

Psychometric Characteristics of a Multidimensional Measure to Assess Impairment: The Child and Adolescent Functional Assessment Scale

Kay Hodges, Ph.D.,^{1,3} and Maria M. Wong, Ph.D.²

The Child and Adolescent Functional Assessment Scale (CAFAS) is a multidimensional measure of degree of impairment in functioning. Interrater reliability data are presented for lay raters, graduate students, and frontline staff. Reliability was high for the total score and behaviorally-oriented scales. Construct, concurrent, and discriminant validity were assessed with the sample of children and adolescents evaluated at the Fort Bragg Demonstration Evaluation Project. Youth and their caregivers were evaluated via interview and self-completed instruments at four time points. Significant correlations were found between the CAFAS and other related constructs. Concurrent validity was demonstrated by logistic regression analyses examining the relationship between CAFAS ratings and problematic behaviors endorsed on measures completed by parents, teachers, or the youth. Youth with higher CAFAS total scores were much more likely to have poor social relationships, difficulties in school, and problems with the law. Discriminant validity was assessed with a repeated measures analysis of variance with intensity of care at intake and time as factors. Youth who were inpatients or in residential treatment centers at intake had higher CAFAS scores than those who were outpatients. These findings provide strong evidence for the reliability and validity of the CAFAS.

KEY WORDS: functioning; children; adolescents; impairment; multidimensional measures; scale.

¹Director of Clinical Training, Department of Psychology, Eastern Michigan University, Ypsilanti, MI.

²Research Fellow, Institute for Social Research, University of Michigan, Ann Arbor, MI.

³Correspondence should be directed to Kay Hodges, 2140 Old Earhart Road, Ann Arbor, MI 48105.

The Child and Adolescent Functional Assessment Scale, (Hodges, 1989) assesses the degree of impairment in functioning in children and adolescents secondary to emotional, behavioral, or substance use problems. Assessing impairment has become a required component of applications for federal block grant funds (Federal Register, p. 29425). With the publication of the most recent edition of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ea.) (*DSM-IV*; American Psychiatric Association, 1994), impairment in functioning became a required diagnostic criteria for most disorders. Impairment has also emerged as a salient concept in the movement toward managed behavioral health care. Assessment of functioning is viewed as important in prioritizing delivery of services and in evaluating efficacy and cost effectiveness of clinical programs.

Until recently, the most widely used measures to assess impairment have been global measures, similar to the scale used for Axis V in *DSM-IV*. While the simplicity of global measures is attractive, they are vulnerable to rater bias, both unintentional and purposeful. The rater typically assigns a score between 0 and 100, with 10 anchor points. No justification or description of the impairing behavior is directly linked to the score. The ten descriptive statements do not systematically describe impairment across domains and many are not objective. Very different presentations of behavior can receive the same ratings. Without any anchoring to behaviors, a global score is seen as so prone to bias that some states have even discontinued using these global scores for program evaluation (Hodges, 1994, Hodges & Gust, 1995).

The major advantage of a multidimensional measure is that several dimensions of functioning can be rated, permitting more precise descriptions. The CAFAS is a multidimensional measure in which the child is rated on five scales and the caregiver on two scales. The rater reviews a list of descriptions of behavior, and chooses the items which capture the child's functioning. Many anchor descriptions are provided, and the score for each scale is based on the items selected. Also, since the total score is derived from adding up the individual scale scores, the total score is potentially less vulnerable to rater bias or demand characteristics.

Reliability and predictive validity data have been previously reported on for the CAFAS, based on the sample of youth assessed in the Ft. Bragg Demonstration Evaluation Study, which hereafter will be referred to as the Evaluation Study (Bickman et al., 1994). Satisfactory interrater reliability was reported for the trained lay raters used in the Evaluation Study (Hodges, Bickman, Kurtz, & Reiter, 1991). The relationship between the CAFAS total score at intake and subsequent service utilization at 6 and 12 months post intake was examined by Hodges and Wong (1996). The CAFAS score predicted restrictiveness of care, cost of services, number of total

services, and number of bed days. Moreover, the CAFAS was a better predictor than instruments assessing symptoms or psychopathology, such as the Child Behavior Checklist, and was also a better predictor than any psychiatric diagnosis.

The present report focuses on the construct and concurrent validity of the CAFAS by examining its relationship with other related constructs and with reports of problematic behaviors obtained from measures independent of the CAFAS. The data were from the Ft. Bragg Evaluation Study. Interrater reliability is also presented for samples of raters other than those used in Ft. Bragg; raters included undergraduate and graduate students and staff working in child service agencies. The reliability study will be presented first, followed by the validity study.

STUDY 1: INTERRATER RELIABILITY

Method

Raters

Three samples of student/lay raters consisted of university student volunteers who received extra credit for participating in the study. The lay rater/undergraduate sample, which consisted of 19 women and 1 man, was solicited from a sophomore level class in Psychology of Women. The graduate samples were first semester graduate students in a clinical psychology master's program. One sample had 10 women and 4 men, and the other sample had 7 women and 4 men.

The agency staff raters sample consisted of front-line social service and juvenile court workers from a county in Michigan. All had a bachelor's or master's degree in social work, criminal justice, education, or psychology. There were 6 women and 3 men.

Procedures for Training and Assessing Reliability

Interrater reliability was established using written vignettes which were the same for all samples. The 20 vignettes included a description of the family constellation, a detailed reporting of the child's responses on a structured interview, and the parent's responses to the same interview when asked about the child's symptoms and behavior. The vignettes were based on disguised actual cases, were lengthy (e.g., five to six typed, single-spaced pages), and were characterized by various problematic issues experienced in the field (e.g., inconsistent reports between the child and the parent).

The training for the student raters was entirely self-administered, which consisted of reviewing instructions for scoring and example scored vignettes (Hodges, 1990a). For the agency staff sample, the raters participated in a workshop, and rated 15 vignettes within the time constraints of the workshop.

Instrument

The CAFAS (Hodges, 1989) yields a total score as well as scores for seven scales. Five scales refer to the child's functioning and are as follows: Role Performance (i.e., how effectively the youth fulfills societal roles at home, in school, and in the community); Thinking (i.e., ability of youth to use rational thought processes); Behavior Toward Others/Self (i.e., appropriateness of youth's daily behavior); Moods/Emotions (i.e., modulation of the youth's emotional life); and Substance Use (i.e., youth's substance use and the extent to which it is inappropriate and disruptive). A score for each scale and a total score (the sum of all five scale scores which assess the child's functioning) are generated. Two additional scales are available to assess the caregiver: Basic Needs (i.e., caregiver's ability to provide for the basic needs of the child including food, shelter, clothing, medical care, and safety), and Family/Social Support (i.e., the degree to which the quality of the relationships and the guidance and nurturance provided within the family meet the child's developmental needs). Of the Caregiver scales, only the Family/Social Support scale was scored in the reliability study.

For each scale, the rater determines the severity level which best describes the youth's most severe level of dysfunction during a specified period. The scores assigned to each of the categories are as follows: 30 for Severe (severe disruption or incapacitation); 20 for Moderate (persistent disruption or major occasional disruption of functioning); 10 for Mild (significant problems or distress); 0 for Minimal or No Impairment (no disruption of functioning). For each scale and each severity level, there are a number of items describing behavior. The rater reviews the items in the Severe category first. If any item describes the child's functioning, the child is assigned a score of "30." If none of the items in the Severe category characterizes the child, the rater continues to the Moderate category, progressing through the remainder of the categories as needed to describe the youth's level of functioning.

If the rater is very familiar with the child's behavior and functioning, this measure takes approximately 10 minutes to complete. Various sources of information can be used to make the CAFAS rating (e.g., interviews with parent and/or child, review of detailed case records).

Analyses

Each rater provided scores for each scale for each vignette. The raters' scores for each scale for each vignette were compared to a criterion score, which was a consensus score between the first author and a board-certified child psychiatrist. Pearson product moment correlations were calculated, and *z* transformations were used to average correlations across raters.

Intraclass correlations (ICC) based on analysis of variance procedures were calculated to provide an estimate of agreement between raters. Guidelines for interpreting ICC values were used as follows: values greater than .75 indicate good reliability, values between .50 and .75 indicate fair reliability, and values below .50 indicate poor reliability (Spitzer, Fleiss, & Endicott, 1978).

Analyses were conducted for the total score and for each scale, except for the Thinking scale. Due to the low frequency of formal thought problems or organicity in the vignettes, which were designed to be representative of typical clinical presentations, calculations could not be made for the Thinking scale. Analyses were also calculated for the Caregiver: Family/Social Support scale.

Results and Discussion

Table 1 shows the Pearson correlations for the four samples. The correlations for the total score and for the Substance Use scale were high, with all correlations being .92 or higher. Correlations were consistently high for two scales: Behavior Toward Others/Self (ranging from .83 to .93) and Role Performance (ranging from .79 to .90). The Moods/Emotions and the

Table 1. Correlations Between Trainees and Criterion for Four Training Samples on the CAFAS

CAFAS Scores	Under- Graduate Students (<i>N</i> = 20)	Graduate Students Sample 1 (<i>N</i> = 14)	Graduate Students Sample 2 (<i>N</i> = 11)	Agency Staff (<i>N</i> = 9)
Child Scores				
Total	.92	.92	.96	.95
Role Performance	.79	.82	.88	.90
Behavior	.83	.88	.87	.93
Moods/Emotions	.74	.77	.88	.94
Substance Use	.99	.98	.99	.98
Caregiver Scores				
Family/Social Support	.82	.73	.82	.84

Caregiver: Family/Social Support scales were characterized by moderately high reliability, with all values being .73 or greater.

The intraclass correlations in Table 2, which reflect on the agreement among raters, show a similar pattern. All correlations (i.e., on each scale for each sample) were characterized by good or fair reliability using the Spitzer et al. (1978) guidelines. Both the total score for the child and the Substance Use scale had good reliability for all samples, with all values being .84 or higher. The ICC values for the Behavior Toward Others/Self and Role Performance scales were lower but still in the good range, except for the undergraduates' rating of Role Performance. As with the correlations, the values for the Moods/Emotions and Family/Social Support scales were lower and mostly in the fair range.

Good reliability was demonstrated for the total CAFAS score for all three types of raters, including naive undergraduates (lay raters), graduate students, and agency staff working with children. Furthermore, the method of training was efficient and cost effective, given that, for three of the samples, the results were found when training was done entirely independently, using a self administered training manual. The ICC values observed for the CAFAS total score were similar to those reported for the most widely used global measure for children, the Children's Global Assessment Scale (CGAS; Shaffer et al., 1983), with ICC values for the CGAS ranging from .84 to .93 across various studies using highly trained professionals (Bird, Canino, Rubio-Stipec, & Ribera, 1987; Garber, Kriss, Koch, & Lindholm, 1988; Shaffer et al., 1983; Steinhausen, 1987). The results for the individual scales indicated that the more behaviorally-oriented scales (i.e., Role Performance, Behavior Toward Others/Self, Substance Use) had the highest reliability. These findings were consistent with other studies, in which higher reliability was generally found for externalizing behavioral problems, compared to in-

Table 2. Intraclass Correlations for Four Training Samples on the CAFAS

CAFAS Scores	Under- Graduate Students (<i>N</i> = 20)	Graduate Students Sample 1 (<i>N</i> = 14)	Graduate Students Sample 2 (<i>N</i> = 11)	Agency Staff (<i>N</i> = 9)
Child Scores				
Total	.84	.89	.88	.87
Role Performance	.68	.74	.79	.79
Behavior	.74	.79	.79	.77
Moods/Emotions	.63	.75	.71	.69
Substance Use	.92	.88	.87	.96
Caregiver Scores				
Family/Social Support	.70	.65	.74	.54

ternalizing behaviors such as anxiety and depression (Edelbrock, Costello, Dulcan, Conover, & Kalas, 1986; Hodges, Gordon, & Lennon, 1990).

STUDY 2: VALIDITY STUDY

Method

Subjects

Subjects were dependents of Army personnel (i.e., active, retired, or discharged) who were referred for mental health services, and agreed to participate in the Ft. Bragg Demonstration Evaluation Study (Breda, 1996). They were recruited from three Army bases: Ft. Bragg, North Carolina, Ft. Campbell, Kentucky, and Ft. Stewart, Georgia, and were evaluated at four time points: intake, 6 months post intake, 12 months post intake, and 18 months post intake. Because the primary focus of this paper was to examine the reliability and validity of the CAFAS, the presence of CAFAS scores was used as a criterion for selecting respondents. Only those respondents whose CAFAS scores were present in the data were included in the study. There were 984 respondents at intake (Wave 1), 780 respondents at 6 months post intake (Wave 2), 617 respondents at 12 months post intake (Wave 3), and 373 respondents at 18 months post intake (Wave 4). The children ranged in age from 5 to 17 years old, with the mean age being 11.1 years at intake.

Table 3 summarizes the demographic characteristics of the validity sample. Approximately two-thirds of the sample were male, and the majority were preadolescents, defined as 13 years old or younger. There were more preadolescent boys (63%) than adolescent boys (37%), whereas there were about equal numbers of preadolescent girls (49%) and adolescent girls (51%). In the modal family, at least one parent figure had some education beyond high school (but was not a college graduate); the household income was between \$30,000 and \$40,000; and there were two caregivers in the home. The caregivers could be a biological parent, stepparent, adoptive parent, or a relative. The demographics of the sample at 6 months, 12 months, and 18 months were very similar to that at intake, with the exception that the mean age increased.

Raters

The CAFAS was rated by 28 interviewers who administered a structured diagnostic interview to the youth and the parent separately. Seventy-five percent of the raters/interviewers were female. They were

Table 3. Characteristics of Respondents in the Validity Study

	Wave 1 (N = 984)		Wave 2 (N = 780)		Wave 3 (N = 617)		Wave 4 (N = 373)	
	n	%	n	%	n	%	n	%
<i>Age</i>								
5-7	236	24.0	156	20.0	90	14.5	18	4.8
8-10	187	19.0	163	20.9	138	22.4	91	24.4
11-13	219	22.2	169	21.7	132	21.4	68	18.2
14 and above	342	34.8	292	37.4	257	41.7	196	52.6
<i>Gender</i>								
Male	622	63.2	493	63.2	387	62.7	229	61.4
Female	362	36.8	287	36.8	230	37.3	144	38.6
<i>Race</i>								
Caucasian	778	79.1	626	80.2	512	83.0	295	79.1
African-American	202	20.5	151	19.4	104	16.9	76	20.4
<i>Family Income</i>								
<\$10,000	15	1.5	15	1.9	11	1.8	6	1.6
\$10,000-14,999	18	1.8	18	2.3	15	2.4	12	3.2
\$15,000-19,999	91	9.2	55	7.0	41	6.6	28	7.5
\$20,000-29,999	165	16.8	111	14.2	89	14.4	50	13.4
\$30,000-39,999	334	33.9	270	34.6	200	32.4	121	32.4
\$40,000-59,999	176	17.9	148	19.0	100	16.2	58	15.5
\$60,000 or more	126	12.8	107	13.7	88	14.1	53	14.2
<i>Caregiver's Education</i>								
Schooling below high school graduate	7	.7	7	.9	5	.8	2	.5
High School Graduate	159	16.2	123	15.7	90	14.6	63	16.9
Some Post-High School	548	55.7	414	53.0	307	49.8	186	49.9
College Graduate	143	14.5	117	15.0	88	14.3	52	13.9
Some Post-College	45	4.6	42	5.4	33	5.3	17	4.6
Advanced Graduate	71	7.2	53	6.8	42	6.8	24	6.4

Note. Percentages do not total 100% for some variables because of insufficient data.

trained by the first author in five separate training sessions, spanning a three-year period. Except for one rater who had an associate's degree, all had an undergraduate degree or were currently enrolled in undergraduate education. None of the interviewers had direct experience providing psychiatric services to children and adolescents, although four raters had Master's Degrees: two in educational guidance and counseling and two in social welfare. Ten other raters had Master's Degrees in non-mental health fields (i.e., theology, music, business, education, political science). The raters were hired for the Evaluation Study and were not involved in the treatment or treatment decisions for any of the subjects. The research pro-

ocols were collected, stored, and processed by the research team employees only.

Training entailed a day-and-a-half workshop and rating of 20 vignettes for the purpose of establishing reliability. The criteria for satisfactory reliability were a Pearson correlation of .85 or higher for the total CAFAS score, and .80 for Role Performance and Behavior Toward Others/Self. Throughout the study, 10% of all of the cases were reviewed via videotapes of the interviews. Remediation was conducted if raters' reliability began to deteriorate.

Measures

The measures used for this study were part of the larger battery used in the Evaluation Study and were described in Breda (1996).

CAFAS. The CAFAS was described in detail above. The time period rated for the CAFAS was the last month. The lay raters completed the CAFAS after interviewing the child and the parent with a structured diagnostic interview.

CAS and PCAS. The Child Assessment Schedule (CAS; Hodges, 1990b, 1990c; Hodges, Kline, Stern, Cytryn, & McKnew, 1982) is a structured diagnostic interview which generates information about diagnoses as well as difficulties across various life areas, such as school, peers, and family. There are parallel versions for the child and the parent, with the latter referred to as the PCAS. Approximately half of the items on the interview are diagnostically related. For each question, there is a response criterion, with the scoring options being: true, false, ambiguous, or not applicable. Data on the reliability and validity of the CAS are summarized in Hodges (1993). For the analyses in the present study, the total number of endorsements of items, which are all scored in the direction of pathology, was used.

Parent Self-Administered Packet. The parents were asked to complete on their own a compendium of widely-used questionnaires inquiring about youth and their families. For topics with insufficient coverage in the questionnaires, the evaluation staff supplemented the packet with additional questions. Of these additional questions, those used in the present study inquired about: (a) social relationships with others, (b) history of involvement with juvenile justice, and (c) behavior at school (i.e., general attitude, grades, specific behavioral problems). The specific questionnaires utilized in this study were the Child Behavior Checklist and the Burden of Care Questionnaire.

The CBCL was designed to obtain ratings of the competencies and behavioral/emotional problems of children aged 4 to 16 years old, as re-

ported by parents. The child is rated on 118 problem items using a three point scale for how true the item is for the child over the last six months. *T* scores provide comparisons with a normative sample of children. The psychometric data on the CBCL are contained in Achenbach (1991). The total problem score for the CBCL was used in the present study, as well as several specific risk behaviors (i.e., physically attacking people, threatening people, talking about killing self) and ratings on problems in social relationships with other children, siblings, and parents (e.g., having problems getting along with other children of the same age).

The BCQ was developed for use in the Evaluation Study to assess the impact on the family of having a child with serious emotional or behavioral problems. It is a 42-item self-report instrument with responses scored on 5- or 3-point scales. It contains items aimed at capturing the caregiver's objective burden (i.e., disruptive occurrences resulting from the child's problems) and subjective burden (i.e., the extent to which the caregiver feels burdened). The total score was used in the present study.

Youth Self-Report. The Youth Self-Report (YSR; Achenbach & Edelbrock, 1983) was contained in a self-report packet that youth ages 12-17 years completed on their own. The YSR was designed to obtain self-reports of competencies and problems from 11- to 18-year-olds. Its format is similar to the CBCL. YSR items regarding physically attacking people, threatening people, and talking about killing self were examined.

Teacher's Report Form. The Teacher's Report Form (TRF; Edelbrock & Achenbach, 1984) has a behavior problem profile that is similar to that of the CBCL. The TRF was designed to collect information on functioning in school. Teacher ratings of the youth's happiness and motivation, as well as the same risk behaviors assessed with the CBCL and the YSR, were analyzed.

Procedures

Children and parents were interviewed at the subjects' convenience, either in their home, at an office, or in a treatment facility. The parents, youth and teachers completed their respective self-administered packets on their own. The information garnered was used for research purposes only.

Analyses and Expectations

As the predictive validity of the CAFAS was demonstrated in a previous paper (Hodges & Wong, in press), the analyses in this paper focused on examining the construct, concurrent and discriminative validity of the

CAFAS. To demonstrate the construct and concurrent validity of the CAFAS, we examined the relationship between the CAFAS total score and: (a) a number of related global constructs regarding psychopathology and caregivers' perceived burden, and (b) specific problematic behaviors.

The global constructs included total psychopathology scores on the CAS, PCAS, CBCL, and BCQ. It was expected that the CAFAS total score, which indicates functional impairment, would have positive correlations with all of the global measures of psychopathology. It was also expected that such positive correlations would be consistent across Waves 1 through 4.

Specific problematic behaviors included behavioral items in the domains of social relationships, risk behaviors, juvenile justice, and school-related behaviors. Information regarding specific problematic behaviors was provided independently via self-completed measures done by parents, teachers, or the youth, all of whom had no knowledge of the CAFAS scores. It was expected that these behavioral items, which indicated important aspects of the role and emotional functioning of children and adolescents, would be significantly related to the CAFAS total score across different waves. Logistic regression analysis was performed separately on each behavioral indicator (0 = absence of pathology, or behavior similar to children of the same age, 1 = presence of pathology or behavior much worse than children of the same age). The CAFAS total score (0 = low, 1 = high; the cutoff point was 80 at Wave 1 and 50 for the other three waves) was used as an independent variable, and five demographics variables (i.e., age, sex, race, caregivers' education, and annual family income) were also included in the analyses to serve as controls. As the dependent variable was binary, we believed that the relationship between the CAFAS and specific behaviors would be best demonstrated by treating the CAFAS as a categorical instead of a continuous variable. There were approximately 20% of respondents who scored above 80 at Wave 1 and above 50 at Waves 2, 3, and 4. We considered these respondents to be seriously impaired.

To demonstrate the discriminative validity of the CAFAS, we examined the changes in the CAFAS total scores of respondents grouped by initial level of care. The youth were classified into three groups (inpatient/residential treatment center [RTC], alternative care, outpatient) according to the services they were receiving at Wave 1. Alternative care included services which were usually seen as an alternative to traditional residential care and included: home-based services, day treatment, specialized foster care (i.e., foster parents are paid extra for providing for a child needing special care, such as behaviorally disordered), and group home. It was expected that the inpatient treatment group would have the highest CAFAS total scores, the outpatient treatment group would have the lowest CAFAS total scores, and the alternative care group would have intermediate scores. As

all respondents received some form of treatment during the study, it was expected that the discriminative power of the CAFAS would be stronger at the initial waves. Repeated measures analysis of variance with level of care and time (Waves 1, 2, 3, and 4) were computed.

Results

Descriptive Statistics

The characteristics of respondents are presented in Table 3. Statistical analyses showed that there were no significant differences between those who continued to participate in the study and those who dropped out at Wave 4 with respect to age, $t(982) = -1.07$, n.s., gender, $\chi^2(1) = .85$, n.s., race, $\chi^2(1) = .00$, n.s., family education, $t(971) = -1.00$, n.s., and annual family income, $t(923) = -.28$, n.s. However, attrition seemed to be related to the severity of respondents' problem, $\chi^2(2) = 14.52$, $p < .001$. Those who dropped out were more likely to be receiving outpatient treatment at the beginning of the study (dropouts 78.7%, non-dropouts 70.4%), whereas those who continued to participate were more likely to be receiving inpatient (dropouts 17.1%, non-dropouts 19.9%) and alternative (dropouts 4.1%, non-dropouts 9.7%) types of treatment. Those who dropped out of the study might have significantly improved over time due to treatment and were not motivated to continue to participate.

The means and standard deviations for the CAFAS total scores for each wave were as follows: Wave 1, $M = 46$, $SD = 26$; Wave 2, $M = 31$, $SD = 26$; Wave 3, $M = 26$, $SD = 24$; and Wave 4, $M = 24$, $SD = 24$. CAFAS total scores decreased over time, indicating that most respondents had become less impaired by the end of the study.

Tables 4 and 5 show the correlations between the five scales of the CAFAS at different waves. The correlations between the scales were moderate as the five scales examined different aspects of impairment. Cronbach's alpha for the CAFAS at Waves 1, 2, 3, and 4 were .63, .68, .67, and .67 respectively.

Demographic Variables

In this section the relationship between the CAFAS and five demographics variables—age, sex, race, caregiver education, and annual family income—was examined. As the analyses involved a large number of tests (five demographic variables \times four waves \times eight CAFAS scores at each wave), there was a possibility of chance significant findings. Thus, the focus

Table 4. Zero-order Correlations Among CAFAS Scales at Waves 1 and 2

CAFAS Scales	Youth Scales				
	Role Performance	Thinking	Behavior Toward Others/Self	Moods/Emotions	Substance Use
<i>Youth</i>					
Role Performance	—	.17***	.53***	.28***	.22***
Thinking	.23***	—	.26***	.22***	.09**
Behavior Toward Others/Self	.61***	.23***	—	.37***	.13***
Moods/Emotions	.41***	.25***	.46***	—	.11***
Substance Use	.25***	.01	.19***	.09*	—

Note. Numbers in the upper triangle of the correlation matrix are Wave 1 data (N = 982-984); numbers in the lower triangle of the correlation matrix are Wave 2 data (N = 781). *p < .05. **p < .01. ***p < .001.

Table 5. Zero-Order Correlations Among CAFAS Scales at Waves 3 and 4

CAFAS Scales	Youth Scales				
	Role Performance	Thinking	Behavior Toward Others/Self	Moods/Emotions	Substance Use
<i>Youth</i>					
Role Performance	—	.25***	.58***	.37***	.32***
Thinking	.12*	—	.20***	.17***	.03
Behavior Toward Others/Self	.63***	.16**	—	.48***	.20***
Moods/Emotions	.42***	.17***	.48***	—	.11**
Substance Use	.29***	-.06	.22***	-.01	—

Note. Numbers in the upper triangle of the correlation matrix are Wave 3 data (N = 617); numbers in the lower triangle of the correlation matrix are Wave 4 data (N = 372-373). *p < .05. **p < .01. ***p < .001.

was to examine patterns of results, and only results that were significant in at least two waves were reported. Also, a more conservative criterion for statistical significance ($p < .01$) was adopted.

Older respondents scored higher on the scales of Role Performance, Wave 2: Beta = .17, $p < .001$; Wave 3: Beta = .19, $p < .001$, Substance Use, Wave 1: Beta = .37, $p < .001$; Wave 2: Beta = .31, $p < .001$; Wave 3: Beta = .27, $p < .001$; Wave 4: Beta = .25, $p < .001$, and Caregiver: Family/Social Support, Wave 1: Beta = .21, $p < .001$; Wave 3: Beta = .15, $p < .001$, than younger respondents. As expected, older respondents also had a higher total score than younger respondents, Wave 1: Beta = .18, $p < .001$; Wave 2: Beta = .17, $p < .001$; Wave 3: Beta = .13, $p < .001$.

.001. Male respondents were more impaired on the Role Performance scale, Wave 1: $t(982) = 3.84, p < .001$; Wave 2: $t(778) = -3.59, p < .001$; Wave 3: $t(615) = -2.61, p < .01$. However, the two sexes did not differ in the total scores at any time of the study. Race did not seem to affect CAFAS scores. Caucasians and respondents from other ethnic groups scored similarly on all the scales and the total scores. Caregivers' education level also did not seem to affect respondents' CAFAS scores. Respondents whose caregivers were well educated had similar CAFAS scores to those whose caregivers were less educated. When compared to respondents with higher family income, those with lower family income tended to have a higher score on the Caregiver: Basic Resources scale, Wave 1: Beta = $-.11, p < .001$; Wave 2: Beta = $-.18, p < .001$. Lower income families more often reported having difficulties meeting the basic material needs of respondents than did higher income families.

To summarize, age, gender, and family income were significantly related to the CAFAS. Comparatively speaking, age seemed to have a stronger relationship with the CAFAS than gender or family income. There were no significant effects of race or caregiver education on CAFAS scores. In another article focusing on Wave 1 data (Hodges & Wong, in press), the following results were reported: Caucasians had a slightly higher CAFAS total score than respondents from other ethnic groups, and respondents with low family income had a higher CAFAS total score than respondents with higher family income. Those findings were not reported here because a more conservative criterion for statistical significance was chosen, as mentioned above. Given that a number of studies have shown that demographic variables affect perceived stress and mental health (e.g., Almeida & Kessler, 1994, Mirowsky & Ross, 1989), all five demographic variables were included as controls in the subsequent analyses on problematic behavior.

The Relationship Between the CAFAS and Other Related Global Constructs

Table 6 presents the correlations between the total scores of the CAFAS and four other related measures, the CBCL, the CAS, the PCAS, and the BCQ. The results suggested that the CAFAS total scores were significantly related to all related constructs at different waves. The patterns of correlations were similar for the different waves. For instance, the zero-order correlations between the CAFAS and the CBCL were .42, .49, .48, .47 at Waves 1, 2, 3, and 4 respectively.

Table 6. Zero-Order Correlations Between CAFAS Total Scores and Global Measures of Psychopathology at All Four Waves

CAFAS Total Score	PCAS Total Score	CBCL Total Score	CAS Total Score	BCQ Total Score
Wave 1	.59*** (N = 984)	.42*** (N = 924)	.54*** (N = 675)	.36*** (N = 932)
Wave 2	.62*** (N = 781)	.49*** (N = 745)	.56*** (N = 556)	.42*** (N = 730)
Wave 3	.58*** (N = 617)	.48*** (N = 593)	.55*** (N = 469)	.43*** (N = 567)
Wave 4	.63*** (N = 373)	.47*** (N = 341)	.52*** (N = 287)	.42*** (N = 329)

Note. CAFAS = Child and Adolescent Functional Assessment Scale; PCAS = Child Assessment Schedule, Parent Version; CBCL = Child Behavior Checklist; CAS = Child Assessment Schedule; BCQ = Burden of Care Questionnaire.
*** $p < .001$.

The Relationship Between the CAFAS and Specific Behaviors

Problems in Social Relationships. Table 7 shows the odds ratios using the CAFAS total score to predict quality of social interactions while controlling for the effects of age, sex, race, caregivers' education, and annual family income. The CAFAS had a highly significant relationship with almost all of the behaviors reported by parents, teachers, and the youth themselves. For instance, the odds ratio for poor relationships with other children at Wave 1 was 3.28, suggesting that those who scored high on the CAFAS were approximately three times more likely as those who scored low on the CAFAS to have problems getting along with other children of the same age. High scorers on the CAFAS were also two to three times more likely to have poor relationships with schoolmates, parents, and teachers at Wave 1. The relationship between CAFAS total scores and poor social relationships remained consistent across different waves of the study.

Risk Behaviors. The CAFAS total score was positively related to behaviors that were directed at harming others and oneself (See Table 7). Respondents with high CAFAS scores had a stronger tendency to physically attack and threaten people than respondents with low CAFAS scores. For instance, according to information provided by parents at Wave 3, respondents with high CAFAS scores were five times more likely than others to have physically attacked people and four times more likely than others to have threatened people. High CAFAS scorers were also more likely to have suicidal thoughts than low CAFAS scorers. For instance, when compared

Table 7. Relationship Between CAFAS Total Score and Problematic Behavior

		Odds Ratio			
		Wave 1	Wave 2	Wave 3	Wave 4
Problems in Social Relationships with:					
Other children	Parents	3.28***	3.69***	2.58**	5.00***
Other students	Parents	3.11***	3.15***	4.73***	—
Siblings	Parents	1.43	3.29***	2.57***	2.05
Parents	Parents	2.53***	3.42***	3.08***	5.71***
Teachers	Parents	2.07*	3.21***	4.85***	—
Risk Behaviors					
Physically attacked people	Parents	2.37***	3.05***	5.69***	2.94**
	Teachers	1.31***	1.67	1.20	—
	Self	2.95***	2.55**	2.09	.48
Threatened people	Parents	3.06***	3.74***	4.24***	8.38***
	Teachers	2.43**	1.35	2.45*	—
	Self	2.82***	1.55	2.17*	2.04
Talked about killing self	Parents	3.70***	3.11***	4.02***	6.90***
	Self	2.73***	2.14*	2.86*	2.57

Note. Logistic regression analyses were performed using CAFAS Total Scores to predict behavior. Age, race, sex, caregiver's education, and annual family income were used as control variables. — = information not available or too few cases of disruptive behavior.

* $p < .05$. ** $p < .01$. *** $p < .001$.

to respondents with low CAFAS scores, those with high CAFAS scores reported twice as often having thoughts about killing oneself.

Involvement with Juvenile Justice. As can be seen in Table 8, there was a significantly positive relationship between the CAFAS total score and involvement with juvenile justice. The CAFAS was positively related to the likelihood of being arrested by police and convicted of a crime. For instance, at Wave 2, when compared to their counterparts, respondents with high CAFAS scores were five times more likely to have been arrested by police or convicted of a crime, three times more likely to have been placed on probation, and four times more likely to have spent time in prison or other correctional facilities. Note that at Wave 1, the time period for involvement with juvenile justice was "ever happened," while the time period for other waves was "last six months." As CAFAS ratings at all waves were based on the youth's functioning during the last month, it was not surprising that the relationship between juvenile justice variables and CAFAS ratings at Waves 2, 3, and 4 were stronger than that at Wave 1.

School-Related Behaviors. Table 9 describes the relationship between the CAFAS total score and behavior at school. A positive relationship was also expected between the CAFAS and negative attitude or disruptive be-

Table 8. Relationship Between CAFAS Total Score and Involvement with Juvenile Justice

	Informant	Odds Ratio			
		Wave 1	Wave 2	Wave 3	Wave 4
Arrested by police	Parents	1.72 ^a	5.65***	3.79**	6.83***
Convicted of a crime	Parents	1.45 ^a	5.29***	8.87***	3.73
Placed on probation	Parents	1.34 ^a	3.69***	8.88***	6.68**
Spent time in prison, or other correlational facilities	Parents	2.46 ^a	4.49***	7.90***	—
Saw probation or law enforcement officers	Parents	1.89**	—	—	—
Detention center	Parents	2.44*	—	—	—

Note. Logistic regression analyses were performed using CAFAS total scores to predict behavior. Age, race, sex, caregiver’s education, and annual family income were used as control variables. — = information not available or too few cases of involvement with juvenile justice.

^aAt Wave 1, the time period was “ever happened.” At Waves 2, 3, 4, the time period was “last six months.”

p* < .05 *p* < .01. ****p* < .001.

Table 9. Relationship Between CAFAS Total Score and School-Related Behavior

	Informant	Odds Ratio			
		Wave 1	Wave 2	Wave 3	Wave 4
Disliked school	Parents	1.83**	2.97***	3.07***	4.20***
Skipped school	Parents	2.49***	4.51***	5.98***	4.84***
Disciplined in school frequently in the last 6 months	Parents	1.92***	2.71***	2.52***	3.24***
Suspended at least once in the last 6 months	Parents	2.21***	1.74*	4.37***	5.19***
Happiness when compared to other students of the same age	Teachers	2.68**	2.85*	—	—
Grades in last 6 months	Parents	1.59*	2.83***	2.09**	2.91**
Worked much less hard than other students of the same age	Teachers	1.66	1.02	2.28*	—
Had to repeat a single grade in past 6 months	Parents	—	—	2.02*	5.16***

Note. Logistic regression analyses were performed using CAFAS Total Scores to predict behavior. Age, race, sex, caregiver’s education, and annual family income were used as control variables. — = information not available or too few cases of problematic school-related behavior.

p* < .05. *p* < .01. ****p* < .001.

havior at school. The more functionally impaired a student was, the less cognitive resources he or she had available to perform well in school. The findings were consistent with our expectations. High CAFAS scorers showed a general negative attitude towards school. For instance, at Waves 1 and 2, when compared to those with low CAFAS scores, respondents

with high CAFAS total scores were approximately twice as likely to dislike school and showed less happiness in school.

The CAFAS was also related to poor academic performance. For instance, at Wave 4, children and adolescents with high CAFAS scores were approximately three times more likely than others to have poor grades and five times more likely than others to have repeated a single grade in the last six months. Moreover, CAFAS total scores were positively related to being frequently disciplined by school authorities, being suspended, and skipping school.

The Development of Impairment over Time

Repeated measures analysis of variance with level of care (inpatient/RTC, alternative, and outpatient) and time (Waves 1, 2, 3, and 4) as factors were performed to examine the changes in CAFAS total scores over time. As there was a large attrition of respondents at Wave 4, the repeated measures analysis of variance was computed twice, once using data from the first three waves only, and once using data from all four waves. Both analyses suggested that the main effects of initial level of care, Waves 1-3: $F(2, 565) = 28.31, p < .001$; Waves 1-4: $F(2, 317) = 5.93, p < .01$, and time, Waves 1-3: $F(2, 1130) = 84.65, p < .001$; Waves 1-4: $F(3, 951) = 66.35, p < .001$, were significant. There was also a significant interaction between level of care and time, Waves 1-3: $F(4, 1130) = 6.29, p < .001$; Waves 1-4: $F(6, 951) = 3.39, p < .01$.

The differences in CAFAS scores at each wave were therefore examined separately. As can be seen in Table 10, the data were consistent with our expectations. One way analysis of variance suggested that there was a significant main effect of initial level of care at Waves 1, 2, and 3. The CAFAS total scores of the three treatment groups were different from one another, with the inpatient group receiving the highest impairment score,

Table 10. Outcome and Treatment Type: Examining Total CAFAS Scores Over Time Grouped by Initial Level of Care

CAFAS Total Score	Initial Level of Care								
	Inpatient/RTC			Alternative			Outpatient		
	<i>M</i>	<i>SE</i>	<i>n</i>	<i>M</i>	<i>SE</i>	<i>n</i>	<i>M</i>	<i>SE</i>	<i>n</i>
Intake	63.99	2.16	178	53.77	3.10	61	40.58	0.88	740
6 Months	40.71	2.54	140	38.70	4.12	46	28.65	1.00	591
12 Months	33.68	2.68	106	28.75	4.05	40	23.25	1.02	467
18 Months	25.41	3.01	74	21.94	3.39	36	23.69	1.50	262

followed by the alternative treatment group and the outpatient group, Wave 1: $F(2, 976) = 67.10, p < .001$; Wave 2: $F(2, 774) = 14.58, p < .001$; Wave 3: $F(2, 610) = 9.05, p < .001$. There was no significant main effect of level of care at Wave 4, probably due to the fact that a large number of respondents had successfully dealt with their problems over time. The means for the three groups ranged from 25 to 22 ($F(2, 368) = .29, n.s.$).

Discussion

The findings provide strong evidence for the reliability and validity of the CAFAS. Interrater reliability was excellent for the CAFAS total score. This was found for both lay raters and frontline professionals who work with children. Moreover, high levels of reliability were achieved even when the raters did training entirely on their own. The data relevant to the concurrent, construct, and discriminant validity were also very supportive of the CAFAS.

The zero-order correlations among individual scales and Cronbach alpha levels indicated that the scales were significantly and moderately correlated with each other. Only the Substance Use scale had insignificant relationships with other scales, which was understandable given that some respondents with substance problems did not have impaired thinking or were not depressed. Moderate correlations would be expected given that each scale reflects a distinct facet of functioning.

The significant correlations between the CAFAS total score and the CBCL as well as the BCQ total scores provided evidence of concurrent validity. The moderate magnitude was expected given that the CBCL is primarily a list of symptoms, with each symptom contributing the same weight toward the total score. In contrast, in the CAFAS, behaviors have different weights depending on the extent of impairment associated with the behavior (e.g., not attending school is given a much higher weight than not obeying school rules). The BCQ was designed to assess the impact of the youth's problems on the family, and thus would be expected to be related positively to the extent of the youth's impairment in functioning.

Analyses relevant to the individual CAFAS scales were supportive of the measure's validity. As expected, older respondents had higher impairment scores than younger respondents on the Substance Use and Role Performance scales. The gender effect observed for the Role Performance scale was consistent with the well-documented male preponderance of youth with conduct disorders (Eme & Kavanaugh, 1995). While role performance at school, at home, or in the community could be impaired for a variety of reasons (e.g., children with depression or posttraumatic stress

disorder might refuse to go to school), the behavior of conduct disordered youth would typically warrant a high score on this scale. The finding that low income families scored significantly higher than high income families on the Basic Needs scale is more evidence of validity given that the intent of the scale is to assess the caregiver's ability to provide for the child's material needs, such as food, shelter, and clothing.

There was a significantly positive relationship between the CAFAS and the independent ratings made by parents, teachers, and the youth themselves regarding specific problems. Youth with higher CAFAS total scores were much more likely to have poor social relationships, to have difficulties in school, and to get into trouble with the law. Furthermore, the findings were observed across spheres of functioning. Interpersonal problems were reported with peers, family members, and teachers at school. School problems ranged across the youth's attitudes, behavior, and academic achievement. The strongest and most consistent findings across time were those indicating involvement with the juvenile justice system. When the parent was the informant, these significant relationships were typically observed at each of the four waves. These findings suggest that the CAFAS is a useful indicator of impaired functioning both at intake and at follow-up.

The repeated measures analyses indicated that youth who were in treatment programs differing in intensity of services (i.e., inpatient/RTC, alternative care [e.g., day treatment], outpatient) at intake had significantly different CAFAS total scores not only at Wave 1 (i.e., intake), but also at 6 and 12 months post intake (i.e., Waves 2 and 3). The youth who were inpatients or in RTCs at intake were the most impaired at all three waves, whereas the outpatients at intake were the least impaired at all three waves. By 18 months, the three groups had similar means which were all in the low range of impairment. There was a significant effect for time, indicating that the impairment scores dropped from Wave 1 through Wave 4.

While it would be tempting to assume that the main effect for time in the repeated measures analyses and the lack of main effect for initial level of care at Wave 4 in the one way analysis of variance were the result of treatment, the design of the Evaluation Study does not permit making this assumption. The Evaluation Study did not have a placebo or a no treatment control group. However, most clinical settings would be pleased to report to their governing boards the clinical findings evident from Table 10. Youth starting out at intake as psychiatric inpatients or in RTCs had impairment levels at 18 months which were considerably lower than the typical outpatient seen at intake. Impairment scores for these youth starting off in residential care were reduced by one and a half standard deviations. Similar findings were also seen for children served in the alternative treatments.

GENERAL DISCUSSION

The reliability and validity of the CAFAS were demonstrated by the two studies. These findings, along with the data on predictive validity reported by Hodges and Wong (1996), present a strong case for the CAFAS. When compared to the state-of-the-art measures used in the Evaluation Study, the CAFAS was the best predictor of subsequent service utilization and associated costs.

The CAFAS can be useful for a variety of administrative purposes, including linking level of care to level of need (Newman, Griffin, Black, & Page, 1989; Newman & Hodges, 1995), program evaluation and planning (Herman & Mowbray, 1991; Pokorny, 1991), conducting client oriented cost outcome studies (Newman, Burwell, & Underhill, 1978), and providing information to consumers, consistent with the notion of developing provider "report cards" (Freeman & Trabin, 1994). The CAFAS is also useful for clinical decision-making about individual consumers. Its focus on behavioral descriptions of impairment and its depiction of the youth's functioning across domains make it a useful organizing tool for treatment plan discussions among staff and with parents. For any measure of impairment to be feasible for administrative purposes, it will likely need to be multidimensional, require behavioral anchors, and be clinically relevant.

ACKNOWLEDGMENTS

This research was supported in part by the U. S. Army Health Services Command (DADA 10-89-C-0013) as a subcontract from the North Carolina Department of Human Resources, Division of Mental Health, Developmental Disabilities, and Substance Abuse Services; and a grant from the National Institute of Mental Health (RO1MH-46136-01), on which Dr. Leonard Bickman was the principal investigator and the first author was a co-investigator. Dr. Lenore Behar was also a co-investigator on the NIMH grant and generously provided guidance and support throughout this research.

The authors would like to acknowledge John Knapp, Thomas Summerfelt, and Warren Lambert for their technical assistance; Jean Gust for her critical review of the manuscript; Sarah Ring-Kurtz who assisted in training; Judy Hovey, Sherry Whalen, Anne Kramer, and Mark Latessa who supported our research in Michigan, and the following students who served as research assistants: Pat Battey, Kyle Bonesteel, Jean Gust, Kristina Kalaj, Brian Libby, Lance Morgan, Julie Peregord, Mary Roberson, Michelle Va-

lentino, and Lisa Vengroff. In addition, we are grateful to all of the volunteers who served as raters and interviewees.

REFERENCES

- Achenbach, T. (1991). *Manual checklist/4-18 and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Almeida, D. M., & Kessler, R. C. (1994). *Daily life events as mediators of the gender-distress relationship*. Unpublished manuscript, Ann Arbor, MI: University of Michigan.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Bickman, L., Guthrie, P., Foster, E. M., Lambert, E. W., Summerfelt, W. T., Breda, C. S., & Heflinger, C. A. (1994). *Final report of the outcome and cost/utilization studies of the Fort Bragg Evaluation Project*. Unpublished manuscript, Nashville, TN: Vanderbilt University Center for Mental Health Policy.
- Bird, H., Canino, G., Rubio-Stipec, M., & Ribera, J. C. (1987). Further measures of the psychometric properties of the Children's Global Assessment Scale (CGAS). *Archives of General Psychiatry*, *44*, 821-824.
- Brannan, A. M., Heflinger, C. A., & Bickman, L. (1995). *The Burden of Care Questionnaire: Measuring the impact on the family of living with a child with serious emotional problems*. Unpublished manuscript, Nashville, TN: Center for Mental Health Policy, Vanderbilt University.
- Breda, C. S. (1996). Methodological issues in evaluating mental health outcomes of a children's mental health managed care demonstration. *The Journal of Mental Health Administration*, *23*, 40-50.
- Edelbrock, C., & Achenbach, T. M. (1984). The teacher version of the Child Behavior Profile: I. Boys aged 6-11. *Journal of Consulting and Clinical Psychology*, *52*, 207-217.
- Edelbrock, C., Costello, A. J., Dulcan, M. K., Conover, N. C., & Kalas, R. (1986). Parent-child agreement on child psychiatric symptoms assessed via structured interview. *Journal of Child Psychology and Psychiatry*, *27*, 181-190.
- Eme, R. F., & Kavanaugh, L. (1995). Sex differences in conduct disorder. *Journal of Clinical Child Psychology*, *24*, 406-426.
- Federal Register*. Substance Abuse and Mental Health Services Administration: Center for Mental Health Services. Washington, DC: U.S. Department of Health & Human Services, Office of Human Development Services, May 20, 1993.
- Freeman, M. A., & Trabin, T. (1994). *Managed behavioral healthcare: History, models, key issues, and future course*. Report prepared for the U.S. Center for Mental Health Services. Rockville, MD: Department of Health & Human Services, Substance Abuse and Mental Health Services Administration.
- Garber, J., Kriss, M. R., Koch, M., & Lindholm, L. (1988). Recurrent depression in adolescents: A follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *27*, 49-54.
- Herman, S. E., & Mowbray, C. T. (1991). Client typology based on functioning level assessments: Utility for service planning and monitoring. *The Journal of Mental Health Administration*, *18*, 101-115.
- Hodges, K. (1989). *Child and Adolescent Functional Assessment Scale*. Unpublished manuscript, Ypsilanti, MI: Eastern Michigan University.
- Hodges K. (1990a). *Child and Adolescent Functional Assessment Scale Self-Training Manual*. Unpublished manuscript, Ypsilanti, MI: Eastern Michigan University.

- Hodges, K. (1990b). *Child Assessment Schedule*. Unpublished manuscript (3rd ed.), Ypsilanti, MI: Eastern Michigan University.
- Hodges, K. (1990c). *Child Assessment Schedule-Parent Form*. Unpublished manuscript (3rd ed.), Ypsilanti, MI: Eastern Michigan University.
- Hodges, K. (1993). Structured interviews for assessing children. *Journal of Child Psychology and Psychiatry*, 34, 49-68.
- Hodges, K. (1994). *Measures for assessing impairment in children and adolescents*. Paper prepared for the U.S. Center for Mental Health Services. Rockville, MD: Department of Health & Human Services, Substance Abuse, and Mental Health Services Administration.
- Hodges, K., Bickman, L., Kurtz, S., & Reiter, M. (1991). A multi-dimensional measure of level of functioning for children and adolescents. In A. Algarin & R. M. Friedman (Eds.), *Proceedings of the fourth annual research conference on a system of care for children's mental health* (pp. 149-154). Tampa, FL: Research and Training Center for Children's Mental Health, University of South Florida.
- Hodges, K., Gordon, Y., & Lennon, M. (1990). Parent-child agreement on symptoms assessed via a clinical research interview for children: The Child Assessment Schedule (CAS). *Journal of Child Psychology and Psychiatry*, 31, 427-436.
- Hodges, K., & Gust, J. (1995). Measures of impairment for children and adolescents. *Journal of Mental Health Administration*, 22, 403-413.
- Hodges, K., Kline, J., Stern, L., Cytryn, L., & McKnew, D. (1982). The development of a child assessment interview for research and clinical use. *Journal of Abnormal Child Psychology*, 10, 173-189.
- Hodges, K., & Wong, M. M. (1996). *Use of the Child and Adolescent Functional Assessment Scale to predict service utilization and cost*. Manuscript submitted for publication.
- Mirowsky, J., & Ross, C. E. (1989). *Social causes of psychological distress*. New York: Aldine de Gruyter.
- Newman, F. L., Burwell, B. A., & Underhill, W. R. (1978). Program analysis using the client oriented cost outcome system. *Journal of Evaluation and Program Planning*, 1, 19-30.
- Newman, F. L., Griffin, B. P., Black, R. W., & Page, S. E. (1989). Linking level of care to level of need: Assessing the need for mental health care for nursing home residents. *American Psychologist*, 44, 1315-1324.
- Newman, F. L., & Hodges, K. (1995). *Developing outcome assessment measures for children, adolescents, and adults to support a provider profile in Indiana*. Paper presented at the meeting of the Second Annual Florida Conference on Behavioral Healthcare Evaluation, Orlando, FL.
- Pokorny, L. J. (1991). A summary measure of client level of functioning: Progress and challenges for use within mental health agencies. *The Journal of Mental Health Administration*, 18, 80-87.
- Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., & Aluwahlia, S. (1983). A Children's Global Assessment Scale (CGAS). *Archives of General Psychiatry*, 40, 1228-1231.
- Spitzer, R. L., Fleiss, J. L., & Endicott, J. (1978). Problems of classification: Reliability and validity. In M. A. Lipton, A. DiMascio, & K. F. Killam (Eds.), *Psychopharmacology: A generation of progress* (pp. 857-869). New York: Raven Press.
- Steinhausen, H. (1987). Global assessment of child psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry*, 26, 203-206.