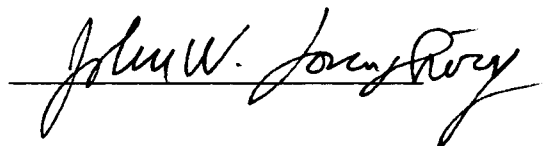


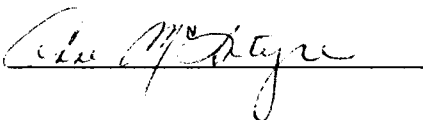
To the Graduate Council:

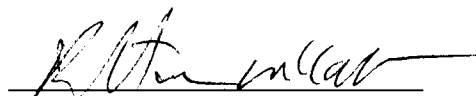
I am submitting herewith a dissertation written by Katherine Hulse-Trotter entitled "Evaluating the Quality of a Behavioral Assessment Instrument: A Treatment Validity Study." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Psychology.


Richard A. Saudargas, Major Professor

We have read this dissertation
and recommend its acceptance:







Accepted for Council:


Associate Vice Chancellor and Dean
of the Graduate School

Evaluating the Quality of a Behavioral Assessment Instrument:

A Treatment Validity Study

A Dissertation Presented for the

Doctor of Philosophy Degree

The University of Tennessee, Knoxville

Katherine Hulse-Trotter

December 1996

ACKNOWLEDGMENTS

Sincere thanks and appreciation go to my committee chair, Dr. Richard A. Saudargas, for his guidance, support, and friendship throughout this project. I would also like to thank the other members of my doctoral committee, Dr. Anne McIntyre, Dr. John Lounsbury, and Dr. Steve McCallum.

I am also thankful to Millie Douglas for her help with data collection and management, her organizational skills, and her dry humor which helped me keep things in perspective. Additional thanks goes to the many observers who gave so much of their time and energy to this project, in particular, Michelle Love.

Special thanks goes also to the teachers and school psychologists who made this research possible, especially Dr. Lucy Haile-Griffey who always made things easier than they could have been.

ABSTRACT

The purpose of this study was to evaluate the treatment validity of the State-Event Classroom Observation System. Treatment validity is one application of the overall notion of validity which is concerned with determining the appropriateness of interpretations made from test scores. No test is valid for all purposes. The degree of validity depends on a particular use; therefore, one must always ask, validity for what? In treatment validity, the question is one of the relationship between assessment and treatment outcome. Treatment validity is established through evidence that an assessment instrument contributes to the effectiveness of treatment. For example, if an instrument identifies appropriate target behaviors for change, these behaviors can serve as pre-intervention measures against which behavior change can be compared. Similarly, if an instrument is sensitive to changes in target behaviors, it can be used to help determine when modifications in recommended intervention procedures are warranted.

Support for the treatment validity of SECOS was provided through evidence that the code is sensitive to both known behavioral differences distinguishing ADHD children from their peers, and the effects of stimulant medications on these behaviors as documented in numerous studies. Eighteen

pairs of hyperactive and normally active children were observed in their regular classrooms both pre- (i.e., at referral) and post-medication of the hyperactive children. At referral, SECOS data indicated that ADHD children were more frequently off-task and exhibited more total disruptive behaviors and more variable behavior than comparison children. Following medication, previously documented enhancements in the behavioral functioning of ADHD children (i.e., increased attention, decreased disruptive behavior and behavior variability) were evident on SECOS. The findings in this study were further substantiated by analogous changes in children's daily academic work. Results are discussed in terms of their implications for identification and treatment of hyperactive children.

TABLE OF CONTENTS

Chapter		Page
1.	INTRODUCTION.....	1
	General.....	1
	Validity and Observational Assessment: A Review of the Literature.....	5
	Study Context.....	11
	Summary.....	15
	Hypotheses.....	15
2.	METHOD.....	17
	Subjects.....	17
	Setting.....	21
	Dependent Measures.....	22
	Experimental Design and Procedure.....	24
	Observer Training.....	24
	Interobserver Agreement.....	25
	Data Analyses.....	27
3.	RESULTS.....	31
	Primary Analyses.....	31
	Secondary Analyses.....	36
4.	DISCUSSION.....	43
	REFERENCES.....	51
	APPENDIXES.....	60
	Appendix A.....	61
	Appendix B.....	73
	VITA.....	74

LIST OF TABLES

Table	Page
1. Diagnostic criteria for Attention-Deficit Hyperactivity Disorder.....	61
2. The State-Event Classroom Observation System (SECOS).....	64
3. Interobserver agreement for ADHD and comparison observations at referral and medication.....	67
4. Means, standard deviations, medians, and ranges for SECOS behaviors at referral and medication.....	69
5. Data used in determining clinical significance.....	72

LIST OF FIGURES

Figure	Page
1. ADHD and comparison children at referral.....	32
2. ADHD and comparison children during the medication phase.....	33
3. Behavior normalized: Fourth-grade male.....	38
4. Behavior improved but not normalized: First-grade male.....	39
5. No change in on-task behavior: First-grade female.....	41

CHAPTER 1

INTRODUCTION

General

Validity of measurement concerns the extent to which a measure reflects "true" behavior, and the extent to which the data obtained provides meaningful and relevant information that can be used to address a specific set of problems and concerns (Messick, 1995; Silva, 1993; Ghiselli, Campbell, and Zedeck, 1981). Validity has become the psychometric standard of quality by which measurement instruments are judged (Silva, 1993). Validity refers to accuracy and precision in measurement. In addition, validity may be conceptualized as "the appropriateness of inferences from test scores or other forms of assessment" (APA, AERA, NCME, 1974, p. 25). Although the concept of validity is essentially unitary, that is, "different kinds of inferences from test scores require different kinds of evidence, not different kinds of validity" (Messick, 1980; p. 1014), several "types" or "classes" of validity have been proposed as guidelines for accumulating evidence in support of the overall validity of a test or measure. These classes of validity include content, criterion-related, construct, and, more recently, treatment validity. Briefly, content validity concerns the relevance and representativeness of categories on a test; criterion-related validity pertains to the relationship between a

set of scores on a measure and a chosen criterion; construct validity is the degree to which a measure or category of measurement reflects a construct; finally, a measure has treatment validity if it is sensitive to changes in behavior that result from intervention and treatment (Messick, 1995; Silva, 1993; Whitley, 1996; Hoge, 1985) .

In literature concerning the psychometrics of behavioral observation, issues pertaining to the content-, criterion-related, and construct validity of these procedures have received considerable attention, and recommendations have been made for improving the quality of data obtained from direct observation measures (i.e., for reducing sources of invalidity) (e.g., Foster & Cone, 1980; Ghiselli, et al., 1981; Hoge, 1985; Silva, 1993; Suen & Ary, 1989). Considerably less attention has been given to treatment validity. This neglect may be attributed in part to the treatment validity model which places an emphasis on *the assessment process* rather than assessment results. In this model, treatment validity has become synonymous with treatment utility, and, fundamentally, is evidenced by a positive relationship between the assessment process and recommended treatments (i.e., the assessment enhances treatment outcome) (Hayes, Nelson, & Jarrett, 1986, 1987; Kratochwill & McGovern, 1996; Nelson & Hayes, 1979). Accordingly, "there may even be times when an assessment *process* could have treatment utility

without the assessment *report* having any reliability or validity whatsoever . . .

Therefore, if this assessment process can have treatment utility, then some tests or procedures may be psychometrically invalid and still have treatment utility on that basis" (Hayes et al., 1987; p. 972). It is not surprising, then, that the treatment validity of observational data is discussed in terms of the conceptual and practical link between observational assessment and treatment rather than experimentally demonstrated results.

To illustrate, in observational assessment, assessment and treatment can be linked in three ways. First, intervention and treatment decisions are based on the direct assessment of skill deficiencies and mediating steps between obtaining and interpreting data are minimized. Second, since behavior is measured in the actual environment in which interventions or treatments typically occur, data from observational assessment are more closely related to decisions about students and more likely to translate directly into valid treatment planning. Finally, behaviors can be monitored over the course of treatment and suggestions made for modification of the treatment plan if necessary (Shapiro; 1987; Shapiro & Kratochwill, 1988). Observational assessment focuses on individual behavior in the current environment, data from observational assessment is relevant to treatment goals, and the methods are low inference, direct, and ongoing. As a

result, direct observation has been characterized as "intrinsically linked to valid interpretation of assessment data and development of systematic intervention plans," and as having a "high degree of treatment validity" (Merrell, 1994; p. 45).

The emphasis here, however, is on the assessment process in general with little attention to the instruments used in the assessment. To date, the treatment validity of direct observation measures has not been established empirically but inferred based on the results of the assessment. For example, if an observation measure detects significant positive changes in a student's on-task behavior in response to changes in teacher attention, this is considered evidence of treatment validity (Hoge, 1985; Nelson, 1988). A problem arises, however, in the evaluation of negative results. If an observation measure indicates no change in behavior following intervention and an alternative assessment method (e.g., a teacher rating scale) indicates significant improvement, one would have to assume either that no change occurred or that the observation measure failed to detect it (Foster & Cone, 1980; Hayes, et al., 1987). If treatment validity is not demonstrated in a given study, the obvious question is: Is it because the assessment itself lacked treatment validity or because the instruments used to measure treatment outcome were not sensitive to treatment effects (i.e., lacked treatment validity)? As noted by Hayes et al. (1987), "It is impossible to build a science of assessment totally on

treatment utility (validity) alone. Knowing the effect on treatment outcome itself requires assessment, and that means that treatment utility (validity) must ultimately be based on measures that are themselves not validated in this manner" (p. 972).

Thus, the treatment validity of assessment and the treatment validity of instruments used in the assessment must be established separately. Furthermore, since types of validity do not imply "sharp distinctions," but may be perceived as cumulative evidence of the overall validity of an instrument (AERA, APA, & NCME, 1985), recommended procedures for increasing an observation code's validity and strategies for gathering evidence of validity will be reviewed.

Validity and Observational Assessment: A Review of the Literature

A valid assessment not only provides accurate data but has relevance for educational decisions. Consequently, "assessments can be judged in terms of their utility in moving the child toward appropriate educational goals" and making treatment-related decisions (Heller, Holtzman, & Messick, 1982, p. 98). An observation instrument can provide information helpful in developing effective educational plans if the "right" behavior is observed in the "right" setting (Foster & Cone, 1980). An observation instrument can be evaluated in traditional psychometric terms by examining both internal (content, criterion-related, construct) and external (convergent, discriminant, treatment) aspects of a code's

validity (Messick, 1995).

Content Validity

The content validity of a classroom observation code is typically assessed through logical evaluation of the behaviors and settings chosen for observation. A code is said to have content validity if it "adequately samples the behavior domain of interest" (Furman & Drabman, 1981). In addition to targeting specific problem behaviors, a classroom observation code should encompass a wide range of positive, alternative behaviors as well as contextual variables that are potentially influencing the observed problem (Keller, 1986). Furthermore, since changes in settings have been found to correlate with changes in behavior, the code should be useful in a variety of classroom settings (e.g., different instructional or physical arrangements) (Merrell, 1994).

Criterion-Related Validity

A classroom observation code that includes behaviors that predict or are related to other important criteria has criterion-related validity (Hoge, 1985; Hoge & Luce, 1979). Variables related to academic achievement and student adjustment include academic engaged time and various contingencies surrounding academic performance. For example, teacher and peer attention, teachers' praise and disapproval, and the pattern of interactions with teachers and peers are all known

to have potentially strong effects on student behavior. Other variables related to student performance include aspects of the teaching style employed by the classroom teacher such as stimulus procedures, prompts, accuracy feedback, and error correction (Shapiro & Skinner, 1990).

Construct Validity

Construct validity has to do with the meaningfulness of behavior categories on an observation code. For example, if a code contains a behavioral category called "disruptiveness," one needs to determine the extent to which this category is actually measuring classroom disruptiveness (Hoge & Luce, 1979). Factors influencing the construct validity of classroom observation codes include the way behaviors are defined and the methods used to record the behaviors.

Behavioral definitions. The behaviors included on an observation code may be defined very broadly or very narrowly. One type of measure includes molar categories of behavior such as "off task," "inappropriate," or "disruptive." A second type of measure includes specific or molecular categories (e.g., raise hand, out of seat) of behavior. In a third type of measure, the molecular-composite, several discrete behaviors are combined to form a single response class. For example, a measure may yield a composite score called "aggression" by summing across relevant subcategories such as kicking, hitting, biting, pushing, and name

calling (Hoge, 1985). Construct validity is less of a problem for low inference (e.g., molecular) measures than for high inference (e.g., molar) measures. The more molar or global behavior categories, since they call for judgment on the part of observers (Alessi & Kaye, 1983; Hoge & Luce, 1979), are apt to lead to inconsistency in measurement and problems in detection (Johnston & Pennypacker, 1980). In addition, relationships among behaviors combined to form a single response class (i.e., molecular composite) should be based on empirical evidence of response interrelationships rather than similarities in topography or presumed functional similarity (Foster & Cone, 1980).

Recording procedures. The extent to which a direct observation measure provides a veridical representation of behavior is partially determined by the method and procedure by which data are recorded. Typically, observation sessions are divided into timed intervals, and the presence or absence of behaviors is recorded within each interval. Common methods for recording behaviors include partial-interval, whole-interval, and momentary time sampling. Each of these recording procedures is differentially sensitive to natural behavior parameters (Saudargas & Lentz, 1986). When compared to natural (i.e., continuous) measures of behavior, momentary time sampling provides the most accurate estimate of actual behavior durations (Murphy & Harrop, 1994; Saudargas & Zanolli, 1990; Klesges,

Woolfrey, & Vollmer, 1985). Partial-interval time sampling consistently overestimates behavior durations, while whole-interval time sampling consistently underestimates durations (Powell, 1984). For all of these recording procedures, longer intervals result in less accurate data (Merrell, 1994), and none of these procedures is sensitive to behavioral frequencies. Where frequency is of interest (e.g., to record discrete body movements or vocalizations which can be approximated as points in time), the most accurate procedure is simply to count behaviors as they occur in the appropriate interval (Saudargas & Lentz, 1986). The construct validity of a code depends on a sampling procedure that is sensitive to the most relevant dimensional qualities (i.e., duration, frequency) of the behaviors targeted for observation, that yields accurate data, and does not introduce unknown measurement error (Foster & Cone, 1980; Johnston & Pennypacker, 1980).

Convergent and Discriminant Validity

Convergent and discriminant validity (both aspects of construct validity) provide external substantiation that the construct(s) represented are consistent (convergent) or not consistent (discriminant) with measures of the same or different constructs, respectively. Traditionally, convergent validity is based on evidence of correspondence between methods that measure the same behavior

while discriminant validity is based on evidence of disparate results between measures designed to measure different behaviors (Messick, 1995). In the context of observational assessment, data from behavioral observations and rating scales designed to measure the same construct (e.g., hyperactivity) have been compared in an effort to demonstrate convergent validity. Discriminant validity has been based on evidence of behavioral differences between groups of children.

Evidence for the convergent validity of rating scale and observation data has not been well-established (Skiba, 1989; Foster & Cone, 1980). Although both are widely recommended practices for assessing behavior problems, it has been suggested that the two measures tap different dimensions of a child's behavior with ratings reflecting enduring properties and observation reflecting behavioral events influenced by situational and contextual variables (Skiba, 1989). Moreover, ratings, unlike observation, are based on both objective behavior and the rater's past experience with the ratee (McMahon, 1984). Many direct observation instruments have demonstrated discriminant validity. For example, a school-based observation code may discriminate between handicapped and nonhandicapped children or between children of different socioeconomic classes.

Treatment Validity

An assessment device or instrument that is sensitive to changes in behavior

that result from intervention and treatment has treatment validity (Hoge, 1985; Cone, 1982). Empirical evidence in support of the treatment validity of direct observation measures is notably absent.

Study Context

The treatment validity of the State-Event Classroom Observation System (SECOS; Saudargas, 1992) was assessed by ascertaining the degree to which SECOS is sensitive to changes in behavior following intervention and treatment of children referred for Attention-deficit Hyperactivity Disorder (ADHD). This disorder was chosen because (1) ADHD is one of the most common Axis I childhood disorders, and (2) a powerful treatment (i.e., stimulant medications) is available to treat the disorder (APA, 1994; Greenhill, 1992). As a result, not only can the treatment validity of SECOS be established without the additional confound of some unknown treatment effect, but predictions can be made as to the type and direction of behavioral change that should be evident on SECOS following intervention and treatment of the ADHD child.

Symptoms of ADHD

ADHD is characterized by inattention, restlessness, impulsivity, and hyperactivity. DSM-IV divides ADHD criteria into three subtypes, inattention, hyperactivity-impulsivity, and combined type. Symptoms of inattention include

such descriptors as "easily distracted," "often does not seem to listen," and "makes careless mistakes." Symptoms of hyperactivity-impulsivity include fidgeting, excessive running and climbing, blurting out answers, and difficulty waiting turn (see Table 1* for DSM-IV criteria) (APA, 1994). To qualify for the ADHD diagnosis, children must exhibit (for a period of at least six months) six or more symptoms from either the inattention or hyperactivity-impulsivity category, and the symptoms must be present in at least two settings. Children who meet the criteria for both categories are coded as being combined type (Power & DuPaul, 1996).

Manifestations of these primary symptoms occur more prominently in situations with a high degree of structure or in those requiring sustained attention to task (e.g., the classroom). In less structured situations (e.g., free play) the behavior of ADHD children may be indistinguishable from that of their peers. In addition to these primary symptoms, ADHD children often have difficulty achieving on grade level and experience serious difficulties in social relationships as their behavior tends to elicit negative interactions from both teachers and peers (Barkley, 1990; Barkley, DuPaul, & McMurray, 1990; Carlson & Rapport, 1989; Frick & Lahey, 1991; Sheridan, Dee, Morgan, McCormick, & Walker, 1996).

Studies employing direct observation in the classroom have focused on a

*All tables are presented in Appendix A.

few negative/disruptive behaviors and attention-to-task to discriminate between normal and hyperactive children. Behaviors subsumed under the heading of "disruptive" include, but are not limited to, interference (i.e., calling out, interrupting others during work periods), solicitation (i.e., seeking teacher attention), noncompliance, out of seat, self stimulation (i.e., fidgeting, tilting chair), destruction of property, name calling, teasing, and swearing. In general, results of observational studies consistently demonstrate that ADHD children are more frequently "off -task" and exhibit more total disruptive behaviors than controls, and these differences are more pronounced in familiar/low structure settings (e.g., independent seatwork) than in novel (e.g., film) or high structure settings (e.g., reading group) (Atkins, Pelham, & Licht, 1985; Barkley, Fischer, Newby, & Breen, 1988; Jacob, O'Leary, & Rosenblad, 1978; McGee, Williams, & Silva, 1984; Zentall, 1980). In addition, ADHD children display significantly more variable classroom behavior than their non-hyperactive peers (Abikoff, Gittelman-Klein, & Klein, 1977; Atkins et al., 1985; Schaughency & Rothlind, 1991).

Psychostimulant Intervention

Numerous studies have documented enhancements in the academic, behavioral, and social functioning of the majority of ADHD children treated with psychostimulant medications. These drugs have been shown to have positive

effects on the ability of such children to sustain attention to task in the classroom; and, using classroom-based measures of academic performance (i.e., children's regularly assigned academic work), medication-induced improvements in academic productivity and accuracy have been found for a large number of ADHD children (Carlson & Bunner, 1993; Carlson, Pelham, Milich, & Dixon, 1992; DuPaul & Rapport, 1993; Fischer & Newby, 1991; Pelham, Bender, Caddell, Booth, & Moorer, 1985; Rapport, Jones, DuPaul, Kelly, Gardner, Tucker, & Shea, 1987; Rapport, Stoner, DuPaul, Birmingham, & Tucker, 1984). Reductions in the day-to-day variability of children's attentive behavior have also been noted (DuPaul & Rapport, 1993). Further, these medications significantly reduce classroom disruptive behavior, task-irrelevant movement and activities, and noncompliance with authority figure commands (Carlson et al., 1992; DuPaul, Barkley, & McMurray, 1991; Fischer & Newby, 1991; Pelham, 1993; Pelham et al., 1985). Research on the effects of medication on the social behavior of ADHD children has produced mixed results. Some studies indicate that while aggressive and noncompliant behavior may decrease, there may not be a concomitant increase in those prosocial behaviors either initiated or received by the ADHD child (Landau & Moore, 1991). Other findings suggest, however, that medication may not only inhibit negative social behaviors (i.e., pushing another child, quitting a game in

anger), but increase positive social behaviors (i.e., waiting turn, playing in a cooperative manner) (Pelham, Walker, Sturges, & Hoza, 1989). Finally, the effects of stimulant medication on the behavior of ADHD children have been shown to be comparable in males and females (Pelham, et al., 1989), and in those children diagnosed ADHD with and without the hyperactivity component (Pelham, 1993).

Summary

Treatment validity concerns the utility of a particular assessment in developing intervention or treatment plans. Although the treatment validity of observational assessment has been widely acclaimed, there is no published empirical evidence of the treatment validity of specific instruments used in the assessment. Moreover, while utility has been characterized as an issue separate from other psychometric standards, without these standards an estimation of treatment utility is impossible. In the present study, the treatment validity of SECOS was assessed in terms of the sensitivity of the instrument to the medication status of ADHD children.

Hypotheses

Hypothesis I

Prior to intervention and treatment, it was expected that the behaviors of

hyperactive and normal children as indicated by SECOS would show the following pattern: Hyperactive children were expected to (1a) engage in less on-task behavior (schoolwork) than normals; (1b) exhibit more disruptive/inappropriate behavior (e.g., look, other activity, out of seat) than normals; (1c) solicit and receive more teacher attention than normals; and (2) exhibit significantly more variable classroom behavior than normals. Although patterns of student interactions were examined (e.g., approach child, other child approach), there was no a priori basis for making specific predictions in regard to these behaviors.

Hypothesis II

Medication was anticipated to normalize the behavior of ADHD children, and the following behavior patterns were expected on SECOS: The behavior of hyperactive children was expected to (1a) change in a positive direction (e.g., increased on-task behavior, decreased disruptive/inappropriate behavior); (1b) be more consistent (i.e., less variable); and (2) statistically equivalent behavior patterns on SECOS were expected for hyperactive and normal children.

CHAPTER 2

METHOD

Subjects

Children (n=18) who were diagnosed with Attention-deficit Hyperactivity Disorder (ADHD) and for whom medical intervention (i.e., psychostimulant medication) was recommended, and their classroom comparisons (n=18) were included in the study. The hyperactive children were compared with 18 normally active children selected from the same classrooms on the basis of sex and teacher recommendation. All children's parents provided informed consent to participate in the study (see Appendix B). Potential participants in the study included 51 elementary school-age children referred by their teachers for hyperactivity. Referred children excluded from the study included (1) children diagnosed with ADHD but for whom medication was not prescribed either because parents opposed medical treatment (n=14) or physicians did not recommend it (n=4); (2) children who did not meet the certification criteria for ADHD (n=8); (3) children for whom there was no follow-up on the teacher's referral by the school psychologist (n=4); and (4) children who left the school system prior to completion of the study (n=3).

The final sample included eighteen pairs of children (12 male; 6 female) in

grades one through five (first grade, $n = 5$; second grade, $n = 4$; third grade, $n = 3$; fourth grade, $n = 4$; fifth grade, $n = 2$). Children in the hyperactive sample were diagnosed with ADHD on the basis of parent and teacher questionnaires, parental reports of a long history of hyperactivity, and evaluation by a licensed psychologist or pediatrician. Children in the comparison sample were reported by the teacher to exhibit average behavior in the classroom and had never been referred for assessment.

Parent Ratings

Parents of ADHD and comparison children completed the Child Behavior Checklist (CBCL; Achenbach, 1991). The CBCL consists of seven competence items and 113 behavior problem items. Competence items provide information about the child's participation in sports, activities, groups, chores, peer relationships, and academics. The behavior problem items cluster into eight subscales: Withdrawn, Somatic Complaints, Anxious/Depressed, Social Problems, Thought Problems, Attention Problems, Delinquent Behavior, and Aggressive Behavior. These subscale scores further combine to yield Externalizing and Internalizing scores; and a Total Problem score is obtained by combining all behavior problem items with the exception of two items. Problem Subscale T scores of 67 to 70 and higher and Total Problem T scores of 60 to 63 and higher

may be considered clinically elevated.

Scores for the control sample on the CBCL were consistent with the normative data available for this scale. The mean Problem Subscale T score was 51 (range 42 to 55), and the mean Total Problem T score for comparison children was 46 (range 34 to 61). For ADHD children, clinically elevated scores were obtained for the Total Problem T score with a mean of 66 (range 55 to 77) reported by parents. The mean Problem Subscale T score was 59 (range 55 to 63). Subscales rated as most problematic by parents included the Externalizing (T = 63), Delinquent (T = 60), and Attention subscales (T = 63).

Teacher Ratings

Teachers completed the Child Behavior Checklist - Teacher Report Form (CBCL-TRF), the School Situations Questionnaire - Revised (SSQ - R) and the Child Attention Problems rating scale (CAP) for ADHD children (Barkley, 1990). The CBCL-TRF includes an Academic / Adaptive functioning section; and, like the CBCL, contains 113 behavior problem items that cluster into eight subscales. Scoring and interpretation of the CBCL and the CBCL-TRF are essentially the same (Kronenberger & Meyer, 1996).

For ADHD children, clinically elevated scores on the CBCL-TRF were obtained for the Total Problem score. The mean Total Problem score reported by

teachers was 70 (range 60 to 87). The mean Problem Subscale T score was 64 (range 56 to 70). Subscales rated as most problematic by teachers included the Externalizing (T = 69), Aggressive (T = 68), Delinquent (T = 65), Attention (T = 70), Social (T = 65), and Thought (T = 67) subscales.

The SSQ-R is an eight-item scale consisting of school situations (e.g., individual desk work, class discussions) that may be problematic for ADHD children. Two scores can be obtained from the scale: the number of problem settings and the mean severity (Barkley, 1990). Normative data are available for this scale for boys and girls between the ages of six and 12 years. For females in the normative population, the mean number of problem settings ranges from 2.00 to 2.84, and mean severity ranges from 2.99 to 3.81. For females in this study referred for ADHD, the mean number of problem settings reported by teachers was 6.2 (range 6 to 7), and the mean severity was 6.4 (range 6 to 8). For males in the normative population, the mean number of problem settings ranges from 2.12 to 3.86, and mean severity ranges from 3.14 to 4.82. For males in this study referred for ADHD, the mean number of problem settings reported by teachers was 7 (range 5 to 8), and the mean severity was 6.5 (range 3 to 8).

The CAP is a 12-item scale consisting of behaviors that may be indicative of ADHD (e.g., talks too much, easily distracted). Three scores can be obtained

from the scale: a total score (upper limit of normal range is 15 for boys, 11 for girls), an overactivity score (upper limit of normal range is 6 for boys, 5 for girls), and an inattention score (upper limit of normal range is 9 for boys, 7 for girls) (Barkley, 1990). On the CAP, scores for ADHD children exceeded the upper limit of the normal range for the total score and the overactivity factor assessed by this scale. For females in the ADHD sample, the mean total score was 17.4 (range 14 to 23), and the mean overactivity score was 10.4 (range 8 to 13). For males, the mean total score was 16.8 (range 5 to 24), and the mean overactivity score was 12.5 (range 8 to 19). Mean inattention scores obtained for males and females were 8.7 (range 7 to 12) and 6.6 (range 4 to 10), respectively.

Setting

The study took place in three elementary schools in a large East Tennessee school district. Both phases (referral and medication) of the study were conducted in the classroom during independent seatwork. In the independent seatwork setting, children work independently at their desks (e.g., completing a workbook assignment) while the teacher works at his or her desk or conducts small reading groups.

Dependent Measures

Observations

Data were collected using the State-Event Classroom Observation System (SECOS; Saudargas, 1992) by trained observers who were blind to the purpose of the study and unaware of the diagnostic or medication status of the participants. SECOS is a classroom observation code designed to be used in the remediation of academic and/or behavior problems experienced by elementary school-aged children. Students in regular and special education classrooms may be observed in a variety of instructional arrangements including independent seatwork, large group (i.e., teacher lecture), and small group (i.e., reading group) settings. The 17 behavior categories on SECOS are low inference and descriptive and encompass a wide range of student behaviors (e.g., schoolwork, other activity), and antecedent (e.g., call out) and consequent (e.g., teacher disapproval) events known to influence behavior and academic performance.

Of the 17 behaviors coded on SECOS, six are child state behaviors, five are child event behaviors, and six are teacher event behaviors (see Table 2). State and event behaviors are not coded simultaneously. Rather, a momentary time sampling procedure is used to code state behaviors every 15 seconds, and event behaviors are counted as they occur within each 15-second interval. State

behaviors are summarized into percent of intervals to provide information on how much time a child engages in each behavior. Event behaviors are summarized into rate per minute to provide information on the frequency of behaviors. Data is coded sequentially so that recommendations for change can be made in the context of ongoing behavioral sequences.

Very acceptable levels of interobserver reliability for state (range .88 to .97) and event (range .71 to 1.00) behaviors have been reported. In addition, significant Kappas ($p < .05$) have been obtained for all state behaviors on SECOS (Saudargas & Lentz, 1986). In regard to validity, the code discriminates between nonhandicapped and learning disabled (Fellers & Saudargas, 1987; Slate & Saudargas, 1986a), behaviorally disordered (Slate & Saudargas, 1986b), and gifted children (Fellers & Saudargas, 1988). The code also has ecological validity since it is sensitive to a variety of instructional arrangements and curricula (Wadlington, 1988). Normative data from a sample of 486 nondisabled elementary students are also available for SECOS (Haile-Griffey, 1992).

Work samples

Teachers provided representative academic work samples for hyperactive and comparison children. These samples were collected both before and after intervention and treatment of the ADHD children. All work samples were scored

for completion (ratio of attempted items to assigned items) and accuracy (ratio of correct items to attempted items).

Experimental Design and Procedure

A pretest-posttest design was employed in this study. Before the initial classroom observation, teachers provided a seating chart and daily schedule of activities to the observer. In addition, parents and teachers were instructed not to divulge who was being observed, but only that someone was coming who was interested in how a classroom worked. All children were observed in their regular classrooms during independent seatwork on six occasions. Three of the observations took place prior to intervention and treatment of the ADHD children (at referral), and three subsequent to intervention and treatment of the ADHD children (medication). In addition, academic work samples were obtained for both the referral and medication conditions.

Observer Training

Observers were undergraduate students at the University of Tennessee, Knoxville, who received academic credit for their participation. Observers were asked to review the SECOS manual and learn the behavioral definitions and coding procedures. Following discussion and tests of their understanding of this material, observers met for two one-hour practice sessions for a period of six weeks.

Practice sessions were conducted using simulated classroom videotapes, and all observers had to reach a minimum of 80% agreement with criterion-scored videotapes before collecting data in the schools. Observers were retrained periodically throughout data collection in the schools, and feedback was provided about definitional problems and coding errors.

Interobserver Agreement

Agreement scores were calculated for each SECOS variable by two methods. First, the frequency of agreement, a percentage based on the total ratio of the total frequency of a behavior as measured by two concurrent observers, was calculated. Second, a more conservative method of establishing reliability, the exact agreement formula ($\text{Agreements} / (\text{Agreements} + \text{Disagreements}) - \text{Agreements}$), was utilized. The exact agreement formula accounts not only for the number of behavior occurrences scored, but the number of times the same behavior is scored within the same time interval by two concurrent observers. Agreement scores were available for 20 percent of the total observations for each group (ADHD and comparison) at both the referral and medication conditions. In addition, these scores were computed separately for ADHD and comparison children as it has been suggested that it may be more difficult to obtain reliable data for the ADHD population.

Reliability of ADHD Observations

The overall interobserver agreement for the state behaviors using the frequency of agreement method was 92.8% and 86.8%, at the referral and medication conditions, respectively. For event behaviors, the frequency of agreement method yielded agreement coefficients of 81.8% at referral and 81.8% at medication. The overall interobserver agreement for state behaviors using the exact agreement method was 84.4% at referral and 80.6% at medication. For event behaviors, the exact agreement method yielded agreement coefficients of 70.5% at referral and 74.8% at medication. The reliability of event behaviors was calculated using scores from only seven of 11 possible categories as four of the behavior categories were not scored by either observer (i.e., opportunity to respond, direction, approval, disapproval). See Table 3 for reliability coefficients obtained for each SECOS behavior.

Reliability of Comparison Observations

The overall interobserver agreement for the state behaviors using the frequency of agreement method was 89.8% at referral and 89.2% at medication. For event behaviors, the frequency of agreement method yielded agreement coefficients of 92.5% at referral and 87.9% at medication. The overall interobserver agreement for state behaviors using the exact agreement method was

81.9% and 80.8% at referral and medication, respectively. For event behaviors, the exact agreement method yielded agreement coefficients of 75.9% at referral and 81.9% at medication. The reliability of event behaviors was calculated using scores from only three of 11 possible categories as eight of the behavior categories were either not scored by either observer or scored for only one observation session (i.e., raise hand, call out, teacher approach school work, teacher approach other, opportunity to respond, direction, approval, disapproval). See Table 3 for reliability coefficients obtained for each SECOS behavior.

Data Analyses

Raw scores obtained for SECOS state and event behaviors were converted to T scores obtained from a norm group of 486 elementary school students in grades one through five. The T scores are standard scores with a mean of 50 and a standard deviation of 10 (Haile-Griffey, 1992). Median T scores and the within-subject range of these T scores (i.e., the range over the three sessions the child was observed at referral and during the medication condition) were used to represent the magnitude and variability of subject behavior in all subsequent data analyses.

In order to simplify data analyses (i.e., to reduce the variable to subject ratio), conceptually related dependent variables on SECOS were combined for all group-level analyses and cronbach's alpha were used to estimate the internal

consistency of individual items. Alpha coefficients should be at least .70, but lower coefficients are expected and acceptable in a highly homogeneous sample (Bernardi, 1994). By this method, 15 of the 17 SECOS behaviors were combined to reflect four dependent variables: off-task / negative teacher attention, social interaction with teacher / positive, social interaction with child, and out of seat. School work and looking around were maintained as separate variables.

Primary Analyses

Behavior patterns of hyperactive and comparison children: referral and medication conditions: Data for each subject (n=36) were summarized by descriptive statistics which provided means, standard deviations, medians, and ranges for SECOS behaviors at referral and during the medication phase. Median and range scores were analyzed via repeated measures multivariate analyses of variance (MANOVA) to test for (1) group differences at referral and following medication, (2) differences that occurred over time (i.e., change from referral to medication condition), and (3) interaction effects. Multivariate analyses were employed to decrease Type 1 error, and because it is the most appropriate procedure for analyzing multiple dependent variables (Bordens & Abbot, 1991). Further, since there are only two levels of the independent variables group (ADHD and comparison) and condition (referral and medication) all multivariate analyses of the

F statistic were made according to the Hotellings criterion.

Secondary Analyses

Individual change. In order to examine change in ADHD children's behavior on an individual basis (i.e., the degree to which post-treatment behavior resembles that of the normative population), difference scores were calculated for SECOS behaviors. Consistent with the procedures described by Jacobson and Truax (1991), the reliable change index and the *c* statistic were used to classify each child's score as exhibiting "no change," "improved but not normalized," or "normalized."

Correspondence between academic measures and SECOS data. Improvements in academic productivity (i.e., percentage of work completed) and efficiency (i.e., percentage of work correct) following medication were expected to be consistent with SECOS data. That is, significant differences were expected at referral between the two groups while the performance of ADHD children was expected to be indistinguishable from comparison children following medication. Additionally, scores for ADHD children on these two measures were expected to increase during the medication phase (analogous to changes in the SECOS dependent measure school work). The statistical equivalence of academic scores for ADHD and comparison children was analyzed at referral and following medication

through a series of independent t tests. To measure change, t tests for paired differences were computed for these scores.

CHAPTER 3

RESULTS

Primary Analyses

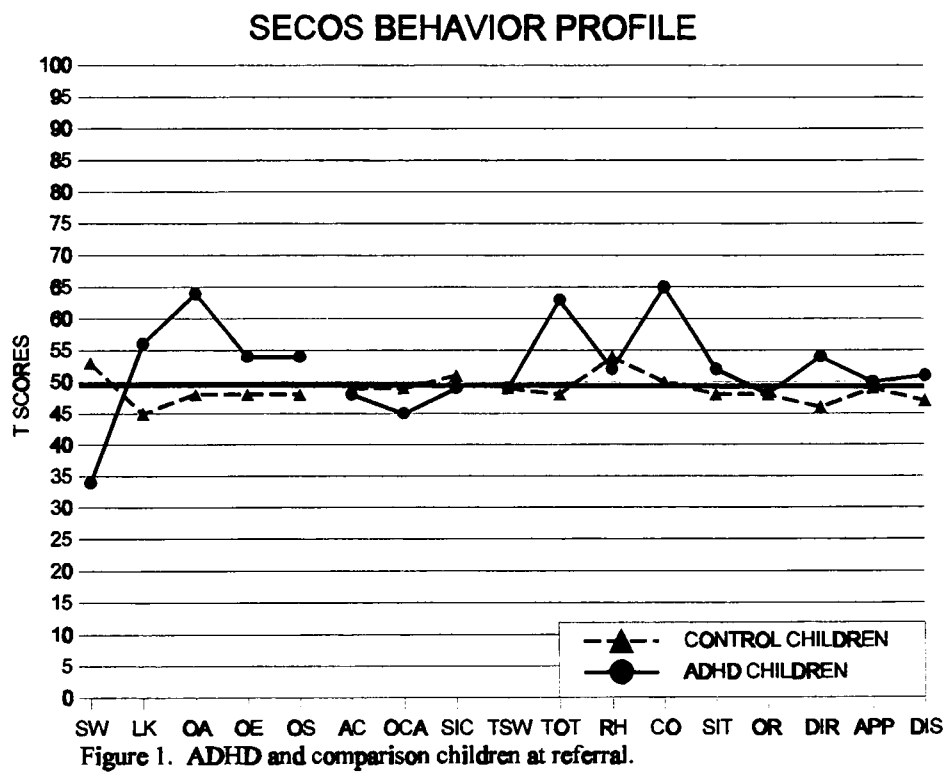
Internal Consistency of Combined Dependent Variables

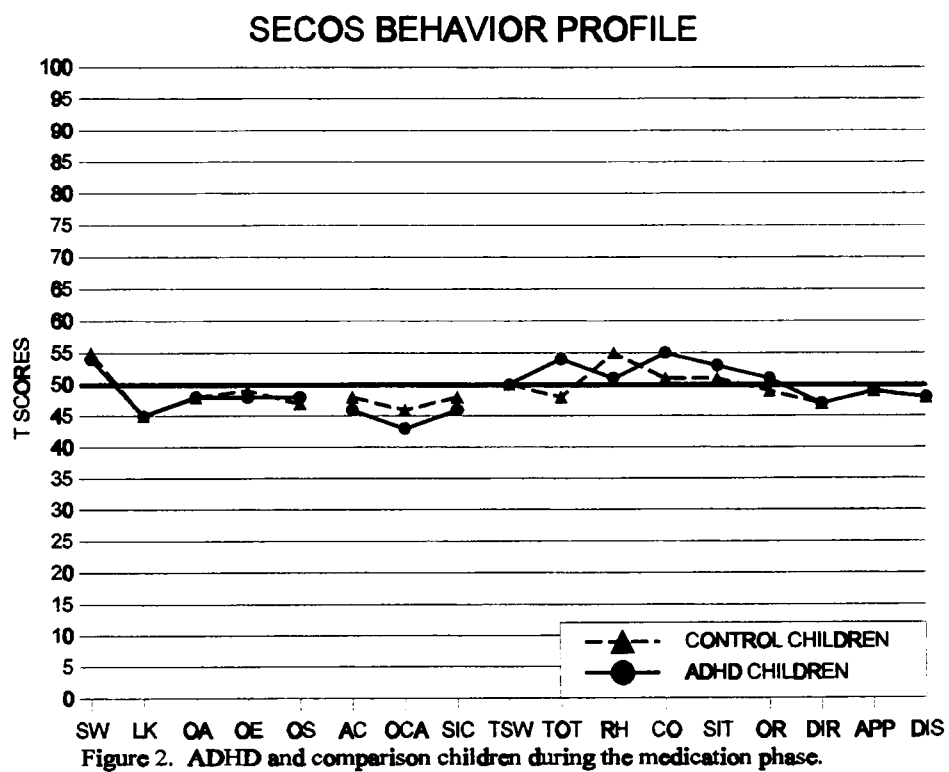
To test for the internal consistency of combined dependent variables, Cronbach's coefficients alpha were calculated. Moderate-to-high coefficients alpha were obtained for off-task / negative teacher attention (other activity + call out + teacher approach other + direction + disapproval; .89), social interaction with teacher / positive (raise hand + social interaction with teacher + teacher approach schoolwork + opportunity to respond + approval; .79), social interaction with child (approach child + other child approach + social interaction with child; .88), and out of seat (out of seat event + out of seat state; .62), suggesting that each is homogeneous. The homogeneity of the combined variables is further supported by the item-total correlations. All item-total correlations within each of the combined variables exceeded .44, and many were in the .70 to .80 range.

Behavior Patterns of Hyperactive and Comparison Children: Referral and

Medication Conditions

Descriptive statistics summarizing SECOS data for each group at referral and during the medication phase are shown in Table 4. Figures 1 and 2 show





SECOS behavior profiles for each group both pre- and post-medication of the ADHD children. To examine changes in these measures, multivariate analyses of variance (MANOVA) with repeated measures (BMDP4V) were performed on the median and range scores for the dependent variables school work, off-task / negative teacher attention, social interaction with teacher / positive, social interaction with child, out of seat, and looking around. For median scores, the 2 (Group: ADHD, comparison) X 2 (Condition: referral, medication) MANOVA showed a significant main effect for group, $F(6, 29) = 5.86, p < .0004$, condition, $F(6, 29) = 4.29, p < .003$, and for the interaction of group and condition, $F(6, 29) = 3.45, p < .01$. For range scores, the 2 (Group: ADHD, comparison) X 2 (Condition: referral, medication) MANOVA showed a trend for group effects, $F(6, 29) = 2.30, p < .06$, and a significant main effect for condition, $F(6, 29) = 4.23, p < .003$, and for the interaction of group and condition $F(6, 29) = 3.84, p < .006$. The following results refer to multivariate and univariate tests of contrast analyses.

Tests of behavior magnitude. Significant between-subjects effects were found for group at referral, $F(6, 29) = 5.10, p < .001$, but not during the medication condition, $F(6, 29) = .99, p < .45$. Univariate tests indicated that, at referral, hyperactive children spent less time engaged in school work than comparison

children and more time engaged in other activity, looking around, and out of seat behavior. Hyperactive children also solicited (called out) and received more negative teacher attention (teacher approach other, direction, disapproval) than comparison children. No significant differences were found for positive social interaction with teacher or social interaction with child. For tests of behavior change from referral to medication, significant effects were found for the ADHD group, $F(6, 29) = 7.16, p < .0001$, but not for the comparison group, $F(6, 29) = .58, p < .74$. Univariate tests indicated significant improvements in ADHD children's attention to task and decreases in maladaptive behavior. No significant differences were found for positive social interaction with teacher or social interaction with child.

Tests of behavior variability. Significant between-subjects effects were found for group at referral, $F(6, 29) = 3.18, p < .01$, but not during the medication condition, $F(6, 29) = .68, p < .66$. Univariate tests indicated that, prior to intervention, hyperactive children had significantly greater within-subject variability than comparison children for school work, looking around, off-task / negative teacher attention, and out of seat behavior. No significant differences in within-subject variability were found for positive social interaction with teacher or social interaction with child. For tests of behavior change from referral to medication,

significant effects were found for the ADHD group, $F(6, 29) = 6.79, p < .0001$, but not for the comparison children, $F(6, 29) = 1.28, p < .29$. Univariate tests indicated significantly reduced variability in on-task behavior, looking around, off-task / negative teacher attention, social interaction with child, and out of seat.

Secondary Analyses

Individual Change

Individual differences in ADHD children's response to treatment were evaluated using criteria proposed by Jacobson and Truax (1991) for evaluating the clinical significance of change: (a) meaningful change and (b) reliable change. Although these criteria were used to classify each ADHD child's treatment response for all dependent measures on SECOS (see Table 5), school work was chosen to represent clinically significant change as it is frequently the standard by which treatment effects are measured.

An individual is considered to meet the condition of meaningful change if his or her posttreatment school work mean is more likely to be drawn from the normative population than from the ADHD population. The meaningful change score is calculated by the following formula:

$$c = \frac{S_0 M_1 + S_1 M_0}{S_0 + S_1} = \frac{10 (33.7) + 16.9 (50)}{10 + 16.9} = 44$$

The condition of reliable change is met if the Reliable Change Index (RCI) is

greater than 1.96 using the following formula:

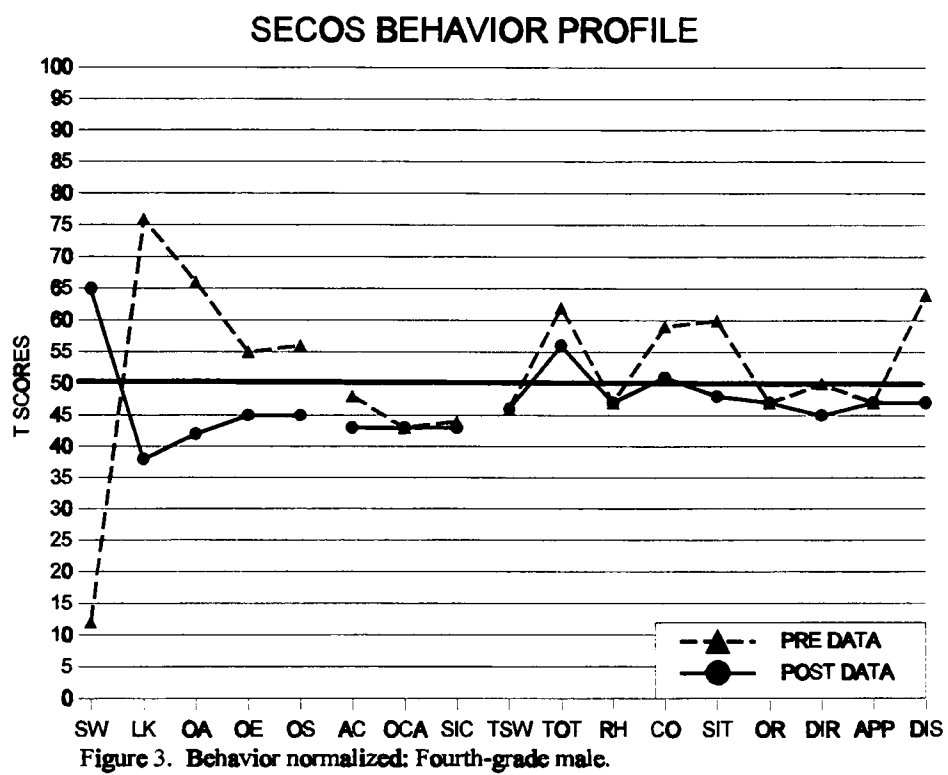
$$RC = (SW_{\text{post}} - SW_{\text{pre}}) / S_{\text{Diff}}$$

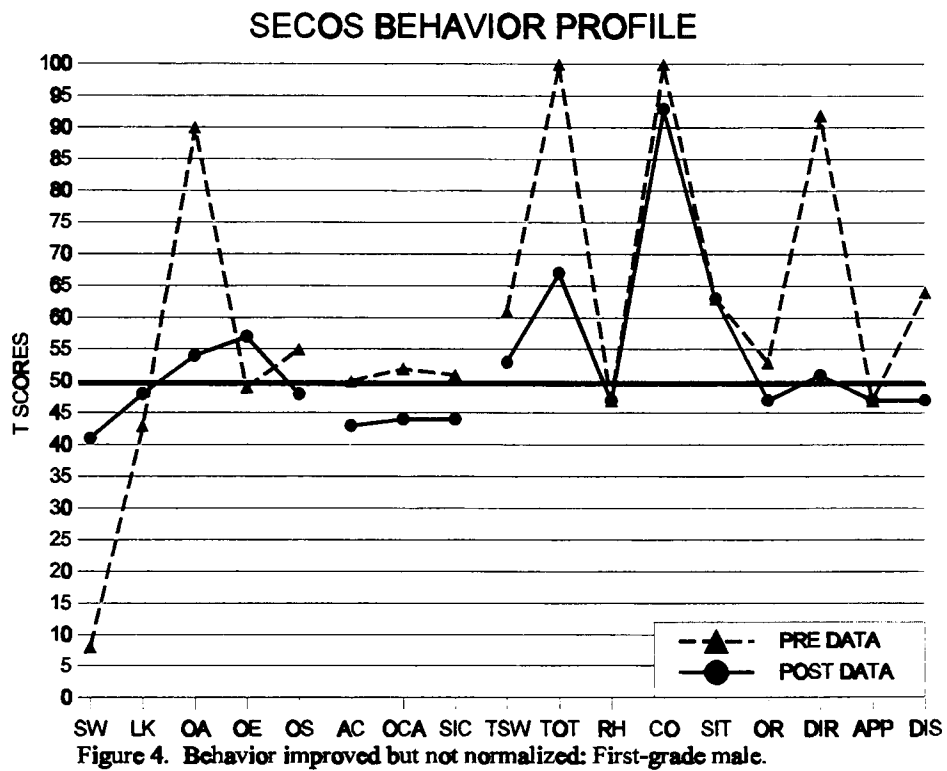
$$S_{\text{Diff}} = \sqrt{2(S_E)^2}$$

$$S_E = S_1 \sqrt{1 - r_{xx}}$$

ADHD children meeting both criteria of meaningful change and reliable change were considered "normalized;" those meeting the criteria of reliable change, but whose behavior remained closer to the mean of the ADHD population than the normal population were considered improved but not normalized; children classified as exhibiting no change did not meet the criterion of reliable change.

For on-task behavior (i.e., school work), the percentage of children whose behavior was normalized was 33% (n=6) (i.e., see Figure 3). One child met the criterion of reliable change, but his posttreatment score for school work was closer to the ADHD population mean (Score = 41) than the normal population mean; his behavior was considered improved but not normalized (see Figure 4). Sixty-two percent of the total sample (n=11) exhibited no change in on-task behavior according to the Jacobson and Truax criterion. However, of those children classified as exhibiting no change in on-task behavior, 36% (n=4) exhibited normalized behavior for all SECOS dependent measures at referral, and 64% (n=7) met the conditions of reliable and meaningful change for other problematic





behaviors following medication. For example, Figure 5 shows a first-grade female for which there was no improvement in on-task behavior from the referral to medication condition. Reliable and meaningful change were obtained, however, for the SECOS behaviors raise hand, calling out, and social interaction with teacher, indicating a significant positive change in this child's classroom behavior. Corresponding reductions in negative responses from the teacher (i.e., teacher approach other, disapproval) were also evident on SECOS.

Correspondence Between Academic Measures and SECOS Data

A series of independent t tests were conducted to assess the statistical equivalence of academic productivity (i.e., percentage of work complete) and academic efficiency (i.e., percentage of work correct) for ADHD and comparison children at referral and during the medication condition. In addition, t tests for paired differences were utilized to test the hypothesis of no change in these scores.

During the referral condition, academic productivity and efficiency scores for ADHD children were significantly different from scores obtained by comparison children ($p < .03$, $p < .0004$, respectively). The percentage of work completed by ADHD children at referral was 62 while comparison children completed 98% of academic work assigned. In addition, ADHD children were correct only 44% of the time while comparison children were correct 95% of the

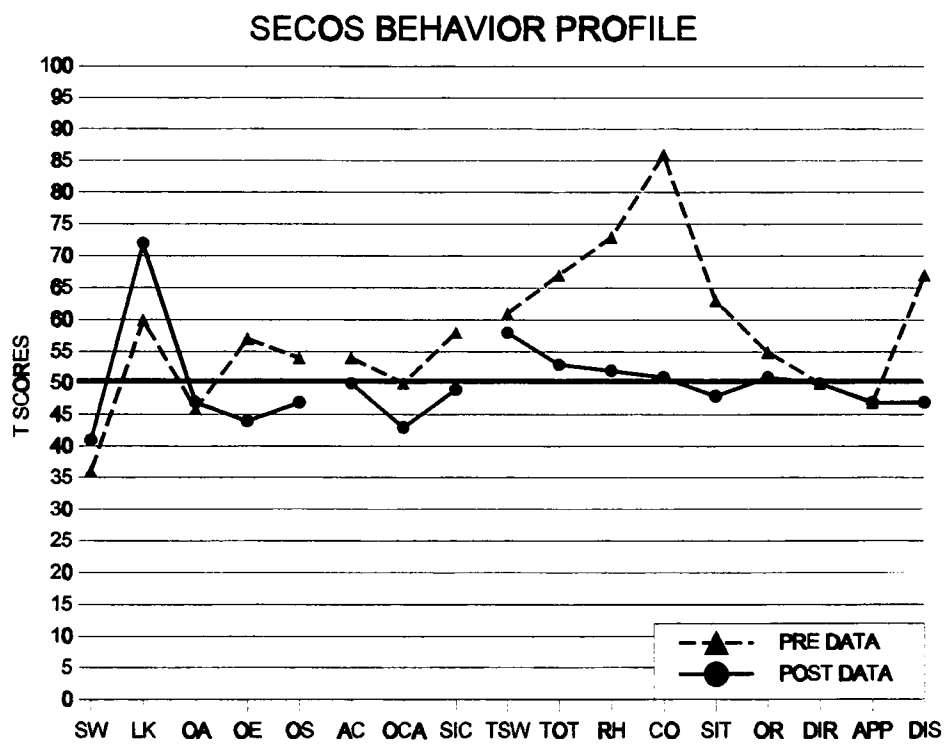


Figure 5. No change in on-task behavior: First-grade female.

time. During the medication condition, no differences were found between the two groups for academic productivity. ADHD children completed assignments 98% of the time and 100% of assignments were completed by comparison children. However, academic efficiency scores were significantly different for the two groups ($p < .004$). ADHD children were correct 87% of the time while comparison children were correct 96% of the time.

In the analyses of change in academic scores from the referral to medication condition, t tests for paired differences were significant for ADHD children but not for comparison children. Both academic productivity and academic efficiency scores increased for ADHD children ($p < .04$, $p < .004$, respectively). The percentage of work completed by ADHD children changed an average of 37 points, and the percentage of work correct changed an average of 43 points.

CHAPTER 4

DISCUSSION

This study provided empirical evidence linking a specific instrument (i.e., SECOS) to a specified treatment (i.e., psychostimulant intervention) and represents a preliminary investigation of the treatment validity of SECOS. For purposes of this study, treatment validity was defined in terms of the sensitivity of a measuring instrument to changes in behavior that result from intervention and treatment (Hoge, 1985). Treatment validity of specific instruments was conceptualized as an issue separate from questions of the utility of the overall assessment process in order to minimize concerns resulting from negative outcomes and strengthen the integrity of treatment-related decisions. In contrast to previous investigations in which treatment validity (i.e., instrument sensitivity) was based on treatment outcome, support for the treatment validity of SECOS in the assessment of ADHD children was provided through evidence that the code is sensitive to both known behavioral differences distinguishing ADHD children from their peers and the effects of stimulant medications on these behaviors as documented in numerous studies. The findings in this study were further substantiated by concomitant changes in children's daily academic work.

Consistent with previous investigations in which children with ADHD were

compared to normal controls (Atkins et al., 1985; Barkley et al., 1988; Zentall, 1980), SECOS data indicated that, at referral, ADHD children were more frequently off-task and exhibited more total disruptive behaviors than controls. That is, ADHD children more frequently engaged in other activity or simply looked around the classroom more than control children. They were also more frequently out of their seats and for longer durations, interrupted more (i.e., called out), and received more negative teacher attention (i.e., teacher approach other, direction, disapproval) than the control group. Data on SECOS were also consistent with findings in previous research in which ADHD children display significantly more variable classroom behavior than controls (Abikoff et al., 1977; Atkins et al., 1985; Schaughency & Rothlind, 1991).

Following medication, previously documented enhancements in the behavioral functioning of ADHD children were evident on SECOS (Carlson, et al., 1992; DuPaul & Rapport, 1993; Pelham, 1993; Pelham et al., 1985; Rapport et al., 1985). ADHD children sustained attention to task at a level indistinguishable from that of their non-hyperactive peers. Significant reductions were also noted in off-task, classroom disruptive behaviors, and negative teacher attention. The overall magnitude of scores obtained by ADHD children at follow-up for looking around, other activity, out of seat-event, out of seat-state, calling out, teacher approach

other, direction, and disapproval were no longer statistically different from scores obtained by comparison children. In addition to the overall level of change observed, significant reductions in the day-to-day variability of the classroom behavior of ADHD children were manifested on SECOS.

Permanent product measures (i.e., academic productivity and efficiency) were included as a source of independent verification of the findings in this study as previous investigations have shown improvements in these measures following treatment with stimulant medication (Carlson et al., 1992; DuPaul & Rapport, 1993; Pelham et al., 1985; Rapport, et al., 1985). In a manner consistent with SECOS data, academic productivity (i.e., percentage of work complete) and efficiency (i.e., percentage of work correct) scores were significantly lower for ADHD children than comparison children at referral. In addition, no differences were found between the two groups during the medication condition for academic productivity; and, although differences were noted in academic efficiency, the percentage of work completed accurately by ADHD children increased from 44% at referral to 87% at medication. Similar changes in academic productivity from the referral to medication condition were evident for the ADHD children. No changes were found in academic productivity or efficiency for the comparison children from the referral to medication condition.

Although not addressed directly in this study, the question of "relative" benefits (i.e., the benefit of a particular assessment procedure compared to available alternative procedures) has been proposed as a further manifestation of treatment validity. Within this framework, the relative benefits of observational assessment must outweigh the costs in terms of time and resources for the procedure to be considered truly beneficial. For example, observational assessment has been routinely compared to data from informant reports or rating scales. While the latter are more cost efficient (i.e., widely available, require little time to administer) and are an important source of information as parents and teachers typically initiate the referral process and have the greatest amount of exposure to the child's behavior (McMahon, 1984), they may be unsuitable for purposes of planning treatment strategies and evaluating outcomes of specific behaviors. There is a general lack of correspondence between information derived from rating scales and direct observation of specific behaviors in a given context (Bailey, Bender, & Montgomery, 1983; Cosper & Erickson, 1984; Shapiro, Lentz, & Sofman, 1985). In addition, ratings are not particularly useful for pinpointing target behaviors for change and are frequently insensitive to changes in behaviors following intervention (Atkins, Pelham, & White, 1990; Atkins & Pelham, 1991; McMahon, 1984). A second set of analyses, designed to examine the sensitivity of

SECOS to individual as opposed to group-level responses to stimulant medication, illustrated the utility of SECOS for treatment selection and evaluation.

Utilizing the reliable change index (RCI) and the *c* statistic, children in the ADHD sample were divided into three groups based on changes in their pre- to posttreatment scores on the SECOS behavior "school work". By this criteria, 33% of ADHD children obtained posttreatment scores within the normal range of functioning for school work, one child's behavior was considered improved but not normalized, and 62% of the ADHD sample exhibited no change in behavior. However, 55% of ADHD children classified as exhibiting no change had pretreatment school work scores in the normalized range before receiving medication; thus, underestimating the percentage of children showing normalized on-task behavior. In addition, while these criteria define an "indisputably improved group," the price of such a rigorous standard is that some children will show moderate or even extreme change (i.e., in the most deviant cases) and still fail to reach clinical significance (Ankuta & Abeles, 1993). In fact, only those children whose school work score increased between three and five standard deviations from baseline to follow-up reached the criteria for clinical significance specified in this study; whereas, by convention, one standard deviation is considered a reliable and valid standard for change (Barkley, DuPaul, & McMurray, 1991; Fischer &

Newby, 1991; Gadow, Nolan, Paolicelli, & Sprafkin, 1991). Had a less rigorous standard for change (i.e., $SD = 1$) been adopted in this study, 72% ($n = 13$) of the ADHD sample would have reached the criteria of clinical significance for on-task behavior.

In practice, however, the Jacobson and Truax (1991) criteria together with a measure (i.e., SECOS) that is sensitive to the child and environmental variables potentially controlling target behaviors do facilitate making essential decisions concerning the treatment of choice for individual children. For example, one first-grade male's behavior improved at follow-up but was not normalized (see Figure 4). Data on SECOS indicate that although there was significant improvement in this child's on-task behavior, he was still performing at a below-average level for his age group. Additionally, while other activity (i.e., drawing in his workbook, taking things in and out of his desk) was normalized, he still solicited (i.e., called out) and received teacher attention (i.e., teacher approach other, social interaction with teacher) at an above-average rate. Data on SECOS clearly indicate specific behaviors to target for intervention. These data suggested that this child's negative behavior may have been reinforced by teacher attention as he received more contacts from the teacher when engaged in some other activity rather than the academic assignment. One strategy for further diminishing inappropriate behavior

is to make teacher attention contingent on appropriate behavior (Abramowitz, & O' Leary, 1991; DuPaul, Guevremont, & Barkley, 1992; Newby, Fischer, & Roman, 1991).

Overall, results of the present study demonstrate the efficacy and utility of SECOS in the assessment of ADHD children. SECOS comprises sensitive and meaningful behaviors that can be used to identify core deficits, suggest treatment strategies, and evaluate individual treatment response. Notwithstanding, a number of issues remain for future research. First, the benefits of using SECOS in the assessment of ADHD children relative to alternative procedures (e.g., rating scales) were not addressed directly in this study. Future investigations could assess relative benefits through a procedure suggested by Hayes et al., 1987. To illustrate, treatment outcomes for two groups of ADHD children would be evaluated: those for whom treatment decisions were based on both rating scale and observation data, and those for whom treatment decisions were based on rating scale data only. If children whose treatments were based on both types of data improve more than those whose treatment is based on rating scale data only, the treatment validity of the observation data would be established. Second, sample sizes were too small to permit an assessment of the ability of SECOS to discriminate between the three subtypes of ADHD now recognized by DSM-IV.

As each of these diagnoses may dictate differential treatments (Power & DuPaul, 1996), the ability of SECOS to discriminate between these subtypes and offer information that leads to appropriate actions in their regard is an important question for future research. Finally, while the conclusions drawn from this study are bolstered by the inclusion of a non-referred comparison group and permanent product measures (i.e., academic assignments), the results of this study would have been strengthened by independent corroboration of stimulant effects (e.g., by parents and teachers) on the behavior of ADHD children. Although adult perceptions of a child's behavior may not reflect the child's actual behavior, improved parent and teacher perceptions may be considered a treatment goal in and of itself so that referrals do not re-occur.

References

References

- Abikoff, H., Gittelman-Klein, R., & Klein, D.F. (1977). Validation of a classroom observation code for hyperactive children. Journal of Consulting and Clinical Psychology, 45(5), 772-783.
- Abramowitz, A.J., & O'Leary, S.G. (1991). Behavioral interventions for the classroom: Implications for students with ADHD. School Psychology Review, 20, 220-234.
- Achenbach, T.M. (1991). Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington: University of Vermont Department of Psychology.
- Alessi, G.J., & Kaye, J.H. (1983). Behavior Assessment for School Psychologists. Kent, OH: National Association of School Psychologists.
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education. (1985). Standards for Educational and Psychological Testing. Washington, D.C.: American Psychological Association.
- American Psychological Association (APA), American Educational Research Association (AERA), National Council on Measurement in Education (NCME). (1974). Standards for Educational and Psychological Tests. Washington, D.C.: American Psychological Association.
- American Psychiatric Association. (1994). Diagnostic and Statistical Manual of Mental Disorders (4th ed.). Washington, DC: Author.
- Ankuta, G.Y., & Abeles, N. (1993). Client satisfaction, clinical significance, and meaningful change in psychotherapy. Professional Psychology: Research and Practice, 24(1), 70-74.
- Atkins, M.S., & Pelham, W.E. (1991). School-based assessment of attention deficit-hyperactivity disorder. Journal of Learning Disabilities, 24(4), 197-205.
- Atkins, M.S., Pelham, W.E., & Licht, M.H. (1985). A comparison of objective classroom measures and teacher ratings of attention deficit disorder. Journal of Abnormal Child Psychology, 13(1), 155-167.
- Atkins, M.S., Pelham, W.E., & White, K.J. (1990). Hyperactivity and attention

- deficit disorders. In M. Hersen and V. Van Hasselt (Eds.), Psychological Aspects of Developmental and Physical Disabilities: A Casebook (pp. 137-156). New York: Pergamon Press.
- Bailey, D.B., Bender, W.N., & Montgomery, D.L. (1983). Comparison of teacher, peer, and self ratings of classroom and social behavior of adolescents. Behavioral Disorders, 8, 153-160.
- Barkley, R.A. (1990). Attention Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment. New York: Guilford Press.
- Barkley, R.A., DuPaul, G.J., & McMurray, M.B. (1990). Comprehensive evaluation of attention deficit disorder with and without hyperactivity as defined by research criteria. Journal of Consulting and Clinical Psychology, 58(6), 775-789.
- Barkley, R.A., DuPaul, G.J., & McMurray, M.B. (1991). Attention deficit disorder with and without hyperactivity: Clinical response to three doses levels of methylphenidate. Pediatrics, 4, 519-532.
- Barkley, R.A., Fischer, M., Newby, R.F., & Breen, M.J. (1988). Development of a multimethod clinical protocol for assessing stimulant drug response in children with attention deficit disorder. Journal of Clinical Child Psychology, 17, 14-24.
- Bernardi, R.A. (1994). Validating research results when cronbach's alpha is below .70: A methodological procedure. Educational and Psychological Measurement, 54(3), 766-775.
- Bordens, K.S., & Abbott, B.B. (1991). Research Design and Methods: A Process Approach. California: Mayfield Publishing Co.
- Carlson, G.A., & Rapport, M.D. (1989). Diagnostic classification issues in attention-deficit hyperactivity disorders. Psychiatric Annals, 19, 576-583.
- Carlson, C.L., & Bunner, M.R. (1993). Effects of methylphenidate on the academic performance of children with attention-deficit hyperactivity disorder and learning disabilities. School Psychology Review, 22, 184-198.
- Carlson, C.L., Pelham, W.E., Milich, R., & Dixon, J. (1992). Single and combined effects of methylphenidate and behavior therapy on the classroom performance of children with attention-deficit hyperactivity disorder. Journal of Abnormal Child Psychology, 20(2), 213-232.

- Cone, J.D. (1982). Validity of direct observation assessment procedures. In D.P. Hartmann (Ed.), Using Observers to Study Behavior: New Directions for Methodology of Social and Behavioral Science (pp. 67-79). San Francisco: Jossey-Bass.
- Cosper, M.R., & Erickson, M.T. (1984). Relationships among observed classroom behavior and three types of teacher ratings. Behavioral Disorders, 9, 189-195.
- DuPaul, G.J., Barkley, R.A., & McMurray, M.B. (1991). Therapeutic effects of medication on ADHD: Implications for School Psychologists. School Psychology Review, 20(2), 203-219.
- DuPaul, G.J., Guevremont, D.C., & Barkley, R.A. (1992). Behavioral treatment of attention-deficit hyperactivity disorder in the classroom. Behavior Modification, 16, 204-225.
- DuPaul, G.J., & Rapport, M.D. (1993). Does methylphenidate normalize the classroom performance of children with attention deficit disorder? Journal of the American Academy of Child and Adolescent Psychiatry, 32(1), 190-198.
- Fellers, G., & Saudargas, R.A. (1987). Classroom behaviors of learning disabled and non-handicapped girls. Learning Disability Quarterly, 10, 231-236.
- Fellers, G., & Saudargas, R.A. (1988). Classroom behavior of gifted elementary school children. Paper presented at the National Association of School Psychologists Convention, Chicago, IL.
- Fischer, M., & Newby, R.F. (1991). Assessment of stimulant response in ADHD children using a refined multimethod clinical protocol. Journal of Clinical Child Psychology, 20(3), 232-244.
- Foster, S.L., & Cone, J.D. (1980). Current issues in direct observation. Behavioral Assessment, 2, 313-338.
- Frick, P.J., & Lahey, B.B. (1991). The nature and characteristics of attention-deficit hyperactivity disorder. School Psychology Review, 20(2), 163-173.
- Furman, W., & Drabman, R. (1981). Methodological Issues in Child Behavior Therapy. In M. Hersen, R.M. Eisler, & P.M. Miller (Eds.), Progress in Behavior Modification (Vol. 11, 31-64). New York: Academic Press.

- Gadow, K.D., Nolan, E.E., Paolicelli, L.M., & Sprafkin, J. (1991). A procedure for assessing the effects of methylphenidate on hyperactive children in public school settings. Journal of Clinical Child Psychology, 20(3), 268-276.
- Ghiselli, E.E., Campbell, J.P., & Zedeck, S. (1981). Measurement Theory for the Behavioral Sciences. San Francisco: Freeman and Company.
- Greenhill, L.L. (1992). Pharmacologic treatment of attention deficit hyperactivity disorder. Pediatric Psychopharmacology, 15, 1-27.
- Haile-Griffey, E.L. (1992). Obtaining Normative Data on the Classroom Behavior of Average Achieving Elementary School Children. Unpublished doctoral dissertation. University of Tennessee, Knoxville, TN.
- Hayes, S.C., Nelson, R.O., & Jarrett, R.B. (1986). Evaluating the quality of behavioral assessment. In R.O. Nelson & S.C. Hayes (Eds.), Conceptual Foundations of Behavioral Assessment. New York: Guilford.
- Hayes, S.C., Nelson, R.O., & Jarrett, R.B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. American Psychologist, 42, 963-974.
- Heller, K., Holtzman, W., & Messick, S. (1982). Placing Children in Special Education: A Strategy for Equity. Washington DC: National Academy Press.
- Hoge, R.D. (1985). The validity of direct observation measures of pupil classroom behavior. Review of Educational Research, 55(4), 469-483.
- Hoge, R.D., & Luce, S. (1979). Predicting achievement from classroom behavior. Review of Educational Research, 49(3), 479-496.
- Jacob, R.G., O'Leary, K.D., & Rosenblad, C. (1978). Formal and informal classroom settings: Effects of hyperactivity. Journal of Abnormal Child Psychology, 6, 47-59.
- Jacobson, N.S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting and Clinical Psychology, 59(1), 12-19.
- Johnston, J.M., & Pennypacker, H.S. (1980). Strategies and Tactics of Human Behavioral Research. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Keller, H.R. (1986). Behavioral observation approaches to assessment. In H. Knoff (Ed.), The Assessment of Child and Adolescent Personality (pp. 353-397). New York: Guilford Press.
- Klesges, R.C., Woolfrey, J., & Vollmer, J. (1985). An evaluation of the reliability of time sampling versus continuous observation data collection. Journal of Behavior Therapy and Experimental Psychiatry, 16(4), 303-307
- Kratochwill, T.R., & McGovern, J.E. (1996). Clinical diagnosis, behavioral assessment, and functional analysis: Examining the connection between assessment and intervention. School Psychology Review, 25(3), 342-354.
- Kronenberger, W.G., & Meyer, R.G. (1996). The Child Clinician's Handbook. Massachusetts: Allyn & Bacon.
- Landau, S., & Moore, L.A. (1991). Social skill deficits in children with attention-deficit hyperactivity disorder. School Psychology Review, 20(2), 235-251.
- McGee, R., Williams, S., & Silva, P.A. (1984). Behavioral and developmental characteristics of aggressive, hyperactive, and aggressive-hyperactive boys. Journal of the American Academy of Child Psychiatry, 23, 270-279.
- McMahon, R.J. (1984). Behavioral checklists and rating scales. In T.H. Ollendick and M. Hersen (Eds.), Child Behavioral Assessment: Principles and Procedures (pp. 80-105). New York: Pergamon.
- Merrell, K.W. (1994). Assessment of Behavioral, Social, and Emotional Problems: Direct and Objective Methods for Use With Children and Adolescents. New York: Longman Publishing.
- Messick, S. (1980). Test validity and the ethics of assessment. American Psychologist, 35, 1012-1027.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741-749.
- Murphy, M.J., & Harrop, A. (1994). Observer error in the use of momentary time sampling and partial interval recording. British Journal of Psychology, 85(2), 169-179.
- Nelson, R.O. (1988). Relationship between assessment and treatment within a behavioral perspective. Journal of Psychopathology and Behavioral

Assessment, 10, 155-170.

- Nelson, R.O., & Hayes, S.C. (1979). The nature of behavioral assessment: A commentary. Journal of Applied Behavior Analysis, 12, 491-500.
- Newby, R.F., Fischer, M., & Roman, M.A. (1991). Parent training for families of children with ADHD. School Psychology Review, 20, 252-265.
- Pelham, W.E., Bender, M.E., Caddell, J., Booth, S., & Moorer, S.H. (1985). Methylphenidate and children with attention deficit disorder: Dose effects on classroom academic and social behavior. Archives of General Psychology, 42, 948-952.
- Pelham, W.E., Walker, J.L., Sturges, J., & Hoza, J. (1989). Comparative effects of methylphenidate on ADD girls and ADD boys. Journal of the American Academy of Child and Adolescent Psychiatry, 28, 773-776.
- Pelham, W.E. (1993). Pharmacotherapy of children with attention-deficit hyperactivity disorder. School Psychology Review, 22, 199-227.
- Powell, J. (1984). On the misrepresentation of behavioral realities by a widely practiced direct observation procedure: Partial interval (one-zero) sampling. Behavioral Assessment, 6, 209-219.
- Power, T.J., & DuPaul, G.J. (1996). Attention-deficit hyperactivity disorder: The reemergence of subtypes. School Psychology Review, 25(3), 284-295.
- Rapport, M.D., Jones, J.T., DuPaul, G.J., Kelly, K.L., Gardner, M.J., Tucker, S.B., & Shea, M.S. (1987). Attention deficit disorder and methylphenidate: Group and single-subject analyses of dose effects on attention in clinic and classroom settings. Journal of Clinical Child Psychology, 16(4), 329-338.
- Rapport, M.D., Stoner, G., DuPaul, G.J., Birmingham, B.K., & Tucker, S. (1985). Methylphenidate in hyperactive children: Differential effects of dose on academic, learning, and social behavior. Journal of Abnormal Child Psychology, 13(2), 227-244.
- Rowley, G. (1989). Reliability and other indicators of data quality in classroom observation research. Journal of Classroom Interaction, 24(1), 22-29.
- Saudargas, R.A. (1992). State-Event Classroom Observation System (SECOS). Unpublished observation manual. University of Tennessee, Knoxville, TN.

- Saudargas, R.A., & Lentz, F.E. (1986). Estimating percent of time and rate via direct observation: A suggested observational format. School Psychology Review, 15(1), 36-48.
- Saudargas, R.A., & Zanolli, K. (1990). Momentary time sampling as an estimate of percentage time: A field validation. Journal of Applied Behavior Analysis, 23(4), 533-537.
- Schaughency, E.A., & Rothlind, J. (1991). Assessment and classification of attention deficit hyperactivity disorders. School Psychology Review, 20(2), 187-202.
- Shapiro, E.S. (1987). Behavioral Assessment in School Psychology. New Jersey: Lawrence Erlbaum Associates, Inc.
- Shapiro, E.S., & Kratochwill, T.R. (1988). Behavioral Assessment in Schools: Conceptual Foundations and Practical Applications. New York: The Guilford Press.
- Shapiro, E.S., Lentz, F.E., & Sofman, R. (1985). Validity of rating scales in assessing aggressive behavior in classroom settings. Journal of School Psychology, 23, 69-80.
- Shapiro, E.S., & Skinner, C.H. (1990). Best practices in observation and ecological assessment. In Best Practices in School Psychology. (2nd ed.) (pp. 507-518). Washington, D.C.: The National Association for School Psychologists.
- Sheridan, S.M., Dee, C.C., Morgan, J.C., McCormick, M.E., & Walker, D. (1996). A multimethod intervention for social skills deficits in children with ADHD and their parents. School Psychology Review, 25(1), 57-76.
- Silva, F. (1993). Psychometric Foundations and Behavioral Assessment. Newbury Park, CA: Sage Publications.
- Skiba, R.J. (1989). The importance of construct validity: Alternative models for the assessment of behavioral disorders. Behavioral Disorders, 14(3), 175-185.
- Slate, J.R., & Saudargas, R.A. (1986a). Differences in learning disabled and average students' classroom behavior. Learning Disability Quarterly, 9, 61-68.

- Slate, J.R., & Saudargas, R.A. (1986b). Differences in the classroom behaviors of behaviorally disordered and regular class children. Behavior Disorders, 12(1), 45-53.
- Suen, H.K., & Ary, D. (1989). Analyzing Quantitative Behavioral Observation Data. Hillsdale, NJ: Lawrence Erlbaum.
- Wadlington, L. (1988). The effect of instructional setting upon student attention. Unpublished doctoral dissertation, University of Tennessee, Knoxville, TN.
- Whitley, B.E. (1996). Principles of Research in Behavioral Science. California: Mayfield Publishing.
- Zentall, S.S. (1980). Behavioral comparisons of hyperactive and normally active children in natural settings. Journal of Abnormal Child Psychology, 8(1), 93-109.

Appendixes

Appendix A

Table 1. Diagnostic criteria for Attention-Deficit Hyperactivity Disorder.

A. Either (1) or (2):

- (1) six (or more) of the following symptoms of inattention that have persisted for at least six months to a degree that is maladaptive and inconsistent with developmental level:

Inattention

- (a) often fails to give close attention to details or makes careless mistakes in school work, work, or other activities
- (b) often has difficulty sustaining attention in tasks or play activities
- (c) often does not seem to listen when spoken to directly
- (d) often does not follow through on instructions and fails to finish school work, chores, or duties in the workplace (not due to oppositional behavior or failure to understand instructions)
- (e) often has difficulty organizing tasks and activities

Table 1. Continued.

(f) often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort (such as school work or homework)

(g) often loses things necessary for tasks and activities (e.g., toys, school assignments, pencils, books, or tools)

(h) is often easily distracted by extraneous stimuli

(i) is often forgetful in daily activities

(2) six (or more) of the following symptoms of hyperactivity-impulsivity have persisted for at least 6 months to a degree that is maladaptive and inconsistent with developmental level:

Hyperactivity

(a) often fidgets with hands or feet or squirms in seat

(b) often leaves seat in classroom or in other situations in which remaining seated is expected

Table 1. Continued.

- (c) often runs about or climbs excessively in situations in which it is inappropriate
(in adolescents or adults, may be limited to subjective feelings of restlessness)
 - (d) often has difficulty playing or engaging in leisure activities quietly
 - (e) is often "on the go" or often acts as if "driven by a motor"
 - (f) often talks excessively
-

Impulsivity

- (g) often blurts out answers before questions have been completed
 - (h) often has difficulty awaiting turn
 - (i) often interrupts or intrudes on others (e.g., butts into conversations or games)
-

Table 2. The State-Event Classroom Observation System (SECOS).

State Behaviors	Description	Examples
School Work.	Student is doing assigned work while having head and eyes oriented towards the material or teacher in a lecture or discussion.	Looking at the teacher giving a lecture.
Look.	Students eyes and head are oriented away from the classroom activity.	Looking at neighbor.
Other Activity.	Student is engaged in a nonacademic activity.	Sharpening pencil.
Out of Seat.	Time spent out of desk or chair for any reason.	Walking around the room.
Social Interaction with Child.	Verbal or nonverbal interactions between target student and any other student.	Talking while standing in line.
Social Interaction with Teacher.	Target student and teacher are interacting either verbally or nonverbally.	Teacher is helping the student with math homework.
Child Event Behaviors	Description	Examples
Out of Seat.	Frequency of getting out of desk or chair.	The act of leaving one's seat.
Approach Child.	Student attempts to obtain attention from another student verbally or physically.	Taps neighbor on arm.

Table 2. Continued.

Other Child Approach.	Another student either physically or verbally attempts to make contact with the student you are observing.	Neighbor calls out to observed student.
Raise Hand.	Student solicits teacher attention by raising hand.	Student raises hand in response to question.
Call Out.	Student verbally attempts to contact the teacher.	Student calls out response to question directed at neighbor.
Teacher Event Behaviors	Description	Examples
Teacher Approach Schoolwork.	Teacher verbally or physically contacts student who is engaged in schoolwork.	Teacher calls on student who is attending to the lecture.
Teacher Approach Other.	Teacher verbally or physically contacts student who is not engaged in schoolwork.	Calls on student who is tying shoes.
Opportunity to Respond.	Teacher asks student to respond to an academic question.	Asks student to read a sentence.
Direction.	Teacher directs student to change activity or behavior.	Tells student to sit down.

Table 2. Continued.

Approval.	Any positive comment regarding social or academic behavior.	"That's really neat writing."
Disapproval.	Any negative comment regarding social or academic behavior.	"Do that again and you'll leave the room."

Table 3. Interobserver agreement for ADHD and comparison observations at referral and medication.

Frequency and Occurrence Agreement at Referral and Medication: Comparison Children

<u>Behavior</u>	<u>Frequency Agreement</u> Referral, Medication	<u>Occurrence Agreement</u> Referral, Medication
School Work.	95.6, 95.1	92.5, 92.8
Look.	88.4, 81.5	76.2, 71.9
Other Activity.	85.7, 86.7	72.6, 72.7
Social Interaction with Child.	83.3, 84.8	73.8, 79.5
Social Interaction with Teacher.	95.8, 88.5	87.5, 78.5
Out of Seat State.	90.2, 98.6	88.9, 89.5
Approach Child.	83.3, 98	65.4, 84.4
Other Child Approach.	94.2, 84.1	81.9, 79.5
Out of Seat Event.	100, 81.8	80.3, 81.8

Frequency and Occurrence Agreement at Referral and Medication: ADHD Children

<u>Behavior</u>	<u>Frequency Agreement</u> Referral, Medication	<u>Occurrence Agreement</u> Referral, Medication
School Work.	90.3, 95.8	84.8, 95.8

Table 3. Continued.

Look.	96.9, 83.1	82.4, 72.6
Other Activity.	93.5, 83.4	78.2, 72.3
Social Interaction with Child.	94.4, 81.3	84.9, 71.3
Social Interaction with Teacher.	95.5, 89.6	93.6, 81.3
Out of Seat State.	86.1, 87.5	82.6, 90.4
Approach Child.	84, 80.6	68.5, 75
Other Child Approach.	77.9, 85.2	62.9, 77.6
Out of Seat Event.	83.5, 83.3	66.9, 75.1
Raise Hand.	90.9, 90.8	81.8, 79.6
Call Out.	71.9, 79.6	61, 74.8
Teacher Approach School Work.	75.7, 77.8	72.7, 75
Teacher Approach Other.	88.6, 75	79.5, 66.7

Table 4. Means, standard deviations, medians, and ranges for SECOS behaviors at referral and medication.

Comparison Children at Referral, Medication				
<u>Behavior</u>	<u>Mean</u>	<u>Standard Deviation</u>	<u>Median</u>	<u>Ranges</u>
School Work.	53 , 55	7, 7	53, 55	41, 32
Look.	45, 45	5, 6	45, 45	19, 23
Other Activity.	48, 49	7, 6	48, 48	19, 23
Social Interaction with Child.	51, 48	7, 6	51, 48	30, 23
Social Interaction with Teacher.	48, 51	3, 6	48, 51	11, 17
Out of Seat State.	48, 47	4, 5	48, 47	13, 16
Approach Child.	49, 48	7, 7	49, 48	21, 27
Other Child Approach.	48, 46	8, 6	49, 46	30, 20
Out of Seat Event.	53, 48	11, 8	48, 49	22, 28
Raise Hand.	54, 55	16, 14	54, 55	53, 46
Call Out.	50, 51	8, 6	50, 51	30, 22
Teacher Approach School Work.	48, 49	3, 6	49, 50	11, 15
Teacher Approach Other.	48, 47	4, 3	48, 48	10, 16

Table 4. Continued.

Opportunity to Respond.	48, 49	2, 4	48, 49	8, 17
Direction.	46, 46	3, 2	46, 47	13, 12
Approval.	49, 49	4, 4	49, 49	16, 19
Disapproval.	47, 47	3, 4	47, 48	0, 9

ADHD Children at Referral, Medication

<u>Behavior</u>	<u>Mean</u>	<u>Standard Deviation</u>	<u>Median</u>	<u>Ranges</u>
School Work.	34, 54	16, 8	34, 54	52, 24
Look.	57, 45	11, 8	56, 45	38, 34
Other Activity.	64, 49	17, 9	64, 48	56, 25
Social Interaction with Child.	50, 46	7, 5	49, 46	27, 15
Social Interaction with Teacher.	53, 53	8, 9	52, 53	17, 34
Out of Seat State.	56, 49	11, 4	54, 48	32, 13
Approach Child.	50, 46	9, 6	48, 46	33, 19

Table 4. Continued.

Other Child Approach.	46, 44	5, 5	45, 43	11, 15
Out of Seat Event.	61, 48	22, 6	54, 48	60, 22
Raise Hand.	53, 51	13, 6	52, 51	28, 16
Call Out.	67, 56	28, 15	65, 55	94, 47
Teacher Approach School Work.	48, 49	4, 4	49, 50	17, 15
Teacher Approach Other.	62, 53	15, 7	63, 54	59, 35
Opportunity to Respond.	48, 50	1, 5	48, 51	8, 21
Direction.	52, 47	12, 4	54, 47	47, 11
Approval.	49, 48	4, 4	50, 49	20, 15
Disapproval.	50, 48	4, 2	51, 48	17, 13

Table 5. Data used in determining clinical significance.

Symbol	Definition
M_0	Mean score for normative population. ^a
M_1	Pretreatment mean score for ADHD population.
S_0	Standard deviation of normative population. ^a
S_1	Standard deviation of ADHD population.
X_{pre}	Pretreatment score for an individual ADHD child.
X_{post}	Posttreatment score for an individual ADHD child.
r^{xx}	Test-retest reliability. ^b
S_{Diff}	Standard error of the difference between the two scores.

^aBased on a normative population of 486 first- through fifth-grade children in a large East Tennessee school system.

^bTest-retest reliability, $r^{xx}=.76$, based on the school work scores of a sample of 102 elementary school-age children with approximately one week elapsed time between observations.

Appendix B

THE UNIVERSITY OF TENNESSEE
KNOXVILLE



College of Liberal Arts
Department of Psychology
307 Austin Peay Building
Knoxville, Tennessee 37996-0900
(615) 974-2531
FAX (615) 974-3330

Dear Parents:

In cooperation with the Lenoir City Schools, Dr. Richard Saudargas and Kathy H. Trotter of the University of Tennessee's Department of Psychology are conducting research on the classroom behavior of elementary school children. The general goal of the research is to determine if children who have been referred for evaluation by the school's assessment specialist behave differently in the classroom than those children who have not been referred. Children who have been referred for assessment and children who have not been referred will be observed. The non-referred children selected for observation will be identified by the classroom teacher as not exhibiting any special problems in the classroom with respect to behavior or schoolwork. All observations will take place in the child's regular classroom during independent seatwork and will be conducted by University of Tennessee undergraduate and graduate students. Each child will be observed on five days for about one hour a day both prior to and following intervention and assessment of the referred child. In order to describe the students participating, we would like access to your child's school records including achievement test scores, intelligence test scores, and birthdate. Further, if your child has been referred for a problem that requires medication, we would like to know the prescribed dosage and approximate time(s) of day the medicine is dispensed.

We would like your permission to observe your child. There are no risks in being observed since codes will be used to identify each child and no names will be used in written reports of the project. Giving your permission is voluntary, and not giving permission will not affect your child's program in any way. In addition, you may withdraw your permission at any time.

We will be happy to answer any questions about the project or discuss the results with you. We can be reached at 974-4361 or 974-8780. To give your permission, please sign below. Return one copy to your child's teacher and keep one copy for yourself. Thank you very much.

Sincerely,

Richard A. Saudargas, Ph.D.

Kathy H. Trotter, M.S.

I understand the purpose of the project described in this letter and give my permission for my child to be observed.

Name of Child: _____ Phone: _____

Parent Signature: _____ Date: _____

Vita

Katherine Hulse-Trotter received both Bachelor of Science (B.S.) and Master of Science (M.S.) degrees from The University of Tennessee at Chattanooga. The Doctor of Philosophy (Ph.D.) degree was awarded by The University of Tennessee at Knoxville in December of 1996. Kathy majored in Psychology, and her program of study specialty area was child development.

While enrolled in graduate school, Kathy worked as a graduate research assistant and as a graduate teaching assistant. Her research interests include typical and atypical development in childhood, attention-deficit hyperactivity disorder, and behavioral assessment in the schools. Her teaching interests parallel these areas.

Kathy is a member of the Phi Kappa Phi National Honor Society. In addition, she received the Science Alliance research award, the Sigma XI research award, and the Student Provost research award. Kathy is a member of the American Psychological Association, the Association for Behavior Analysis, and the Society for Research in Child Development.