchaine & Beauchaine, 2002; Meehl, 1995).

Some elaboration on this two-stage procedure is warranted, given that it has been criticized as a means of stacking the deck in favor of taxonic outcomes (Widiger, 2001). Provided that one uses variables from several levels of analysis and examines multiple consistency tests (Meehl, 1995; Waller & Meehl, 1998), this is unlikely to be the case. Consistency tests refer to comparisons of CCK-derived values of a host of parameters describing the latent distribution of a taxon. Suppose, for example, that I begin with 15 putative indicators of schizotypy, which are winnowed down to 7 through initial MAMBAC analyses. The first consistency test is to examine the base rate estimates of the taxon in each of the MAMBAC runs. Seven indicators yield $2\binom{k}{2}$ or 42 combinations of variables for analysis (each indicator within a variable pair can be used as both input and output). If there is a true latent taxon present, then valid indicators should produce consistent estimates of the taxon base rate, regardless of the variable pairings. Thus, only when a preponderance of base rates fall within a small range of values is the taxon hypothesis tentatively supported. If such consistency is found across MAMBAC runs, then the indicators are subjected to MAXCOV analyses. For the 7 indicator case, there are

$$i \times \frac{(i-1)!}{(i-3)!2!},$$

or 105 trivariate combinations available for analysis. Each of these combinations produces estimates of the taxon base rate, the taxon group mean, the HITMAX value, and the true and false positive rates. Again, only when a preponderance of these parameters cluster around similar values is a taxonic interpretation supported. In addition, recent work suggests that a goodness of fit index (Jöreskog & Sörbom, 2001) exceeding .90 is highly suggestive of taxonicity (Waller & Meehl, 1998).

The likelihood of several indicators drawn

from biological, psychological, and etiological levels of analysis converging on single base rates, HITMAX values, and taxon group means across 105 separate MAXCOV runs is exceedingly low if there is no taxon present. In fact, Monte Carlo simulations of 100s of 1000s of MAXCOV runs suggest that false positive taxonic findings occur at a negligible rate when valid indicators are used (Beauchaine & Beauchaine, 2002). Thus, although Widiger (2001) may be correct in suggesting that taxonic distributions of attitudes are common, the same cannot be said for basic behavioral traits that are measured at several levels of analysis.

Sampling

When conducting taxometrics research, both sample size and sampling procedures must be considered. Monte Carlo simulations suggest that the absolute minimum sample size for reliable resolution of taxonic structure is around 200 with valid indicators, but sample sizes of 300 or more are preferred (Beauchaine & Beauchaine, 2002; Meehl, 1995). Unfortunately, this precludes taxometric searches with many data sets, aside from those derived from epidemiological samples. Moreover, there are no shortcuts around the sample size issue; the more participants the better when conducting taxometrics. Nevertheless, if variables are selected carefully across studies and equivalent sampling procedures are used, data sets may be combined toward addressing taxonic questions. In some laboratories, this may mean combining data collected across several years of study.

Finally, samples must be selected that do not generate taxonic structure artificially. The term pseudotaxonicity has been used to refer to false positive taxonic findings resulting from biases in sample selection (e.g., Beauchaine & Waters, 2003; Brown, 2001; Meehl, 1996). If, for example, one recruits a group of psychiatrically impaired individuals who exceed an extreme threshold of symptomatology and compares them with a normative control group, then it should be no surprise if CCKs reveal a latent taxon. Confirmation of taxonicity in a sample recruited for bimodality is simply tautological. Although this may seem obvious, evidence of taxonic structure in eating disorder symptoms may derive from such a sampling bias. Williamson, Womble, Smeets, Netemeyer, Thaw, Kutlesic, and Gleaves (2002) compared *DSM-IV* symptoms from 201 women diagnosed as having an eating disorder with those from 116 normal-weight controls. Members of the former group were all receiving treatment for an eating disorder, whereas members of the latter group were recruited from psychology classes. Based on taxometric analyses, the authors concluded that bulimia and binge eating disorder are discrete syndromes. The appropriate sampling procedure for testing taxonic hypotheses is to recruit across a wide range of symptomatology for a given disorder, with a *representative* proportion of participants at each symptom level. Although some authors have suggested that the ultimate utility of CCKs depends upon their ability to differentiate real taxa from pseudotaxa (e.g., Waldman & Lilienfeld, 2001), no statistical procedure is immune to sampling biases and holding CCKs to a higher standard than other inferential statistics is probably unjustified. Thus, we need to think carefully about the nature of the data we subject to taxometric analyses or any other method.

Taxometrics and Developmental Psychopathology

Putting the requirements and limitations of taxometric methods aside, what specific areas of interest to developmental psychopathologists can be addressed by using the techniques? This question is best answered by first considering that developmental psychopathology is an integrative discipline concerned with complex processes through which multiple causal agents interact to produce diverse adjustment outcomes (Cicchetti, 1993; Cicchetti & Rogosch, 2002; Cicchetti & Toth, 1998). Because this framework places an emphasis on individual coping in response to intricate combinations of psychological, biological, and environmental risk, person-centered approaches to studying psychopathology have received increasing attention in the field (e.g., Caspi, 1998; Cicchetti

& Rogosch, 1999; Richters, 1997). Given that taxometric methods represent a person-centered approach, it is all the more surprising that they have gone largely unnoticed by developmental psychopathologists. Moreover, a number of important issues emerge from the developmental psychopathology perspective that can be addressed in part through well-conceived taxometrics studies. Several of these are presented below.

Identifying children who are at risk for future psychopathology

As noted above, at least two discretely distributed traits that indicate elevated risk for psychopathology have been discovered in children through taxometrics investigations. Erlenmeyer–Kimling et al. (1989) demonstrated that measures of cognitive functioning and neuromotor performance can identify schizotypy in 7- to 12-year-olds. Moreover, nearly 50% of children with at least one parent with schizophrenia belonged to the schizotypy taxon group, compared with only 4% of age-matched controls. Thus, among children known to be at heightened risk for developing schizophrenia because one or both parents are afflicted, taxometric methods can pinpoint those individuals of particular vulnerability. Similar results have been reported in adolescent samples (Tyrka et al., 1995). I have already noted the potential implications that these findings have for prevention, given the role of familial and other environmental influences on the onset and expression of schizophrenia.

The second high-risk trait that has been identified in children is behavioral reactivity. In a taxometric investigation of the responses of 4-month-old infants to a series of visual, auditory, and olfactory stimuli, Woodward et al. (2000) identified an extremely reactive taxon group, who engaged in more arching, crying, hyperextension, and leg movements during stimulus presentations than other infants. Furthermore, the 10% of infants who belonged to the taxon group were behaviorally inhibited at age 4.5, marking a potential predisposition to later anxiety disorders (see Kagan, 1994). Here again, the ability to identify at-risk children, particularly at such a young age, may have important implications for prevention and intervention.

Identifying subtypes of disorders

Since its inception about two decades ago, a core theme in developmental psychopathology has been that observable syndromes arise from many developmental pathways and etiological causes and that our current diagnostic system fails to capture heterogeneity among individuals within psychiatric classes (e.g., Cicchetti & Rogosch, 1996). Some cases of depression, for example, appear to be influenced more by environmental risk, whereas others appear to be influenced more by biological risk (see Cicchetti & Rogosch, 2002; Harrington, Rutter, & Fombonne, 1996). Moreover, these depression subtypes cannot be differentiated based solely on the behavioral criteria outlined in the *DSM-IV*. Given this, it is curious that potential biological markers of depression and other psychiatric syndromes are often eschewed in favor of strictly behavioral criteria. The dexamethasone suppression test (DST), for example, an indicator of hypothalamic–pituitary–adrenal axis (HPA) reactivity, was once considered a promising biological marker of endogenous depression. Due to moderate specificity, however, some authors concluded that clinical symptoms are better indicators of the endogenous depression construct (e.g., Casat & Powell, 1988; Lu, Ho, Huang, & Lin, 1988). Thus, rather than entertaining the possibility that behavioral symptoms lack specificity in identifying biologically based depressions, the utility of the marker was questioned and a potential opportunity to refine our diagnostic system was lost. More recent evidence suggests that depressed patients who exhibit dexamethasone nonsupression are at nearly 10 times the risk for future suicide than those who exhibit normal DST results (Coryell & Schlesser, 2001). Thus, abnormal HPA reactivity may indeed mark a more virulent subtype of depression.

It should be emphasized that treating biological markers as subordinate to behavioral symptoms makes little sense scientifically. Be-

cause behavioral symptoms and biological markers are both manifest indicators of latent psychopathology constructs, there is no a priori reason to suspect that the former are more valid or reliable than the latter. Nevertheless, behavioral syndromes are often used as the gold standard against which the validities of biological markers are judged. In taxometrics research, no such preference exists. Rather, any available indicator of adequate validity, either behavioral or biological, can and should be used in efforts to identify latent subgroups within current diagnostic categories. In the case of depression, promising candidates for use as indicators in taxometric analyses include HPA reactivity (e.g., Coryell & Schlesser, 2001) and longitudinal symptom course (e.g., Cole et al., 2002), both of which are more precise than cross-sectional symptom patterns (see above). Symptoms of CD might also serve as an indicator in differentiating between depression subgroups. This follows from the contention that depressions with and without CD symptoms are etiologically distinct, with the latter deriving more from biological determinants (e.g., Harrington, Fudge, Rutter, Pickles, & Hill, 1991; Meller & Borchard, 1996; Panak & Garber, 1992). If taxometric analyses indicate discrete depression subtypes, then a modification of our current diagnostic system might be warranted. Preliminary evidence suggests that endogenous depression may represent a discrete class among both adolescents and adults (Ambrosini et al., 2002; Beach & Amir, 2003; Haslam & Beck, 1994), although analyses to date have been conducted exclusively with self-report indicators. If discrete endogenous and exogenous depression subtypes are confirmed, it does not imply that environmentally induced or influenced depressions are unimportant or undeserving of treatment. We are simply in a better position to formulate effective interventions when we know as much as possible about the etiological origins of a disorder (Cicchetti & Hinshaw, 2002), and different treatment approaches may be required for different subtypes of depression.

A second diagnostic question that might be addressed through the use of formal taxometric methods is whether autistic disorder and Asperger's disorder reflect discrete behavioral syndromes or fall at different points along a continuum of an autistic spectrum. This question has received considerable attention in the child psychopathology literature since the appearance of Asperger's disorder in the *DSM-IV* and has yet to be resolved (Volkmar, Klin, Schultz, Rubin, & Bronen, 2000). Both disorders are characterized by impaired social interaction, stereotyped behaviors, and restricted interests and activities. Autism is diagnosed when these symptoms are accompanied by delays in verbal communication and/or cognitive development. When delays are not observed in either of these domains, Asperger's disorder is diagnosed.

Based on different concordance rates across disorders in the first degree relatives of probands (Volkmar, Klin, & Pauls, 1998), different patterns of comorbidity across diagnostic groups (Ghaziuddin, Weidmer–Mikhail, & Ghaziuddin, 1998), and the aforementioned differences in patterns of verbal and nonverbal skills (e.g., Klin, Volkmar, Sparrow, Cicchetti, & Roarke, 1995; Volkmar et al., 1994), some authors have concluded that autistic disorder and Asperger's disorder are indeed discrete (see also Klin, 1994; Ozonoff, Rogers, & Pennington, 1991). Others, however, have argued for a continuous autism spectrum model based on similarities in clinical features (e.g., Eisenmajer, Prior, Leekam, Wing, Gould, Welham, & Ong, 1996), neuropsychological test performance (Miller & Ozonoff, 2000), and brain stem abnormities (Bauman, 1996).

As noted by Meehl (1995), taxonomic questions such as these are irresolvable if not tested empirically by subjecting phenotypic indicators to formal taxometric analyses. Autism spectrum disorders provide an especially ripe opportunity for testing taxonic conjectures because the autism phenotype has been specified at several levels of analysis, with a number of precise measures (see Akshoomoff, Pierce, & Courchesne, 2002; Dawson, Webb, Schellenberg, Dager, Friedman, Aylward, & Richards, 2002; Klin, Jones, Schultz, Volkmar, & Cohen, 2002). At the psychophysiological level, for example, children with autism spectrum disorders exhibit attenuated event-related potentials when presented with

familiar faces (Dawson, Carver, Meltzoff, Panagiotides, McPartland, & Webb, 2002). At the social level, infants with autism direct their gaze toward others less frequently and orient to their names less often than controls (Osterling, Dawson, & Munson, 2002). Moreover, adults with autism attend less to the eyes of others when observing social interactions (Klin et al., 2002). Finally, accelerated trajectories in brain growth are observed in autism, with normal brain volume at birth but larger than normal brain volume emerging between ages 2 and 4 years (Courchesne et al., 2001). As noted earlier, developmental trajectories such as these provide increased precision over cross-sectional measures, and may therefore be particularly useful in taxometric studies. Note that these are but three examples from a number of potential indicators of the autism phenotype (see Dawson, Webb, et al., 2002). If these traits are distributed discretely, thereby indicating a typological distinction between autistic disorder and Asperger's disorder, follow-up studies aimed at elucidating differential etiological mechanisms should be pursued. As noted above, such differences are likely to have implications for both treatment and long-term course.

There are additional examples of diagnostic categories that might be refined based on taxometrics investigations. One clear candidate is attention-deficit/hyperactivity disorder (ADHD), where a debate exists in the literature regarding the distinctiveness of the *combined* type (ADHD/C), which is characterized by hyperactivity/impulsivity and inattention, and the *inattentive* type (ADHD/I), which is characterized primarily by the latter. Some authors have argued that the disorders are likely to be discrete (e.g., Milich, Balentine, & Lynam, 2001), whereas others have argued that current data are inconclusive (e.g., Barkley, 2001; Hinshaw, 2001; Lahey, 2001). Evidence offered for discrete subtypes has come from factor analytic studies that differentiate among symptoms and from cluster analytic studies that differentiate among children based on symptoms. As noted earlier, these methods are not suitable for distinguishing types from continua, and results have not been fully consistent across studies (see Milich et al., 2001). In addition, few studies have included information about biological markers or symptom course. However, the question of discreteness could be addressed by subjecting carefully selected variables to taxometric analyses. Moreover, given the size and comprehensiveness of existing data sets, taxometric analyses of ADHD might be possible sooner rather than later.

Similar to the case of depression, putative indicators of inattention should extend beyond *DSM-IV* criteria. One potential indicator is *sluggish cognitive tempo* (SCT), a behavioral construct comprising symptoms of lethargy, daydreaming, drowsiness, and hypoactivity (Carlson & Mann, 2002; Lahey, Carlson, & Frick, 1997). Despite correlating specifically with ADHD/I in the *DSM-IV* field trials, SCT items were dropped from the final criterion list because the workgroup sought a single set of inattentivity items for both ADHD subtypes (Carlson & Mann, 2002; Frick et al., 1994). Nevertheless, measures of SCT are elevated among ADHD/I children compared with other ADHD groups (McBurnett, Pfiffner, & Frick, 2001; Milich et al., 2001). Although this does not guarantee that ADHD/I marks a discrete disorder, it does suggest that markers of SCT might be useful in taxometrics investigations.

It should be noted, however, that SCT and inattention are not fully overlapping and therefore cannot be used as proxies for one another. Rather, SCT appears to mark a subset of ADHD/I children who are also differentiated from ADHD/C children on other symptoms, including lower levels of externalizing behaviors and higher levels of anxiety, depression, and social withdrawal (Carlson & Mann, 2002). Indeed, Carlson and Mann concluded that ADHD/I children high on SCT represent a distinct diagnostic group, whereas ADHD/I children low on SCT are more similar to children with ADHD/C. If taxometrics investigations sort children along these lines, a new diagnostic category might be indicated (see also Milich et al., 2001).

Research on the biology of ADHD also suggests some potential indicators for taxometric investigations. For example, several studies have demonstrated reduced urinary 3-methoxy-4-hydroxyphenylglycol (MHPG), a norepinephrine metabolite, among children with ADHD

(e.g., Shekim, Dekirmenjian, Chapel, & Davis, 1982; Shekim, Sinclair, Glaser, Horwitz, Javaid, & Bylund, 1987). Moreover, MHPG deficiencies have been tied specifically to the biobehavioral substrates of impulsivity (see Beauchaine, 2001), and may therefore be associated more strongly with the ADHD/C subtype. Consistent with this interpretation, anxious children exhibit increased urinary MHPG (Kagan, Reznick, & Snidman, 1987). Given the elevated symptoms of anxiety observed in ADHD/I children (see Hinshaw, 2002a), particularly those with SCT (Carlson & Mann, 2002), urinary MHPG may differentiate between ADHD subgroups in taxometric analyses.

Finally, several studies have suggested that the ADHD/C subtype is characterized by earlier ages of onset and referral than the ADHD/I subtype (e.g., Lahey et al., 1997; Faraone, Biederman, Weber, & Russell, 1998). One potential explanation for this is that hyperactive/impulsive behaviors become unmanageable for parents and teachers before symptoms of inattention are noticed (Milich et al., 2001). If this explanation is ruled out in future studies, age of onset might also be used as an indicator in taxometric analyses of ADHD subtypes.

Other markers are unlikely to serve as suitable indicators in taxometric analyses, either because they do not differentiate between ADHD/C and ADHD/I subgroups, or because they are of small effect size. For example, both the *DAT1* and *DRD4* dopamine genes have been identified as candidates in contributing to the ADHD phenotype. In theory, allelic variants of these genes should contribute specifically to hyperactivity/impulsivity and be related more strongly to ADHD/C than to ADHD/I. However, both population- and family-based studies suggest relatively low risk values, indicating modest genetic associations between *DAT1* and *DRD4* allelic variants and symptoms of hyperactivity/impulsivity (e.g., LaHoste et al., 1996; Smalley et al., 1998). Thus, effect sizes may be inadequate for use in taxometric investigations.

Locating bifurcation points in the development of discrete traits

As noted above, there are several psychopathological traits that appear to be distributed as discrete latent classes among adults, including schizotypy (e.g., Blanchard et al., 2000; Korfine & Lenzenweger, 1995; Lenzenweger, 1999; Tyrka, Cannon, et al., 1995), dissociative experiences (Waller et al., 1996; Waller & Ross, 1997), and psychopathy (Harris et al., 1994). What is not clear is when in development these traits emerge as discrete entities. The case of psychopathy provides a particularly good example, as there has been much speculation about the developmental precursors of the construct and whether fledgling psychopaths can be identified premorbidly. One school of thought is that comorbid conduct problems and hyperactivity/impulsivity mark future psychopathy (e.g., Lynam, 1996, 1998). Because the prevalence rate of ADHD among children and adolescents with CD approaches 70%, however (e.g., Klein et al., 1997), it is unlikely that comorbid ADHD marks a predisposition for psychopathy with enough specificity to be used as an indicator in taxometrics research. As noted earlier, variables that hold more promise include callous unemotional traits and physiological markers of underarousal (e.g., Barry et al., 2000; Beauchaine, 2001; Beauchaine et al., 2001; Frick et al., 2000).

Determining when the trajectory toward psychopathy emerges as a discrete class will require carefully designed taxometrics studies with progressively younger samples[4] and may have important implications for treatment. It is well established that interventions for serious conduct problems are of limited effect by the time probands reach adolescence (e.g., Dishion & Patterson, 1992; Ruma, Burke, & Thompson, 1996), but are more successful with younger children (e.g., Webster–Stratton & Hammond, 1997; Webster–Stratton, Reid, & Hammond, 2002). Moreover, biological markers that were once considered to be stable indices of a diathesis for severe conduct problems now appear to be malleable in very young children. For example, Raine, Venables, Dalais, Mellingen, Reynolds, and Mednick (2001) reported that an enriched preschool environment, including child social skills train-

4. Preliminary evidence suggests that a predisposition toward psychopathy may be distributed as a discrete latent class among children in Grades 4–8 (Skilling et al., 2001).

ing, extensive parental involvement, thorough instruction of teachers in behavioral management, and weekly counseling sessions for parents, conferred a 61% increase in electrodermal activity on children 6–8 years later, compared with controls who were assigned randomly to a no treatment condition. These data suggest that long-term changes in the functioning of biological systems implicated in aggression can be effected through multifaceted interventions in the preschool years. After such systems consolidate into discretely distributed patterns of functioning, however, interventions may simply be too late. Identifying the point at which bifurcation into discrete trajectories occurs may therefore inform efforts to prevent and/or alter psychopathic traits. Although some might assume that genetically influenced disorders are characterized by clear, immutable biological markers very early in life, this is typically not the case. There is increasing recognition, for example, that environmental factors contribute to the expression of genes and that genetic effects on behavior increase across the life span, in part due to exposure to accumulated environmental risk (see, e.g., Goldsmith, Gottesman, & Lemery, 1997; Rutter, Dunn, Plomin, Simonoff, Pickles, Maughan, Ormel, Meyer, & Eaves, 1997). For example, children who are impulsive, a trait that is highly heritable, are more likely to develop conduct problems when placed in high risk environments (Burt, Krueger, McGue, & Iacono, 2001; Lynam, Caspi, Moffitt, Wikström, Loeber, & Novak, 2000; Patterson, DeGarmo, & Knutson, 2000). Interventions that occur prior to the behavioral expression of a genetically influenced trait may therefore hold considerable promise in altering trajectories toward psychopathology. In contrast, interventions delivered after a threshold of accumulated risk has been reached may not prevent or alter the expression of a genetic liability. To the extent that discretely distributed traits mark genetic liabilities for psychopathology (e.g., Meehl, 1995), identifying bifurcation points may help us converge on ages for optimal intervention effects.

Identifying moderators of treatment outcome

Taxometric investigations can also help to identify person-specific moderators of treatment outcome. In the context of intervention research, a moderator is any variable present at baseline that discriminates among subgroups of individuals who respond differentially to treatment (e.g., Kraemer, Stice, Kazdin, Offord, & Kupfer, 2001). Because CCKs identify subpopulations who are different in kind, they can provide us with non-arbitrary groups for whom differential treatment response might be assessed. This in turn allows us to zero in on causal processes that enhance or diminish the impact of an intervention in different groups (see Hinshaw, 2002b). Indeed, in their recent report on effective interventions for conduct problems, Brestan and Eyberg (1998) noted that a considerable challenge facing the field in formulating the next generation of treatments is to identify those children who are not served by current interventions and to determine what additional resources they require. It is known, for example, that one-third of children do not benefit from the most successful interventions for conduct problems (Webster–Stratton & Hammond, 1997; Webster–Stratton et al., 2002). Yet little is known about child-specific predictors of treatment response. Because well-conducted taxometrics research includes specification of individual characteristics at behavioral, biological, psychological, and etiological levels, explorations of moderation will necessarily be enriched compared with the current strategy of subgrouping children based on often arbitrary distinctions in behavioral symptoms alone.

In addition, treatment response itself might become an indicator in taxometric studies. In the case of psychopathy, for example, treatment resistance may be an additional marker of an underlying trait that results in emotional unresponsiveness, physiological underarousal, and a deficiency in learning from corrective experiences. All of these characteristics make it difficult to establish therapeutic leverage and may result in treatment failure. By tracking a number of theoretically important indicators throughout the intervention process, a more refined basic understanding of the characteristics of treatment nonresponders will be attained. This knowledge can in turn be used to identify treatment-resistant individuals prospectively and to channel them away from interventions that are unlikely to benefit them

and toward new interventions targeting their specific behavioral traits.

Elucidating mechanisms of equifinality and multifinality

As noted above, equifinality suggests that a given disorder can be the end state of numerous developmental pathways, and multifinality suggests that children in similar high-risk situations can diverge toward quite disparate end states, only some of which will be disordered (Cicchetti & Rogosch, 1996). This framework implies that a greater understanding of the diversity of individual outcomes will be attained through longitudinal analyses of child-specific variables and how these interact with environmental experiences to produce disordered and nondisordered outcomes. Thus, similar to the case of identifying moderators of treatment, elucidating mechanisms of equifinality and multifinality requires that homogeneous subgroups of individuals who are at differential risk for psychopathology be identified and followed throughout the natural course of development (see Hinshaw, 2002b; Richters & Cicchetti, 1993). Recall that Woodward et al. (2000) reported evidence that behavioral reactivity is distributed discretely in childhood, with roughly 10% of 4-month-old infants falling into an extremely reactive taxon group who scored high on measures of behavioral inhibition 4 years later. Given that the taxon can be identified at a very young age, these findings provide a unique opportunity to explore longitudinal outcomes among qualitatively distinct groups of children who are at differential biological risk for later psychopathology (see Kagan, 1997).

A number of questions emerge from the Woodward et al. (2000) findings. What sets of environmental experiences (e.g., family, school, neighborhood) result in taxon group members developing later anxiety disorders or depression? What sets of environmental experiences protect taxon group members from developing later psychopathology? What differences in environmental experiences result in equifinal outcomes for taxon group members compared with nontaxon group members? What differences in environmental experiences result in multifinal outcomes for members within the taxon and nontaxon groups? These and other questions take on considerably more meaning when subgroups of children are defined by nonarbitrary cutoffs. Answering each question will require carefully designed studies evaluating multiple risk models for child psychopathology, an approach that has already proven fruitful with normative samples (e.g., Lengua, 2002).

Summary and Recommendations for Future Research

To date, taxometric methods have been used quite sparingly by developmental psychopathologists. One probable reason for this is a well-founded wariness of categorical classification systems, which fail to account for within-group heterogeneity in biological, developmental, etiological, and cultural influences on behavior. I have argued, however, that the limitations of categorizing have very little to do with the ontological status of specific traits or disorders as dimensional versus discrete. I have argued further that developmental psychopathologists should be concerned with identifying discrete behavioral syndromes, because doing so could result in a number of research advances, both basic and applied. These include (a) identifying children who are at increased risk for developing psychopathology, (b) identifying discrete subtypes of disorders within current diagnostic classes, (c) locating sensitive periods in the development of discrete pathological traits, (d) discovering moderators of treatment-outcome, and (e) elucidating mechanisms of equifinality and multifinality. Finally, I have provided several examples of diagnostic classes for which our understanding might be refined through formal taxometrics investigations, including anxiety disorders, ADHD, autism spectrum disorders, conduct problems, depression, and schizophrenia.

However, despite the potential utility of taxometric methods in clarifying some of the most fundamental diagnostic questions in psychopathology, they are unlikely to be used in the future by developmental psychopathologists if two barriers are not overcome. First, there has been a lack of communication across interdisciplinary boundaries, with almost all taxometric investigations appearing in the adult

psychopathology literature. Thus, many developmental psychopathologists may simply be unaware of the techniques. Although this appears to be changing based on a small set of recent taxometrics studies with juvenile samples (Fraley & Spieker, in press; Skilling et al., 2001; Woodward et al., 2000), increased interdisciplinary communication is essential if the pitfalls of taxometrics research identified in the adult psychopathology literature are to be avoided. As reviewed in previous sections, these include problems with invalid variable selection, inadequate measurement precision, misinterpretation of null findings, and misguided sampling procedures, among others. Developmental psychopathologists are in a unique position to benefit from the work of adult psychopathologists, and it would be unfortunate indeed if lessons learned in their research efforts were ignored.

Perhaps the most critical issue identified to date is the need for taxometric hypotheses to be tested in studies that are carefully and prospectively planned. As outlined in earlier sections, much of the extant taxometrics research has been conducted with datasets of convenience, using primarily Likert-type scales as indicators. Although understandable given the large sample sizes required for taxometric analyses, such ad hoc approaches are unlikely to provide variables from multiple levels of analysis that are of adequate precision for testing taxonic conjectures. Rather, putative markers of a hypothesized latent taxon must be selected in advance based on strong theory, and must be measured with minimal error. As such, taxometrics studies are likely to be informed best by interdisciplinary research teams, who will be in the greatest position to select appropriately precise indicators from behavioral, biological, developmental, physiological, and etiological levels of analysis.

A second barrier to conducting taxometrics research has been a lack of readily available software for performing analyses. In fact, there are currently no statistics packages that offer a taxometrics module. Thus, interested researchers have been forced to write their own software, a somewhat daunting task given the complexity of CCK algorithms (see, e.g., Meehl & Yonce, 1994, 1996; Waller & Meehl, 1998). However, several taxometrics programs have been written by the Waller, Meehl, and Yonce group, and are now available on their taxometrics home page (http://peabody.vanderbilt.edu/depts/psych_and_hd/faculty/wallern/tx.html). These programs run in the R language, which is both free and relatively easy to learn. Our group has also posted free taxometrics programs on our website, which are written in Statistica (1998) BASIC (http://tbeauchaine.psych.washington.edu/tb/cbs/taxometrics.htm). Moreover, both the Waller group and our group are currently developing additional programs that should make taxometric methods accessible to a broadened range of researchers.

Taxometric methods are now 30 years old. Their advantages, limitations, and operating characteristics have been outlined in a number of theoretical, empirical, and simulation studies. It is hoped that developmental psychopathologists will draw from the lessons learned from these studies in designing methodologically rigorous and theoretically informed taxometrics research that will advance the field in the emergent 21st century.

References

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

Achenbach, T. M. (1993). *Empirically based taxonomy*. Burlington, VT: University of Vermont Department of Psychiatry.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the Strange Situation*. Hillsdale, NJ: Erlbaum.

Akshoomoff, N., Pierce, K., & Courchesne, E. (2002). The neurobiological basis of autism from a developmental perspective. *Development and Psychopathology, 14,* 613–634.

Ambrosini, P. J., Bennett, D. S., Cleland, C. M., & Haslam, N. (2002). Taxonicity of adolescent melancholia: A categorical or dimensional construct? *Journal of Psychiatric Research, 36,* 247–256.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: Author.

Atlas, R. S., & Overall, J. E. (1994). Comparative evaluation of two superior stopping rules for hierarchical cluster analysis. *Psychometrika, 59,* 581–591.

Avila, M. T., McMahon, R. P., Elliott, A. R., & Thaker, G. K. (2002). Neuropsychological markers of vulnerability to schizophrenia: Sensitivity and specificity of specific quantitative eye movement measures. *Journal of Abnormal Psychology, 111,* 259–267.

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.

Bauman, M. L. (1996). Neuroanotomic observations of the brain in pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 26,* 199–203.

Barry, C. T., Frick, P. J., DeShazo, T. M., McCoy, M. Ellis, M., & Loney, B. R. (2000). The importance of callous–unemotional traits for extending the concept of psychopathy to children. *Journal of Abnormal Psychology, 109,* 335–340.

Beach, S. R., & Amir, N. (2003). Is depression taxonic, dimensional, or both? *Journal of Abnormal Psychology, 112,* 228–236.

Beauchaine, T. P. (1999). Definitions and levels of analysis. *American Psychologist, 54,* 439–440.

Beauchaine, T. P. (2001). Vagal tone, development, and Gray's motivational theory: Toward an integrated model of autonomic nervous system functioning in psychopathology. *Development and Psychopathology, 13,* 183–214.

Beauchaine, T. P., & Beauchaine III, 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., Katkin, E. S., Strassberg, Z., & Snarr, J. (2001). Disinhibitory psychopathology in male adolescents: Discriminating conduct disorder from attention-deficit/hyperactivity disorder through concurrent assessment of multiple autonomic states. *Journal of Abnormal Psychology, 110,* 610–624.

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.

Blanchard, J. J., Gangestad, S. W., Brown, S. A., & Horan, W. P. (2000). Hedonic capacity and schizotypy revisited: A taxometric analysis of social anhedonia. *Journal of Abnormal Psychology, 109,* 87–95.

Blashfield, R. K., & Aldenderfer, M. S. (1988). The methods and problems of cluster analysis. In J. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 447–474). New York: Plenum Press.

Bremner, J. D., & Vermetten, E. (2001). Stress and development: Behavioral and biological consequences. *Development and Psychopathology, 13,* 473–489.

Brestan, E. V., & Eyberg, S. M. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *Journal of Clinical Child Psychology, 27,* 180–189.

Brooks, D. S., Murphy, D., Janota, I., & Lishman, W. A. (1987). Early-onset Huntington's chorea. *British Journal of Psychiatry, 151,* 850–852.

Brown, T. A. (2001). Taxometric methods and the classification and comorbidity of mental disorders: Methodological and conceptual consideration. *Clinical Psychology Science and Practice, 8,* 534–541.

Burt, S. A., Krueger, R. F., McGue, M., & Iacono, W. G. (2001). Sources of covariation among attention-deficit/hyperactivity disorder, oppositional defiant disorder, and conduct disorder: The importance of shared environment. *Journal of Abnormal Psychology, 110,* 516–525.

Cantor, N., & Genero, N. (1986). Psychiatric diagnosis and natural categorization: A close analogy. In T. Milton & G. Klerman (Eds.), *Contemporary directions in psychopathology: Toward the DSM-IV* (pp. 233–256). New York: Guilford Press.

Cantor, N., & Mischel, W. (1979). Prototypes in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 12). Orlando, FL: Academic Press.

Cantwell, D. P. (1996). Classification of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry, 37,* 3–12.

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

Casat, C. D., & Powell, K. (1988). The dexamethasone suppression test in children and adolescents with major depressive disorder: A review. *Journal of Clinical Psychiatry, 10,* 390–393.

Caspi, A. (1998). Personality development across the life course. In W. Damon (Series Ed.) & N. Eisenberg (Ed.). *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 312–371). New York: Wiley.

Cicchetti, D. (1993). Developmental psychopathology: Reactions, reflections, projections. *Developmental Review, 13,* 471–502.

Cicchetti, D., & Dawson, G. (Eds.). (2002). Multiple levels of analysis [Special issue]. *Development and Psychopathology, 14*(3).

Cicchetti, D., & Hinshaw, S. P. (Eds.). (2002). Prevention and intervention science: Implications for developmental theory [Special Issue]. *Development and Psychopathology, 14*(4).

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology, 8,* 597–600.

Cicchetti, D., & Rogosch, F. A. (1999). Conceptual and methodological issues in developmental psychopathology research. In P. C. Kendall, J. N. Butcher, & G. N. Holmbeck (Eds.), *Handbook of research methods in clinical psychology* (2nd ed., pp. 433–465). New York: Wiley.

Cicchetti, D., & Rogosch, F. A. (2002). A developmental psychopathology perspective on adolescence. *Journal of Consulting and Clinical Psychology, 70,* 6–20.

Cicchetti, D., & Toth, S. L. (1998). Perspectives on research and practice in developmental psychopathology. In W. Damon (Series Ed.), I. E. Sigel, & K. A. Renninger (Vol. Eds.), *Handbook of child psychopathology: Vol 4. Child psychology in practice* (5th ed., pp. 479–583). New York: Wiley.

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification in psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology, 46,* 121–153.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Cole, D. A., Tram, J. M., Martin, J. M., Hoffman, K. B., Ruiz, M. D., Jacquez, F. M., & Maschman, T. L. (2002). Individual differences in the emergence of depressive symptoms in children and adolescents: A longitudinal investigation of parent and child reports. *Journal of Abnormal Psychology, 111,* 156–165.

Cornblatt, B., Obuchowski, M., Roberts, S., Pollack, S., & Erlenmeyer–Kimling, L. (1999). Cognitive and

behavioral precursors of schizophrenia. *Development and Psychopathology, 11,* 487–508.

Coryell, W., & Schlesser, M. (2001). The dexamethasone suppression test and suicide prediction. *American Journal of Psychiatry, 158,* 748–753.

Courchesne, E., Karns, C., Davis, H. R., Ziccardi, R., Carper, R., Tigue, Z., Chisum, H. J., Moses, P., Pierce, K., Lord, C., Lincoln, A. J., Pizzo, S., Schreibman, L., Haas, R. H., Akshoomoff, N., & Courchesne, R. Y. (2001). Unusual brain growth patterns in early life in patients with autistic disorder. *Neurology, 57,* 245–254.

Cummings, E. M., Davies, P. T., & Campbell, S. B. (2000). *Developmental psychopathology and family process*. New York: Guilford Press.

Curtis, C. E., Calkins, M. E., Grove, W. M., Feil, K. J., & Iacono, W. G. (2001). Saccadic disinhibition in patients with acute and remitted schizophrenia and their first-degree biological relatives. *American Journal of Psychiatry, 158,* 100–106.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243,* 1668–1674.

Dawson, G., Carver, L., Meltzoff, A. N., Panagiotides, H., McPartland, J., & Webb, S. J. (2002). Neural correlates of face and object recognition in young children with autism spectrum disorder, developmental delay, and typical development. *Child Development, 73,* 700–717.

Dawson, G., Webb, S., Schellenberg, G. D., Dager, S., Friedman, S., Aylward, E., & Richards, T. (2002). Defining the broader phenotype of autism: Genetic, brain, and behavioral perspectives. *Development and Psychopathology, 14,* 581–611.

Dishion, T. J., & Patterson, G. R. (1992). Age effects in parent training outcome. *Behavior Therapy, 23,* 719–729.

Dulmus, C. N., & Smyth, N. J. (2000). Early-onset schizophrenia: A literature review of empirically based interventions. *Child and Adolescent Social Work Journal, 17,* 55–69.

Eisenmajer, R., Prior, M., Leekam, S., Wing, L., Gould, J., Welham, M., & Ong, B. (1996). Comparison of clinical symptoms in autism and Asperger's disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 1523–1531.

Erlenmeyer–Kimling, L., Golden, R. R., & Cornblatt, B. A. (1989). A taxometric analysis of cognitive and neuromotor variables in children at risk for schizophrenia. *Journal of Abnormal Psychology, 98,* 203–208.

Falloon, I. R. H., Boyd, J. L., McGill, C. W., Williamson, M., Razani, J., Moss, H. B., Gilderman, A. M., & Simpson, G. M. (1985). Family management in the prevention of morbidity in schizophrenia. *Archives of General Psychiatry, 42,* 887–896.

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.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry, 26,* 57–63.

Flanagan, E. H., & Blashfield, R. K. (2002). Psychiatric classification through the lens of ethnobiology. In L. E. Beutler & M. L. Malik (Eds.), *Rethinking the DSM* (pp. 121–145). Washington, DC: American Psychological Association.

Fleiss, J. L. (1972). Classification of the depressive disorders by numerical taxonomy. *Journal of Psychiatric Research, 9,* 141–153.

Fraley, R. C., & Spieker, S. J. (2003). Are infant attachment patterns continuously or categorically distributed? A taxometric analysis of Strange Situation behavior. *Developmental Psychology, 39,* 387–404.

Fraley, R. C., & Waller, N. G. (1998). Adult attachment patterns: A test of the typological model. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 77–114). New York: Guilford Press.

Frick, P. J., Bodin, S. D., & Barry, C. T. (2000). Psychopathic traits and conduct problems in community and clinic-referred samples of children: Further development of the Psychopathy Screening Device. *Psychological Assessment, 12,* 382–393.

Frick, P. J., Lahey, B. B., Applegate, B., Kerdyck, L., Ollendick, T., Hynd, G. W., Garfinkel, B., Greenhill, L., Biederman, J., Barkley, R., McBurnett, K., Newcorn, J., & Waldman, I. (1994). DSM-IV field trials for the disruptive behavior disorders: Symptom utility estimates. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 529–539.

Ghaziuddin, M., Weidmer–Mikhail, E., & Ghaziuddin, N. (1998). Comorbidity of Asperger syndrome: A preliminary report. *Journal of Intellectual Disabilities Research, 42,* 279–283.

Golden, R. R., & Meehl, P. E. (1979). Detection of the schizoid taxon with MMPI indicators. *Journal of Abnormal Psychology, 88,* 212–233.

Goldsmith, H. H., Gottesman, I. I., & Lemery, K. S. (1997). Epigenetic approaches to developmental psychopathology. *Development and Psychopathology, 9,* 365–387.

Grayson, D. A. (1987). Can categorical and dimensional views of psychiatric illness be distinguished? *British Journal of Psychiatry, 151,* 355–361.

Grove, W. M., & Meehl, P. E. (1993). Simple regression-based procedures for taxometric investigation. *Psychological Reports, 73,* 707–737.

Hamilton, E. B., Asarnow, J. R., & Tompson, M. C. (1999). Family interaction styles of children with depressive disorders, schizophrenia-spectrum disorders, and normal controls. *Family Process, 38,* 463–476.

Harrington, R., Fudge, H., Rutter, M., Pickles, A., & Hill, J. (1991). Adult outcome of childhood and adolescent depression: II. Links with antisocial disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 434–439.

Harrington, R., Rutter, M., & Fombonne, E. (1996). Developmental pathways in depression: Multiple meanings, antecedents, and endpoints. *Development and Psychopathology, 8,* 601–616.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1994). Psychopathy as a taxon: Evidence that psychopaths are a discrete class. *Journal of Consulting and Clinical Psychology, 62,* 387–397.

Haslam, N., & Beck, A. T. (1994). Subtyping major depression: A taxometric analysis. *Journal of Abnormal Psychology, 103,* 686–692.

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

Hinshaw, S. P. (2002a). Is ADHD an impairing condition

in childhood and adolescence? In P. S. Jensen & J. R. Cooper (Eds.). *Attention deficit hyperactivity disorder* (pp. 5.1–5.21). Kingston, NJ: Civic Research Institute.

Hinshaw, S. P. (2002b). Prevention/intervention trials and developmental theory: Commentary on the Fast Track special section. *Journal of Abnormal Child Psychology, 30,* 53–59.

Hinshaw, S. P., Lahey, B. B., & Hart, E. L. (1993). Issues of taxonomy and comorbidity in the development of conduct disorder. *Development and Psychopathology, 5,* 31–49.

Hinshaw, S. P., & Lee, S. S. (2003). Conduct and oppositional defiant disorders. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (2nd ed., pp. 144–198). New York: Guilford Press.

Hinshaw, S. P., & Park, T. (1999). Research problems and issues: Toward a more definitive science of disruptive behavior disorders. In H. C. Quay & A. E. Hogan (Eds.), *Handbook of disruptive behavior disorders* (pp. 593–620). New York: Plenum Press.

Hogarty, G. E., Anderson, C. M., Teiss, D. J., Kornblith, S. J., Greenwald, D. P., Ulrich, R. F., & Carter, M. (1991). Family psychoeducation, social skills training, and maintenance chemotherapy in the aftercare treatment of schizophrenia. *Archives of General Psychiatry, 48,* 340–347.

Houts, A. C. (2002). Discovery, invention, and the expansion of the modern *Diagnostic and Statistical Manual of Mental Disorders*. In L. E. Beutler & M. L. Malik (Eds.), *Rethinking the DSM* (pp. 17–65). Washington, DC: American Psychological Association.

Hoyt, W. T., & Kerns, M. D. (1999). Magnitude and moderators of bias in observer ratings: A meta-analysis. *Psychological Methods, 4,* 403–424.

Jensen, P. S., & Hoagwood, K. (1997). The book of names: *DSM-IV* in context. *Development and Psychopathology, 9,* 231–249.

Jöreskog, K. G., & Sörbom, D. (2001). *LISREL 8: User's reference guide*. Lincolnwood, IL: Scientific Software International.

Kagan, J. (1994). *Galen's prophecy*. New York: Basic Books.

Kagan, J. (1997). Conceptualizing psychopathology: The importance of developmental profiles. *Development and Psychopathology, 9,* 321–334.

Kagan, J., Reznick, J. S., & Snidman, N. (1987). The physiology and psychology of behavioral inhibition in children. *Child Development, 58,* 1459–1473.

Kendell, R. E. (1989). Clinical validity. *Psychological Medicine, 19,* 45–55.

Klein, D. N., & Riso, L. P. (1993). Psychiatric disorders: Problems of boundaries and comorbidity. In C. G. Costello (Ed.), *Basic issues in psychopathology* (pp. 19–66). New York: Guilford Press.

Klein, R. G., Abikoff, H., Klass, E., Ganeles, D., Seese, L. M., & Pollack, S. (1997). Clinical efficacy of methylphenidate in conduct disorder with and without attention deficit hyperactivity disorder. *Archives of General Psychiatry, 54*, 1073–1080.

Klin, A. (1994). Asperger syndrome. *Psychiatric Clinics of North America, 3,* 131–148.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Defining and quantifying the social phenotype of autism. *American Journal of Psychiatry, 159,* 895–908.

Klin, A., Volkmar, F. R., Sparrow, S. S., Cicchetti, D. V., & Roarke, B. P. (1995). Validity and neuropsychological characterization of Asperger syndrome: Convergence with nonverbal learning disabilities syndrome. *Journal of Child Psychology and Psychiatry, 36,* 1127–1140

Korfine, L., & Lenzenweger, M. F. (1995). The taxonicity of schizotypy: A replication. *Journal of Abnormal Psychology, 104,* 26–31.

Kovacs, M. (1992). *Children's Depression Inventory*. North Tonawanda, NY: Multi-Health Systems.

Kraemer, H. C., Stice, E., Kazdin, A., Offord, D., & Kupfer, D. (2001). How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *American Journal of Psychiatry, 158,* 848–856.

Krieger, A. M., & Green, P. E. (1999). A cautionary note on using internal cross validation to select the number of clusters. *Psychometrika, 64,* 341–353.

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.

Lahey, B. B., Carlson, C. L., & Frick, P. J. (1997). Attention-deficit disorder without hyperactivity. In T. A. Widiger, A. J. Frances, H. A. Pincus, R. Ross., M. B. First, & W. Davis (Eds.), *DSM-IV sourcebook* (Vol. 3, pp. 163–188). Washington, DC: American Psychiatric Association.

LaHoste, G. J., Swanson, J. M., Wigal, S. B., Glabe, C., Wigal, T., King, N., & Kennedy, J. L. (1996). Dopamine D4 receptor polymorphism is associated with attention deficit hyperactivity disorder. *Molecular Psychiatry, 1,* 121–124.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Boston: Houghton Mifflin.

Lengua, L. J. (2002). The contribution of emotionality and self-regulation to the understanding of children's response to multiple risk. *Child Development, 73,* 144–161.

Lenzenweger, M. F. (1999). Deeper into the schizotypy taxon: On the robust nature of maximum covariance analysis. *Journal of Abnormal Psychology, 108,* 182–187.

Lenzenweger, M. F., & Korfine, L. (1992). Confirming the latent structure and base rate of schizotypy: A taxometric analysis. *Journal of Abnormal Psychology, 101,* 567–571.

Levy, D. L., Holzman, P. S., Matthysse, S., & Mendell, N. R. (1993). Eye tracking dysfunction and schizophrenia. *Schizophrenia Bulletin, 19,* 461–536.

Lilienfeld, S. O., & Marino, L. (1995). Mental disorder as a Roschian concept: A critique of Wakefield's "harmful dysfunction" analysis. *Journal of Abnormal Psychology, 104,* 411–420.

Lilienfeld, S. O., & Marino, L. (1999). Essentialism revisited: Evolutionary theory and concept of mental disorder. *Journal of Abnormal Psychology, 108,* 400–411.

Lu, R., Ho, S., Huang, H., & Lin, Y. (1988). The specificity of the dexamethasone suppression test in endogenous depressive patients. *Neuropsychopharmacology, 1,* 157–162.

Lykken, D. T. (1995). *The antisocial personalities*. Mahwah, NJ: Erlbaum.

Lynam, D. R. (1996). The early identification of chronic offenders: Who is the fledgling psychopath? *Psychological Bulletin, 120,* 209–234.

Lynam, D. R. (1998). Early identification of the fledgling psychopath: Locating the psychopathic child in the

current nomenclature. *Journal of Abnormal Psychology, 107,* 566–575.

Lynam, D. R., Caspi, A., Moffitt, T. E., Wikström, P. H., Loeber, R., & Novak, S. (2000). The interaction between impulsivity and neighborhood context on offending: The effects of impulsivity are stronger in poorer neighborhoods. *Journal of Abnormal Psychology, 109,* 563–574.

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.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory and "nonparametric" indices. *Psychological Bulletin, 107,* 401–413.

Malt, B. (1993). Concept structure and category boundaries. In G. V. Nakamura, D. L. Medin, & R. Taraban (Eds.), *Categorization by humans and machines. The psychology of learning and motivation: Advances in research and theory* (Vol. 29, pp. 363–390). San Diego, CA: Academic Press.

Mayr, E. (1942). *Systematics and the Origin of Species*. New York: Columbia University Press.

McBurnett, K., Pfiffner, L. J., & Frick, P. J. (2001). Symptom properties as a function of ADHD type: An argument for continued study of sluggish cognitive tempo. *Journal of Abnormal Child Psychology, 29,* 207–213.

Meehl, P. E. (1962). Schizotaxia, schizotypy, schizophrenia. *American Psychologist, 17,* 827–838.

Meehl, P. E. (1973). MAXCOV-HITMAX: A taxonomic search method for loose genetic syndromes. In *Psychodiagnosis: Selected papers* (pp. 200–224). Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1975). Hedonic capacity: Some conjectures. *Bulletin of the Menninger Clinic, 39,* 295–307.

Meehl, P. E. (1979). A funny thing happened to us on the way to the latent entities. *Journal of Personality Assessment, 43,* 563–581.

Meehl, P. E. (1989). Schizotaxia revisited. *Archives of General Psychiatry, 46,* 935–944.

Meehl, P. E. (1990). Toward an integrated theory of schizotaxia, schizotypy, and schizophrenia. *Journal of Personality Disorders, 4,* 1–99.

Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality, 60,* 117–174.

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

Meehl, P. E. (1996). MOXCOV pseudotaxonicity. *American Psychologist, 51,* 1184–1186.

Meehl, P. E. (1999). Clarifications about taxometric method. *Applied and Preventive Psychology, 8,* 165–174.

Meehl, P. E. (2001). Comorbidity and taxometrics. *Clinical Psychology Science and Practice, 8,* 507–519.

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 analyses II. Detecting taxonicity using covariance of two quantitative indicators in successive intervals of a third indicator. *Psychological Reports, 78,* 1091–1227.

Meller, W. H., & Borchardt, C. M. (1996). Comorbidity of major depression and conduct disorder. *Journal of Affective Disorders, 39,* 123–126.

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.

Miller, J. N., & Ozonoff, S. (2000). The external validity of Asperger disorder: Lack of evidence from the domain of neuropsychology. *Journal of Abnormal Psychology, 109,* 227–238.

Murphy, E. A. (1964). One cause? Many causes? The argument from the bimodal distribution. *Journal of Chronic Disease, 17,* 301–324.

Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods, 4,* 139–157.

Nagin, D. S., & Tremblay, R. E. (2001). Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods, 6,* 18–34.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw–Hill.

Osterling, J. A., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorders versus mental retardation. *Development and Psychopathology, 14,* 239–251.

Ozonoff, S., Rogers, S. J., & Pennington, B. F. (1991). Asperger's syndrome: Evidence of an empirical distinction from high-functioning autism. *Journal of Child Psychology and Psychiatry, 32,* 1107–1122.

Panak, W. F., & Garber, J. (1992). Role of aggression, rejection, and attributions in the prediction of depression in children. *Development and Psychopathology, 4,* 145–165.

Patterson, G. R., DeGarmo, D. S., & Knutson, N. (2000). Hyperactive and antisocial behaviors: Comorbid or two points in the same process? *Development and Psychopathology, 12,* 91–106.

Raine, A., Venables, P. H., Dalais, C., Mellingen, K., Reynolds, C., & Mednick, S. A. (2001). Early educational and health enrichment at age 3–5 years is associated with increased autonomic and central nervous system arousal and orienting at age 11 years: Evidence from the Mauritius Child Health Project. *Psychophysiology, 38,* 254–266.

Rajendar, G. K. (1996). The influence of positive and negative wording and issue involvement on responses to Likert scales in marketing research. *Journal of the Market Research Society, 38,* 235–246.

Regier, D. A., Narrow, W. E., Rae, D.S., Manderscheid, R. W., Locke, B. Z., & Goodwin, F. K. (1993). The de facto mental and addictive disorders service system. Epidemiologic Catchment Area prospective 1-year prevalence rates of disorders and services. *Archives of General Psychiatry, 50,* 85–94.

Richters, J. E. (1997). The Hubble hypothesis and the developmentalist's dilemma. *Development and Psychopathology, 9,* 193–229.

Richters, J. E., & Cicchetti, D. (1993). Mark Twain meets *DSM-III-R*: Conduct disorder, development, and the concept of harmful dysfunction. *Development and Psychopathology, 5,* 5–29.

Richters, J. E., & Hinshaw, S. P. (1999). The abduction of disorder in psychiatry. *Journal of Abnormal Psychology, 108,* 438–445.

Robins, E., & Guze, S. B. (1970). Establishment of diagnostic validity in psychiatric illness: Its application to

schizophrenia. *American Journal of Psychiatry, 126,* 983–987.

Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 92,* 726–748.

Rosch, E., & Lloyd, B. (1978). *Cognition and categorization*. New York: Wiley.

Ruma, P. R., Burke, R. V., & Thompson, R. W. (1996). Group parent training: Is it effective for children of all ages? *Behavior Therapy, 27,* 159–169.

Rutter, M., Dunn, J., Plomin, R., Simonoff, E., Pickles, A., Maughan, B., Ormel, J., Meyer, J., & Eaves, L. (1997). Integrating nature and nurture: Implications of person–environment correlations and interactions for developmental psychopathology. *Development and Psychopathology, 9,* 335–364.

Rutter, M., & Sroufe, L. A. (2000). Developmental psychopathology: Concepts and challenges. *Development and Psychopathology, 12,* 265–296.

Saal, F. E., Downey, R. G., & Lahey, M. A. (1980). Rating the ratings: Assessing the psychometric quality of rating data. *Psychological Bulletin, 88,* 413–428.

Semin, G. R., & Rosch, E. (1981). Activation of bipolar prototypes in attribute inferences. *Journal of Experimental Social Psychology, 17,* 472–484.

Shekim, W. O., Dekirmenjian, H., Chapel, J. L., & Davis, J. M. (1982). Effects of d-amphetamine on urinary metabolites of dopamine and norepinephrine in hyperactive boys. *American Journal of Psychiatry, 139,* 485–488.

Shekim, W. O., Sinclair, E., Glaser, R., Horwitz, E., Javaid, J., & Bylund, D. B. (1987). Norepinephrine and dopamine metabolites and educational variables in boys with attention deficit disorder and hyperactivity. *Journal of Child Neurology, 2,* 50–56.

Simon, D., Pham, L. B., Le, Q. A., & Holyoak, K. J. (2001). The emergence of coherence over the course of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1250–1260.

Skilling, T. A., Quinsey, V. L., & Craig, W. M. (2001). Evidence of a taxon underlying serious antisocial behavior in boys. *Criminal Justice and Behavior, 28,* 450–470.

Smalley, S. L., Bailey, J. N., Palmer, C. G., Cantwell, D. P., McGough, J. J., Del-Homme, M. A., Asarnow, J. R., Woodward, J. A., Ramsey, C., & Nelson, S. F. (1998). Evidence that the dopamine D4 receptor is a susceptibility gene in attention deficit hyperactivity disorder. *Molecular Psychiatry, 3,* 427–430.

Smith, E. E. (1995). Concepts and categorization. In E. Smith & D. Osherson (Eds.), *Thinking: An invitation to cognitive science*. (2nd ed., Vol. 3, pp. 3–33). Cambridge, MA: MIT Press.

Sonuga–Barke, E. J. S. (1998). Categorical models of childhood disorder: A conceptual and empirical analysis. *Journal of Child Psychology and Psychiatry, 39,* 115–133.

Sroufe, L. A. (1997). Psychopathology as an outcome of development. *Development and Psychopathology, 9*, 251–268.

Sroufe, L. A., & Rutter, M. (1984). The domain of developmental psychopathology. *Child Development, 55,* 17–29.

Sroufe, L. A., & Waters, E. (1977). Attachment as an organizational construct. *Child Development, 48,* 1184–1199.

Statistica [Computer software]. (1998). Tulsa, OK: Statsoft.

Stevens, S. S. (1951). *Handbook of experimental psychology*. New York: Wiley.

Strube, M. J. (1989). Evidence for the type in Type A behavior: A taxometric analysis. *Journal of Personality and Social Psychology, 56,* 972–987.

Szasz, T. (2000). Second commentary on "Aristotle's function argument." *Philosophy, Psychiatry, and Psychology, 7*, 3–16.

Tyrka, A., Haslam, N., & Cannon, T. D. (1995). Detection of a longitudinally-stable taxon of individuals at risk for schizophrenia spectrum disorders. In A. Raine, T. Lencz, & S. A. Mednick (Eds.) *Schizotypal personality disorder* (pp. 168–191). New York: Cambridge University Press.

Tyrka, A. R., Cannon, T. D., Haslam, N., Mednick, S. A., Schulsinger, F., Schulsinger, H., & Parnas, J. (1995). The latent structure of schizotypy: I. Premorbid indicators of a taxon in individuals at risk for schizophrenia-spectrum disorders. *Journal of Abnormal Psychology, 104,* 173–183.

Volkmar, F. R., Klin, A., & Pauls, D. (1998). Nosological and genetic aspects of Asperger syndrome. *Journal of Autism and Developmental Disorders, 28,* 457–463.

Volkmar, F. R., Klin, A., Schultz, R. T., Rubin, E., & Bronen, R. (2000). Asperger's disorder. *American Journal of Psychiatry, 157,* 262–267.

Volkmar, F. R., Klin, A., Siegel, B., Szatmari, P., Lord, C., Campbell, M., Freeman, B. J., Cicchetti, D. V., Rutter, M., Kline, W., Buitelaar, J., Hattab, Y., Fombonne, E., Fuentes, J., Werry, J., Stone, W., Kerbeshian, J., Hoshino, Y., Bregman, J., Loveland, K., Szymanski, L., & Towbin, K. (1994). Field trial for autistic disorder in DSM-IV. *American Journal of Psychiatry, 151,* 1361–1367.

Wakefield, J. C. (1992). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist, 47,* 373–388.

Wakefield, J. C. (1993). Limits of operationalism: A critique of Spitzer and Endicott's (1978) proposed operation criteria for mental disorder. *Journal of Abnormal Psychology, 102,* 160–172.

Wakefield, J. C. (1997). When is development disordered? Developmental psychopathology and the harmful dysfunction analysis of mental disorder. *Development and Psychopathology, 9,* 269–290.

Wakefield, J. C. (1999). Evolutionary versus prototype analysis of the concept of disorder. *Journal of Abnormal Psychology, 108,* 374–399.

Wakefield, J. C., Pottick, K. J., & Kirk, S. A. (2002). Should the DSM-IV diagnostic criteria for conduct disorder consider social context? *American Journal of Psychiatry, 159,* 380–386.

Waldman, I. D., & Lilienfeld, S. O. (2001). Applications of taxometric methods to problems of comorbidity: Perspectives and challenges. *Clinical Psychology Science and Practice, 8,* 520–527.

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

Waller, N. G., Putnam, F. W., & Carlson, E. B. (1996). Types of dissociation and dissociative types: A taxometric analysis of dissociative experiences. *Psychological Methods, 1,* 300–321.

Waller, N. G., & Ross, C. A. (1997). The prevalence and the biometric structure of pathological dissociation in the general population: Taxometric and behavior ge-

netic findings. *Journal of Abnormal Psychology, 106,* 499–510.

Waters, E., & Beauchaine, T. P. (2003). Are there really patterns of attachment? Comment on Fraley and Spieker. *Developmental Psychology, 39,* 417–422.

Webster–Stratton, C., & Hammond, M. (1997). Treating children with early-onset conduct problems: A comparison of child and parent training interventions. *Journal of Consulting and Clinical Psychology, 65,* 93–109.

Webster–Stratton, C., Reid, M. J., & Hammond, M. (2002). Social skills and problem solving training for children with early-onset conduct problems: Who benefits? *Journal of Child Psychology and Psychiatry, 42,* 943–952.

Widiger, T. A. (2001). What can be learned from taxometric analyses? *Clinical Psychology Science and Practice, 8,* 528–533.

Williamson, D. A., Womble, L. G., Smeets, M. A. M., Netemeyer, R. G., Thaw, J. M., Kutlesic, V., & Gleaves, D. H. (2002). Latent structure of eating disorders symptoms: A factor analytic and taxometric investigation. *American Journal of Psychiatry, 159,* 412–418.

Woodward, S. A., Lenzenweger, M. F., Kagan, J., Snidman, N., & Arcus, D. (2000). Taxonic structure of infant reactivity: Evidence from a taxometric perspective. *Psychological Science, 11,* 296–301.