Differences in Levels of Functional Impairment and Rates of Serious Emotional Disturbance Between Youth With Internalizing and Externalizing Disorders When Using the CAFAS or GAF to Assess Functional Impairment

Sarah E. Francis1, Chad Ebesutani2, and Bruce F. Chorpita3

Abstract
This investigation examined how two commonly employed measures of functional impairment (the Child and Adolescent Functional Assessment Scale [CAFAS] and the Global Assessment of Functioning [GAF] scale) perform with respect to assessing levels of functional impairment and, hence, identifying rates of serious emotional disturbance (SED) across youth diagnosed with internalizing or externalizing disorders. GAF scores indicated comparable levels of functional impairment between groups. CAFAS scores suggested significantly higher levels of functional impairment for youth diagnosed with externalizing disorders. The CAFAS also identified significantly higher rates of SED among youth with externalizing than those with internalizing disorders, whereas the GAF identified comparable rates of SED between groups. These findings suggest caution when relying on a single measure to assess functional impairment and to identify SED, particularly when these measures are used in decisions concerning service allocation. Implications of these findings are addressed, including the need for further research to identify the optimal assessment strategy to assess functional impairment and identify SED among youth diagnosed with internalizing or externalizing disorders.

Keywords
functional impairment, serious emotional disturbance, assessment, internalizing disorders, externalizing disorders

Functional impairment, broadly defined as “specific deficits in multiple domains of functioning developing subsequent to a disorder” (Winters, Collett, & Myers, 2005, p. 309), has become an increasingly important construct in the field of child mental health services and research, as is evident from the ways in which functional impairment is used (a) to define “caseness,” (b) to predict treatment needs and outcomes, and (c) as a federal agency criterion for service allocation (Canino, Costello, & Angold, 1999; Winters et al., 2005). Increased interest in assessing and documenting functional impairment has been attributed to (a) changes in the Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition (DSM-IV; American Psychiatric Association [APA], 2000) diagnostic criteria that necessitate functional impairment to obtain a diagnosis, (b) a demand for assessing changes in functioning when assessing treatment outcomes, and (c) the idea that symptom amelioration is not synonymous with improvements in daily functioning (see Winters et al., 2005, for a complete discussion of these points).

Indeed, clinical scientists and mental health services researchers routinely incorporate measures of functional impairment into treatment outcome and utilization studies, and improvements in “functioning” are quickly becoming the accepted standard by which treatment success is gauged (Canino et al., 1999; Winters et al., 2005). A key benefit of such measures is that they are intended to represent a
standard metric across many types of children in the mental health system, rather than varying across different diagnostic groups, as seen with many symptom measures (Bates, 2001). In addition, as the articulated goal of mental health services is to minimize functional impairment (rather than to eliminate a diagnosis), it is fitting that measures of functional impairment play an increasingly central role in decisions pertaining to service allocation, use, and outcome.

Children and adolescents deemed to be in need of mental health services are typically identified on the basis of two criteria: (a) the presence of a mental health diagnosis, and (b) evidence of clinically significant functional impairment. The co-occurrence of these two criteria is required for a child to be identified with serious emotional disturbance (SED), and it is this label that is often required at both federal and state levels for a child to receive community-based mental health services (Costello et al., 1996). Functional impairment is distinct from the specific symptoms experienced by the individual, such that deficits in functioning (i.e., functional impairment) are not to be confounded with symptomatology or severity of the disorder (see Winters et al., 2005, for a discussion of this distinction). However, it is the interplay of these two constructs, symptoms and functional impairment, that is used to determine SED and, hence, service allocation. As such, SED is composed in part of diagnosis but also requires impairments in an individual’s day-to-day functioning.

A wide range of assessment measures has been developed to assess and document functional impairment (see Winters et al., 2005, for a comprehensive review). The availability of numerous measures of functional impairment requires that decisions be made with respect to which measure to employ given the population, the assessment purpose, and the intended use of the rating (Bird & Gould, 1995). Regardless of the specific measure employed, however, it should be free of systematic bias, particularly given the effect that a functional impairment score may have on the determination of youth SED status, which in turn informs the provision of mental health services. Indeed, “good” measures of functional impairment should perform consistently across similar groups of children (e.g., behaviorally and emotionally disturbed youth, boys and girls), and scores on such measures should not vary solely as a function of diagnosis (Bird & Gould, 1995; Canino et al., 1999). More specifically, functional impairment is a construct distinct from symptomatology (e.g., Bird, Canino, Rubio-Stipec, & Ribera, 1987; Pickles et al., 2001; Simonoff et al., 1997; Winters et al., 2005), and as such, measures of this construct should not assess symptom severity or vary systematically in relation to specific diagnoses. Similarly, certain disorders are more prevalent among boys versus girls (e.g., behavioral disorders are more common among boys). In attempting to ascertain whether a given measure of functional impairment operates consistently across diagnostic profiles, a related question concerns whether such a measure also performs consistently across genders. For example, if a measure of functional impairment performs differently across diagnostic groups, it would be important to determine whether this difference is actually attributable to diagnostic status or whether it is instead due to gender. Significant diagnostic group differences that persist when evaluated within each gender group yield greater confidence that the actual differences are associated with diagnosis and that diagnosis is not acting as a proxy variable for gender. To date, gender differences with respect to functional impairment scores have not been reported for either the CAFAS (e.g., Ezpeleta, Granero, de la Osa, Domenech, & Bonillo, 2006; Goldston et al., 2007; Hodges & Wong, 1996) or the GAF (Manassis & Hood, 1998).

The aim of the present investigation was to examine how two commonly employed measures of functional impairment, namely, the Child and Adolescent Functional Assessment Scale (CAFAS) and the Global Assessment of Functioning (GAF) scale, perform with respect to identifying rates of SED across youth diagnosed with internalizing disorders and youth diagnosed with externalizing disorders. These two measures were selected for examination given the widespread use of both the CAFAS (e.g., Hodges, Wong, & Latessa, 1998; Resendez, Quist, & Matshazi, 2000) and the GAF (e.g., Asarnow, Aoki, & Elson, 1996; Yeh et al., 2002) with respect to assessing functional impairment, particularly in determinations of SED status and subsequent decisions concerning the receipt of mental health services. The use of these two measures has become standard practice for individuals working with child and adolescent populations. Specifically, the GAF is the most widely used unidimensional assessment measure of functional status (Blake, Cangelosi, Johnson-Brooks, & Belcher, 2007) and the CAFAS, as a more specific and complex multidimensional measure, has been widely employed on a state level to assess youth impairment (Bates, 2001). As such, these two measures represent the two types of scales currently available to assess functional impairment. Accordingly, this investigation sought to examine how these two measures perform with respect to identifying SED status across internalizing and externalizing diagnoses to better inform the use of these measures. Each of these measures will be reviewed briefly below.

The multidimensional Child and Adolescent Functional Assessment Scale (Hodges & Wong, 1996) is a widely used measure of functional impairment across several settings, including community mental health settings (Bates, 2001; Winters et al., 2005). The CAFAS consists of multiple scales designed to identify impairment specific to each of eight domains related to different areas of child functioning. Several strengths of the CAFAS have been noted, including that the total scores on this instrument have been
found to differentiate between the presence and absence of psychopathology among children and adolescents (Ezpeleta et al., 2006) and that the CAFAS total score at intake has been found to predict subsequent service use at both 6 and 12 months post-intake (Hodges & Wong, 1997). Scores on the CAFAS also distinguish between children who do and do not receive services in education/mental health programs (Rosenblatt & Rosenblatt, 1999). Importantly the CAFAS has been found to perform similarly across males and females with respect to measuring overall functioning (e.g., Ezpeleta et al., 2006; Goldston et al., 2007).

Data generated from two large-scale evaluation studies have indicated that the CAFAS demonstrates good internal consistency (ranging from .63 to .78) and high interrater reliability (above .92 for the total CAFAS score, and above .83 for the individual scales; Hodges & Wong, 1996). The content, concurrent, and predictive validity of the CAFAS have also been examined, suggesting that the CAFAS correlates significantly and positively with other indicators of impairment, including severity of psychiatric diagnosis and subsequent service utilization (Hodges & Wong, 1996). Furthermore, the CAFAS has been shown to serve as a good predictor of service utilization and treatment cost (Hodges & Wong, 1996).

In contrast to the multidimensional CAFAS, the GAF (APA, 2000) uses a unidimensional approach to measuring functional impairment in that it provides a single score describing the overall functioning of the individual across psychological, social, and occupational (school) domains. The relatively simple structure and unidimensional nature of the GAF allow for raters to provide a reliable and valid index of symptoms and functioning to summarize the individual’s clinical presentation (Goldman, Skodol, & Lave, 1992; Jones, Thornicroft, Coffey, & Dunn, 1995). The GAF has been employed in multiple studies with youth as a measure of functional status (e.g., Asarnow et al., 1996; Blake et al., 2007; Evans, Radunovich, Cornette, Wiens, & Roy, 2008; Gold, Shera, & Clarkson, 1993; Manassis & Hood, 1998; Rey, Starling, Weaver, Dossetor, & Plapp, 1995; Vyas, Hadjulis, Vourdas, Byrne, & Frangou, 2007; Yeh et al., 2002).

A psychometric investigation of the GAF suggested moderate internal consistency (ranging from .62 to .90; Hall, 1995). Concurrent validity of GAF scores has also been suggested by its correlations with other measures of support needs (Jones et al., 1995) and severity of depression (Hall, 1995). Investigations of the reliability of the GAF among child clinical populations have indicated moderate levels of interrater agreement among child outpatients (r = .54; Rey et al., 1995) and among youth referred to an anxiety disorder clinic (r = .72; Manassis & Hood, 1998). Good interrater reliability estimates have been suggested for the GAF for both current (.74) and past 6 month (.73) ratings (Bird et al., 1987). Scores on the GAF have been reported to be consistent across gender (Manassis & Hood, 1998).

Although the CAFAS is designed exclusively for use with child and adolescent populations, the GAF is employed with both children and adults. Both scales have a range of applications including in mental health services research, by health care administrative organizations, and in the course of in- and out-patient assessment and service delivery. Both scales are used with individuals with various mental health diagnoses (e.g., Bryson, Corrigan, McDonald, & Holmes, 2008; Fields & Ogles, 2002; Walrath, Nickerson, Crowel, & Leaf, 1998; Yeh et al., 2002). Both scales are also employed on a state level to assist in determining SED status among youth (e.g., Bates, 2001; Fields & Ogles, 2002). Given that neither has yet been deemed the “gold standard” for measuring functional impairment, decisions to use either or both of these scales depend largely on the assessment and/or research question of interest as well as whether agencies have instituted mandated assessments that include either of these measures of functional impairment.

As with any measure of functional impairment, specific concerns with the CAFAS and GAF exist that have been detailed elsewhere (Bates, 2001; Bates, Furlong, & Green, 2006; Winters et al., 2005). For example, it has been suggested that unidimensional measures of functional impairment might equate functioning with symptomatology (Winters et al., 2005). In addition, although such measures can be completed quickly, they cannot provide estimates of impairment across multiple domains of functioning. Moreover, given the moderate reliability estimates (ranging from .54 to .72) yielded when using the GAF to assess functional impairment with children and adolescents, it has been suggested that specialized, multidimensional measures might be more appropriate for this population (Rey et al., 1995). However, although multidimensional measures of functional impairment allow for the assessment of functioning across multiple domains and the ascertainment of specific areas in need of clinical attention, the CAFAS has been criticized for including items that assess both functioning and symptoms, as well as for employing a scoring structure that requires endorsement of only the most severe item in a given domain (Winters et al., 2005).

Previous research has thus outlined the limitations associated with the GAF and the CAFAS. Of relevance to the present study, however, is how these two measures perform with respect to measuring levels of functional impairment (and hence identifying rates of SED) across two groups of youth: those diagnosed with internalizing disorders and those diagnosed with externalizing disorders. This question is of particular interest to the present study given the role of SED status in youths’ access to and receipt of mental health services. Specifically, given that in multiple jurisdictions, a minimum level of functional impairment (in addition to a
mental health diagnosis) is required for youth to be eligible for services, the way in which measures of functional impairment perform across these diagnostic groups is of concern not only at the individual level but at a systems level as well.

Of additional interest to the present study are the structural and scoring differences apparent between the CAFAS and the GAF, as representative of multidimensional and unidimensional, respectively, measures of functional impairment. Specifically, the GAF does not allow for determinations of functioning within specific domains but instead provides anchors that refer to functioning across symptomological, social, occupational, school, and interpersonal relationship domains. In contrast, the CAFAS is composed of eight scales, each of which corresponds to a particular domain of functioning. One artifact of this structural difference is the content assessed by each of the various CAFAS scales. Specifically, visual inspection of these scales suggests that proportionally more scales are designed to assess functional status relevant to externalizing problems than internalizing problems. Indeed, results of an exploratory factor analysis of the CAFAS among a clinically referred sample of children and adolescents suggested that five of the subscales assess content related to externalizing behaviors (the School, Home Community, Behavior Toward Others, and Substance Use sub-scales), whereas only three subscales assess content related to internalizing behaviors (the Moods and Emotions, Self-Harm, and Thinking subscales; Ebesutani, Francis, & Chorpita, 2008).

The primary aim of the current investigation was to examine how total GAF and CAFAS scores each measure levels of functional impairment and, hence, identify rates of SED among youth diagnosed with internalizing disorders and youth diagnosed with externalizing disorders. We reasoned that an enhanced understanding of how these measures perform with respect to SED identification among youth will serve to inform mental health services research specific to youth access to and receipt of services. Rates of SED status when based on the GAF and CAFAS were compared for each of the two groups (i.e., youth diagnosed with internalizing and externalizing disorders). To determine diagnostic group membership, we used a semi-structured clinical interview, the Anxiety Disorders Interview Schedule for DSM-IV, Child and Parent Versions (ADIS-IV-C/P; Silverman & Albano, 1996). The ADIS-IV-C/P was used as the diagnostic interview within the context of a larger study of anxiety symptoms among youth referred for mental health treatment. Specific hypotheses were as follows: (a) the GAF, given its unidimensional structure, will reveal comparable levels of functional impairment across both youth with internalizing disorders and youth with externalizing disorders; (b) the CAFAS, given a disproportionate number of scales assessing externalizing as compared to internalizing problems, will indicate significantly greater levels of functional impairment for youth with externalizing disorders than for youth with internalizing disorders; (c) given the absence of gender effects on functional impairment noted in previous studies (e.g., Ezpeleta et al., 2006; Goldston et al., 2007; Hodges & Wong, 1996; Manassis & Hood, 1998), levels of functional impairment, as measured by the CAFAS and GAF, will not vary by gender such that gender will not account for findings observed with the CAFAS and the GAF with respect to identifying youth as SED; (d) using the presence of a diagnosis and functional impairment cutoff scores to identify cases of SED, the CAFAS will identify significantly more youth diagnosed with externalizing disorders than youth diagnosed with internalizing disorders; and (e) these same differences in SED identification will not be observed when using the GAF to measure functional impairment.

Method

Participants

Participants in the present study included an ethnically diverse sample of 617 children and adolescents (ages 5 to 17) referred for a mental health evaluation at the Center for Cognitive Behavior Therapy (CCBT) at the University of Hawai‘i at Manoa. The majority of these children were referred by school personnel, although a subset of the sample was referred by primary caregivers. Of the total sample, 68.4% were boys (n = 422) and 43.3% were between the ages of 5 and 12 (n = 267). The median annual income was $18,500 (SD = $66,175). Of the present sample, 17.0% identified primarily with an Asian ethnicity (i.e., Chinese, Filipino, Japanese, and Korean), 6.8% with Pacific Islander (i.e., Hawaiian and Samoan), 12.5% with Caucasian, 2.3% with Black, 0.3% with Hispanic, and 40.2% reported a mixed ethnic background; 6.2% reported an “other” ethnic identity and the remaining 14.7% of the sample did not provide information with respect to ethnicity.

Participant diagnoses. Following a comprehensive mental health evaluation, each participant was assigned a diagnostic profile by the intake assessor. Youth in the present sample were then divided into two groups: (a) those with a primary diagnosis of an internalizing disorder (see Table 1) in the absence of an externalizing disorder (30.3%, n = 187), and (b) those with a primary diagnosis of an externalizing disorder (see Table 1) in the absence of an internalizing disorder (69.7%, n = 430). Youth with co-occurring internalizing and externalizing disorders were excluded from analyses. Among those children classified as having one or more internalizing disorders in the present sample, the age range was from 5 to 17, 38.0% of the sample were...
between the ages of 5 and 12 ($n = 71$), and 45.5% were boys ($n = 85$). Among those children receiving a diagnosis of one or more externalizing disorders, the age range was from 5 to 17, 45.6% were between the ages of 5 and 12 ($n = 196$), and 78.4% were boys ($n = 337$).

**Measures**

*Child and Adolescent Functional Assessment Scale.* The CAFAS (Hodges, 1994) is a clinician-reported, multidimensional rating scale designed to assess functional impairment as experienced by children and adolescents, ages 5 to 17, across eight domains of functioning (Hodges & Wong, 1996, 1997). Raters are provided with a list of behavioral descriptors for each of the subscales, from which they must choose those items that are most congruent with the youth’s most severe level of dysfunction during the month preceding the assessment (Hodges & Wong, 1996). Items within each subscale are grouped according to four degrees of impairment: severe, moderate, mild, and minimal or no impairment, respectively: “expelled or equivalent from school;” “frequently truant (i.e., approximately once every 2 weeks or for several consecutive days);” “occasionally disobeys school rules, with no harm to others or to property, more than other youth;” and “reasonably comfortable and competent in relevant roles.” Total scores on the CAFAS, derived by summing across the eight scales, can range from 0 to 240. In the present investigation, the CAFAS total score was significantly positively correlated with the total number of psychosocial and environmental problems reported on Axis IV of the DSM ($r = .39$, $p < .01$), suggesting that CAFAS scores were providing an estimate of psychosocial functioning.

In the present investigation, the CAFAS was completed by a doctoral student in clinical psychology (working in the Center in the capacity of an intake assessor) who had been trained to reliability using this instrument. Each assessor had completed the training procedures in the CAFAS self-training manual (Hodges, 1994). Specifically, CAFAS training involved reading instructions for scoring the CAFAS and subsequently completing a series of 10 sample vignettes (for which the correct scoring was supplied) and 10 test vignettes. Completing the test vignettes with at least

---

**Table 1. Number of Youth With Principal Internalizing and Externalizing Diagnoses and the Percentage of the Total Sample Diagnosed With Each Disorder**

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>$n^a$</th>
<th>$%^b$</th>
<th>Diagnosis</th>
<th>$n^a$</th>
<th>$%^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social phobia</td>
<td>56</td>
<td>9.1</td>
<td>ADHD</td>
<td>156</td>
<td>25.3</td>
</tr>
<tr>
<td>Generalized anxiety disorder</td>
<td>24</td>
<td>3.9</td>
<td>Conduct disorder</td>
<td>137</td>
<td>22.2</td>
</tr>
<tr>
<td>Separation anxiety disorder</td>
<td>15</td>
<td>2.4</td>
<td>Oppositional defiant disorder</td>
<td>128</td>
<td>20.7</td>
</tr>
<tr>
<td>Posttraumatic stress disorder</td>
<td>13</td>
<td>2.1</td>
<td>DBD NOS</td>
<td>38</td>
<td>6.2</td>
</tr>
<tr>
<td>OCD</td>
<td>12</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific phobia</td>
<td>9</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety disorder NOS</td>
<td>9</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panic disorder with agoraphobia</td>
<td>6</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agoraphobia</td>
<td>1</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Depressive disorders</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDD, single episode</td>
<td>14</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dysthmic disorder</td>
<td>11</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDD, recurrent</td>
<td>9</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressive disorder NOS</td>
<td>5</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Adjustment disorders</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjustment disorder with anxiety</td>
<td>7</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjustment disorder with mixed anxiety and depressed mood</td>
<td>7</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjustment disorder with depressed mood</td>
<td>5</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: ADHD = attention-deficit/hyperactivity disorder; NOS = not otherwise specified; DBD = disruptive behavior disorder; OCD = obsessive compulsive disorders; MDD = major depressive disorder.

1. The number of participants receiving each diagnosis does not add up to the total number of study participants given that co-principal diagnoses are tabulated here.

2. The percentage of participants in each diagnostic category does not add up to 100 given that co-principal diagnoses are tabulated here.
85% accuracy was required to be considered a reliable CAFAS rater.

**Global Assessment of Functioning.** The GAF scale (APA, 2000) is a clinician-derived measure of an individual’s psychological, social, and occupational (school) functioning. Clinicians assign each individual a score ranging from 1 to 100, designed to indicate the nature and current severity of the individual’s present difficulties, with lower scores indicating greater degrees of functional impairment (Jones et al., 1995). Scores are assigned using a set of 10 descriptive anchors that range from “persistent danger of severely hurting self or others” to “superior functioning in a wide range of activities,” corresponding to the 1st and 10th deciles, respectively. A sample anchor from the GAF, corresponding to a score of 61 to 70, reads, “Some mild symptoms (e.g., depressed mood and mild insomnia) OR some difficulty in social, occupational, or school functioning (e.g., occasional truancy or theft within the household), but generally functioning pretty well, has some meaningful interpersonal relationships.” In the current investigation, GAF scores were reflective of the individual’s functioning at the time of the assessment. Among the present sample, the GAF was significantly negatively correlated with the CAFAS total score ($r = -0.61$), such that lower GAF scores were associated with higher CAFAS scores, each of which indicate greater levels of functional impairment. The GAF was also significantly negatively correlated with each of the eight subscales (sub-scale correlations ranged from $-0.19$ for the Moods and Emotions scale to $-0.39$ for the Behavior Toward Others scale), suggesting that it measures a construct distinct from that assessed by the CAFAS. Moreover, similar to findings reported with the CAFAS, GAF scores were also significantly negatively correlated with the total number of psychosocial and environmental problems reported on Axis IV of the DSM ($r = -0.29$, $p < .01$), supporting the notion that the GAF provided estimates of functional impairment among the present sample.

**CAFAS and GAF cutoff scores for classifying youth as SED.** Because the developers of the CAFAS recommended a total cutoff score of 80 as an indicator of the level of functional impairment warranting classification as SED (Hodges & Wong, 1996), this same cutoff score was employed in this investigation. In addition, state guidelines for determining eligibility for mental health services in Hawaii, where this study was conducted, require a total CAFAS score of 80 (in addition to the presence of an Axis I diagnosis) for classification as SED. Accordingly, youth in this sample receiving total CAFAS scores ranging from 0 to 70 were identified as not being functionally impaired, whereas youth receiving scores ranging from 80 to 240 were identified as being functionally impaired. Consistent with previous investigations, the total CAFAS score was derived by summing all eight CAFAS scales.

Given that GAF cutoff scores for determining SED status among youth could not be located in the literature, functional impairment studies with the *Children's Global Assessment Scale* (CGAS; Shaffer et al., 1983) were used as a guide to determine appropriate GAF cutoff scores for determining SED status. The CGAS is the child version of the GAF and closely parallels the GAF (Blake et al., 2007) with respect to content (e.g., common wording for the anchors), structure (identical numerical anchors), and scoring procedures. Indeed, correlations between the CGAS and GAF (used to rate children) have been reported to be sufficiently high so as to make the use of both measures redundant ($r = .95$; Gold et al., 1993). Although several previous investigations of functional impairment among youth have employed the CGAS, a recent study has suggested that the GAF is a slightly more reliable measure of functioning among youth (Blake et al., 2007); for these reasons, the GAF was employed in the present investigation. Given that a cutoff score of 60 on the CGAS is widely used to determine functional impairment and to distinguish cases from non-cases (Bird et al., 1990; Canino et al., 1999), the present study designated impaired functioning as a score of 60 and below on the GAF. As such, youth receiving scores ranging from 61 to 100 on the GAF were identified as not being functionally impaired, whereas youth receiving scores ranging from 1 to 60 were identified as being functionally impaired.

**Anxiety Disorders Interview Schedule for DSM-IV, Child and Parent Versions.** The ADIS-IV-C/P (Silverman & Albano, 1996), a revision of the ADIS-C/P (Silverman & Nelles, 1988), which was a downward extension of the adult ADIS (DiNardo, O’Brien, Barlow, Waddell, & Blanchard, 1983), is a semistructured diagnostic interview designed for the assessment of DSM-IV diagnoses of childhood anxiety, mood, behavioral, and attentional disturbances. Excellent interrater reliability has previously been suggested for the ADIS-IV-C/P ($\kappa = .92$ for principal diagnoses; Lyneham, Abbott, & Rapee, 2007), as well as acceptable to excellent 7- to 14-day test–retest reliability estimates for child ($\kappa = .61–.80$), parent ($\kappa = .65–1.00$), and combined diagnoses ($\kappa = .63–1.00$) (Silverman, Saavedra, & Pina, 2001). The anxiety disorders portions of the ADIS-IV-C/P have demonstrated concurrent validity (Wood, Piacentini, Bergman, McCracken, & Barrios, 2002) with respect to convergence with the *Multidimensional Anxiety Scale for Children* (MASC; March, 1998). Evidence of sensitivity to treatment effects has also been demonstrated for the ADIS-IV-C in studies of children with anxiety disorders (e.g., Dadds, Heard, & Rapee, 1992; Kendall et al., 1997). The ADIS-IV-C/P is designed to yield diagnoses based on child report only, parent report only, and a combined consensus report. Child and parent diagnoses are assigned on the basis of the information gathered only during the course of the child or
parent interview, respectively. In the present study, consensus diagnoses were assigned based on (a) child report obtained during the child interview, (b) parent report obtained during the parent interview, (c) child self-report measures, (d) parent-report measures of child behavioral and emotional functioning, and (e) information obtained from relevant collateral sources (e.g., teachers). Additional standardized measures were not employed when obtaining information from teachers or other school professionals; rather, such individuals were interviewed by the intake assessor who made queries specific to each child’s referral question. For the purposes of the present investigation, diagnosis will refer to the consensus diagnosis, based on all of the information available to the clinician at the time of the assessment (see Note 1). In addition, all sections of the ADIS-IV-C/P were administered to each participant and parent. Given that the parent ADIS-IV contains sections pertaining to both internalizing and externalizing symptomatology, the data collected from this interview provided the clinician with the information necessary to complete the measures of functional impairment employed in this study.

The ADIS-IV-C/P was administered by a doctoral student in clinical psychology (working in the Center in the capacity of an intake assessor) who had been trained to reliability using this structured interview. Training in the administration of the ADIS-IV-C/P followed the standardized procedures as outlined by the authors (Silverman & Albano, 1996) and involved (a) observation of three ADIS-IV-C/P interviews conducted by trained assessors, (b) conducting a series of five ADIS-IV-C/P interviews while being observed by trained assessors, and (c) matching the experienced assessors on all clinical diagnoses in three of the five interviews.

Procedure

At the time of the initial visit, all participants and their legal guardians underwent standardized Institutional Review Board-approved notice of privacy and consent procedures prior to any data collection. The data reported here were collected as part of ongoing clinical services at a university-based stress and anxiety disorders clinic. Families presenting to the clinic were provided with a comprehensive mental health evaluation in which the assessor spent approximately 3 hours with the child and parent. Families were neither charged nor received reimbursement for their participation in the evaluation. All participants and their primary caregiver(s) were administered the ADIS-IV-C/P. For any participant younger than age 7, as well as for any youth who experienced difficulty understanding questions presented on the ADIS-IV, interview questions were reworded on a case-by-case basis to meet the idiosyncratic needs of each participant (which is consistent with the use of the ADIS as a semistructured interview) and to facilitate the child’s understanding of the questions posited. To ascertain the child’s functional capabilities specific to the domains of school and peer relationships, we also sought information from teachers and guidance counselors, both through telephone interviews and paper-and-pencil measures received through the mail. On the basis of all information obtained during the assessment process, assessors completed the CAFAS and assigned a GAF score to each participant; CAFAS and GAF ratings were made by the same intake assessor who had conducted the diagnostic interview. Given that the clinic in which this investigation was conducted had consistently used the CAFAS and GAF in all assessment cases, each intake clinician received extensive training in the use of each measure. Standard assessment protocol at the Center included assessors completing the CAFAS and GAF immediately upon collecting all pertinent assessment information from the child, parent, and collateral sources. Completed CAFAS forms and GAF ratings were also reviewed by supervisors in weekly assessment supervision meetings, ensuring that CAFAS forms and GAF ratings were completed at least within the week of the completion of the assessment. All intake therapists in the present investigation were blind to all of the hypotheses posited in the present investigation.

Results

To examine whether the GAF yielded scores that indicated comparable levels of functional impairment across youth diagnosed with internalizing disorders and youth diagnosed with externalizing disorders, we conducted a 2 × 2 analysis of variance (ANOVA) in which we evaluated the main effects of diagnostic group and gender as well as the interaction between diagnostic classification and gender. Results indicated no significant between-group differences, \( F(1, 613) = 0.33, p > .05, \eta^2 = .00 \), with respect to GAF scores for youth with internalizing diagnoses (\( M = 57.51, SD = 9.71 \)) and youth with externalizing disorders (\( M = 57.41, SD = 9.16 \)). We observed neither a significant main effect of gender, \( F(1, 613) = 2.14, p > .05, \eta^2 = .00 \), nor a significant interaction of gender and diagnostic classification, \( F(1, 613) = 0.01, p > .05, \eta^2 = .00 \). These results indicated that youth with internalizing diagnoses and youth with externalizing diagnoses exhibited comparable levels of functional impairment on the GAF.

To examine whether the CAFAS also yielded comparable levels of functional impairment across youth with internalizing diagnoses and youth with externalizing diagnoses, we conducted a 2 × 2 ANOVA in which we evaluated the main effects of diagnostic group and gender as well as the interaction between diagnostic classification and gender. Results indicated a significant main effect of diagnostic
classification, $F(1, 613) = 23.35, p < .01, \eta^2 = .04$, such that youth diagnosed with externalizing disorders scored higher on the CAFAS ($M = 66.95, SD = 35.99$) than youth diagnosed with internalizing disorders ($M = 53.10, SD = 30.43$). Consistent with hypotheses, we observed neither a significant main effect of gender, $F(1, 613) = 0.89, p > .05, \eta^2 = .00$, nor a significant interaction effect, $F(1, 613) = 0.97, p > .05, \eta^2 = .00$.

We next examined the percentage of youth with externalizing and internalizing diagnoses identified as functionally impaired on the CAFAS and GAF. Tables 2 and 3 show the numbers and percentages of youth scoring at each level of functioning on the CAFAS and GAF, respectively. First, we examined the percentages of the sample identified as exhibiting clinically significant functional impairment, as defined by predetermined cutoff scores of 80 on the CAFAS (Hodges & Wong, 1996) and 60 on the GAF (cf. Bird et al., 1990; Canino et al., 1999). Using a cutoff score of 80 on the CAFAS, 29.3% of the sample exhibited clinically significant levels of functional impairment. Using a cutoff score of 60 on the GAF, 64% of the full sample exhibited such impairment. These results suggested that markedly different rates of clinically significant functional impairment were identified by each measure.

Because of the possibility that different rates of clinically significant functional impairment could be artifacts of the selected cutoffs of 80 and 60 employed for the CAFAS and GAF, respectively, rather than due to the differential sensitivity of the measures, we attempted to control for this artifact by testing a second cutoff score on the GAF. We determined this cutoff by examining the distribution of GAF scores for this sample and selecting a cutoff score that would identify approximately 30% of the sample as SED to be consistent with the percentage of youth identified by the CAFAS. A cutoff score of 52 on the GAF was found to identify 29.3% of youth with significant impairment. As such, all subsequent analyses for the GAF will report findings using both a cutoff score of 60, which has been suggested by studies employing the CGAS, and a cutoff score of 52, which identifies the same percentage of youth as functionally impaired in this sample as the cutoff of 80 on the CAFAS.

To examine whether proportionately more youth with externalizing diagnoses were identified as SED on the CAFAS and GAF than youth with internalizing diagnoses, we used the cutoff scores mentioned above (i.e., GAF = 60, GAF = 52, CAFAS = 80) to classify youth for whom clinically significant functional impairment was present or absent. Given that all youth in the sample had received an Axis I DSM-IV-TR diagnosis, the identification of cases with clinically significant functional impairment “present” was tantamount to identifying cases as SED. We used the chi-square statistic to compare the percentages of youth with internalizing and externalizing diagnoses identified as SED using each functional impairment measure (and their respective cutoff scores). Consistent with hypotheses, using a cutoff score of 80 on the CAFAS to indicate SED resulted in the identification of proportionately more youth with externalizing diagnoses (34.2%, $n = 147$) than youth with internalizing diagnoses (18.2%, $n = 34$). These differences were significant, $\chi^2(1, n = 617) = 16.10, p < .01$. Moreover, these differences remained significant when examined among boys, $\chi^2(1, n = 617) = 5.38, p < .01$, and girls, $\chi^2(1, n = 617) = 11.94, p < .01$. Specifically, a cutoff score of 80 on the CAFAS identified 32.9% ($n = 111$) of boys with an externalizing disorder as clinically impaired (i.e., obtaining scores that exceeded the clinical cutoff) but only 20.0% ($n = 17$) of boys diagnosed with an internalizing disorder as clinically impaired. Similarly, a cutoff score of 80 on the CAFAS identified 38.7% ($n = 36$) of girls with an externalizing disorder as clinically impaired, but only 16.7% ($n = 17$) of girls with an internalizing disorder as clinically impaired. These findings suggest that the significant difference in rates of SED between youth with internalizing and youth with externalizing diagnoses was not a function of gender.

 Whereas the CAFAS led to a disproportionate percentage of youth with externalizing diagnoses meeting the cutoff for SED as compared to youth with internalizing diagnoses, the GAF did not result in a disproportionate number of youth meeting the cutoff for SED (based on GAF cutoff of 60) between externalizers (64%, $n = 275$) and internalizers (64.2%, $n = 120$), $\chi^2(1, n = 617) = 0.00, p > .05$. These differences remained nonsignificant among both boys, $\chi^2(1, n = 617) = 0.20, p > .05$, and girls, $\chi^2(1, n = 617) = 0.03, p > .05$. Specifically, whereas a cutoff score of 60 on the GAF identified 62.6% ($n = 211$) of boys with an externalizing disorder, 60% ($n = 51$) of boys with an internalizing disorder were identified as clinically impaired using this cutoff. Similarly, 68.8% ($n = 64$) of girls with an externalizing disorder were identified as clinically impaired using this GAF cutoff, and 67.6% ($n = 69$) of girls with an internalizing disorder were similarly identified. Moreover, the GAF also did not result in a disproportionate number of youth meeting the cutoff for SED (based on GAF cutoff of 52) between externalizers (29.3%, $n = 126$) and internalizers (29.4%, $n = 55$), $\chi^2(1, n = 617) = 0.00, p > .05$. Again, these differences remained nonsignificant among both boys, $\chi^2(1, n = 617) = 0.04, p > .05$, and girls, $\chi^2(1, n = 617) = 0.09, p > .05$. Specifically, a cutoff score of 52 on the GAF identified 28.2% ($n = 95$) of boys with externalizing disorders and 27.1% ($n = 23$) of boys with an internalizing disorder as clinically impaired; similarly, 33.3% ($n = 31$) of girls with an externalizing disorder and 31.4% ($n = 32$) of girls with an internalizing disorder were identified as clinically impaired using this GAF cutoff.
Discussion

This study reported on levels of functional impairment as well as rates of identified SED youth using two measures, each representative of the two types of functional impairment assessment tools, namely, the multidimensional CAFAS and the unidimensional GAF. Given the key role that SED status plays in service allocation decisions for youth, this study sought to examine whether comparable rates of SED would be identified amongst youth diagnosed with internalizing disorders and youth diagnosed with externalizing disorders using each of these instruments. This question was examined amongst a sample of clinically referred youth, all of whom had received a diagnosis of either an internalizing or externalizing disorder. Given that diagnosis is only one half of the criteria used to identify youth as SED—the other half being functional impairment—this investigation examined levels of functional impairment and rates of SED using the multidimensional CAFAS as well as those observed when using the unidimensional GAF score.

Mean GAF scores observed in this study were consistent with those reported elsewhere (Manassis & Hood, 1998; Rey et al., 1995). The proportion of youths identified as functionally impaired using specific GAF cutoff scores, however, has not yet been documented in previous studies on child clinical samples. Comparisons of these proportions based on the GAF across samples could thus not be evaluated. With respect to mean CAFAS scores found among other child clinical samples, normative data have not yet been reported for the CAFAS (Winters et al., 2005), making it difficult to compare the mean CAFAS values observed in this study to those obtained elsewhere. Although the existence of normative data from a research perspective would facilitate cross-sample comparisons of functional status, such data are perhaps not clinically necessary given the use of the CAFAS to determine the appropriate level of service in a specific setting. However, Hodges and Wong (1996) reported that a cutoff score of 80 identified approximately 20% of their sample as seriously impaired; amongst the present sample, 29.3% of cases received scores of 80 or above on the CAFAS, suggesting that the CAFAS eight-scale total summed score identified a rate of SED relatively comparable to rates previously reported.

It is notable, however, that markedly different rates of clinically significant functional impairment were identified using the CAFAS and the GAF when previously established cutoff scores were used to designate the presence of such impairment. Whereas approximately 30% of the present sample was identified as experiencing serious impairment...
when using a cutoff score of 80 on the CAFAS (a score identified by Hodges & Wong, 1996), more than 60% of the sample was identified as such when using a cutoff score of 60 on the GAF (a score identified by consulting research with the CGAS; e.g., Bird et al., 1990). It must be noted that there is not currently an established cutoff score for the GAF to identify SED status among youth. However, these results nonetheless suggest that these two measures of functional impairment performed quite differently within the same sample of diagnosed youth with respect to identifying clinically significant levels of functional impairment. Although such differences might speak to the distinct strengths and unique aspects of each of these measures, these findings also suggest caution when relying on a single measure of functional impairment to identify cases warranting mental health services.

The key finding of the present investigation is the markedly different rates of SED identification among youth with internalizing and externalizing disorders observed when using the CAFAS, and the concurrent absence of such differences when using the GAF, to assess functional impairment. Significantly more youth with externalizing disorders than internalizing disorders were identified as SED using a CAFAS total score (derived by summing scores obtained on the eight scales). A specific cutoff score designed to identify SED status with the GAF has not been identified in the literature, thus requiring consultation of research using the CGAS as a guide. However, no differences were found between the number of youths with externalizing and internalizing diagnoses identified as SED using the GAF (and the suggested cutoff score of 60, indicating moderate impairment). Even when using a GAF cutoff score of 52 (with the intent of identifying impairment among a similar percentage of youth as identified by the CAFAS), equivalent proportions of youth with internalizing and externalizing disorders were identified as impaired. This specific finding suggests that the differential rates of clinically significant functional impairment identified for these two diagnostic groups might not be an artifact of the cutoff scores employed, particularly given that equivalent proportions of youth with internalizing and externalizing diagnoses were identified at each impairment range of the GAF. Moreover, these findings remained consistent when evaluated by gender, suggesting that differential rates of clinically significant impairment identified for youth with internalizing and externalizing diagnoses are not accounted for by gender.

One possible explanation of the observed findings suggests that externalizing disorders might not be inherently associated with greater levels of functional impairment. Indeed, others have suggested that diagnostic severity and functional impairment are two orthogonal constructs, such that the number or type of symptoms experienced by the individual is not necessarily equated with the degree of functional impairment present (e.g., Bird et al., 1987; Winters et al., 2005). Moreover, major depressive disorder, an internalizing disorder, has been found to be associated with as much, and in some cases more, impairment than oppositional defiant and conduct disorders (Pickles et al., 2001; Simonoff et al., 1997). In the present sample, this hypothesis is supported by the data showing that (a) youth with externalizing diagnoses did not obtain significantly higher GAF scores than youth with internalizing diagnoses and (b) equivalent proportions of youth with internalizing and externalizing disorders were identified as impaired on the GAF at all score ranges (see Table 3). If these results are robust and replicable in future research, the observed differences here in impairment scores between internalizing and externalizing youth might be an artifact of the CAFAS’s scale structure and method of scoring. This interpretation would therefore suggest that the CAFAS is differentially sensitive to externalizing impairment, particularly given the number of scales assessing functioning related to externalizing behaviors (i.e., five scales) relative to those assessing functioning related to internalizing behaviors (i.e., three scales). Indeed, levels of functioning tend to vary across domains, each of which will be affected differentially according to the symptoms experienced (Winters et al., 2005). However, an equally plausible alternative explanation for these findings is that the GAF is insensitive to the differences associated with internalizing and externalizing impairment and fails to detect the more severe levels of functional impairment associated with externalizing disorders. These findings raise the important question of whether the observed differences in SED rates amongst internalizing and externalizing youth when using the CAFAS are due to (a) externalizing disorders being associated with an inherently greater level of functional impairment than are internalizing disorders or (b) an artifact of the measurement strategy itself.

The differential rates of SED identified amongst youth with internalizing and externalizing disorders when using the CAFAS and GAF bring to light several clinical and research implications, the most salient of which is access to mental health services for youth. Consistent with recommendations for functional impairment instruments specifying that impairments in functioning should not vary uniquely as a function of the disorder (Bird & Gould, 1995; Canino et al., 1999), such measures should not perform differentially across different diagnostic groups. However, the CAFAS eight-scale total score appears to identify more youth with externalizing disorders as impaired than youth with internalizing disorders. As such, from a clinical perspective, using the CAFAS to define the functional impairment component of the SED definition might result not only in an underidentification of internalizing youth but also in a failure to allocate the necessary treatment resources...
to this group of youth. As such, these findings suggest that the particular measure employed to assess functional impairment could result in differential access to mental health services for children and adolescents. From an epidemiological perspective, using a total score cutoff on the CAFAS of 80 to identify SED might result in either inflated rates of clinically impaired functioning amongst youth diagnosed with externalizing disorders or suppressed rates amongst youth diagnosed with internalizing disorders. Alternatively, using the GAF to measure functional impairment might result in misrepresentations of the levels of functional impairment experienced by youth diagnosed with externalizing disorders. Accordingly, relying on the GAF as a sole indicator of functional impairment might result in the inadequate allocation of treatment resources to this diagnostic group. Collectively, the findings presented here stress the importance for future investigations to examine whether the CAFAS is indeed overly sensitive to impairment associated with externalizing disorders or whether the GAF is undersensitive to such impairment.

Given that the CAFAS and GAF do not perform similarly with respect to measuring clinically significant functional impairment in youth with internalizing and externalizing diagnoses, these data also show that special considerations may be needed depending on the goal of the assessment. For example, should the goal of the assessment be identification of youth with SED in order to make decisions concerning service allocation, relying solely on either the eight-scale CAFAS total score or the GAF total score to identify such youth is potentially problematic. Accordingly, one option when selecting measures of functional impairment might entail relying on multiple measures to make decisions concerning service allocation. Alternatively, users of the CAFAS might also examine different ways of scoring this instrument (rather than relying on summing across the eight scales) to achieve more equitable identification of youth with internalizing and youth with externalizing disorders across diagnostic groups. Evaluating different means of scoring the CAFAS to obtain an overall estimate of functioning might be a particularly relevant option given the varied dimensions of impairment assessed by this instrument (e.g., behavior toward others, self-harm).

More generally, the present findings suggest that unidimensional measures of functional impairment might be less prone to differentially identify youth as SED according to diagnosis. Given that unidimensional scales provide an overall indicator of global functioning, such measures might be preferred when attempting to capture a snapshot of adaptive functioning across domains. However, when the goal of assessment is to determine relative functioning in each of several domains, particularly with an eye toward specific treatment planning, a multidimensional measure might be preferred. Given that multidimensional measures, by their nature, will query about functioning in a variety of settings and circumstances, each differentially related to specific symptoms, these results suggest that such measures might be more prone to systematic bias with respect to the youth identified as SED as a function of the scale composition (e.g., a multidimensional scale assessing functioning across more domains sensitive to internalizing distress might identify significantly more youth with internalizing disorders). Based on the present findings, clinicians should carefully consider the measure of functional impairment employed given that the CAFAS and GAF are associated with differential rates of SED identification and, hence, carry the potential of differential allocation of mental health services to youth.

One limitation of the present investigation is that the vast majority of children in the present sample were identified and referred for a mental health evaluation by their schoolteacher or counselor, thereby allowing for a potential referral bias. Those children who come to the attention of school personnel may in some way be distinct from those who are referred by their parents or through other health care professionals. Similarly, given that children identified initially at school tend to be those who are disruptive or otherwise problematic in the school environment, the present sample was skewed to include a larger number of children presenting with externalizing problems rather than internalizing problems, thus potentially limiting the generalizability of the present findings. In addition, the marked differences in rates of clinically significant functional impairment identified using the CAFAS and the GAF suggest that further research efforts be directed toward establishing normative rates of serious impairment in functioning amongst samples of diagnosed youth. To date, the literature has provided few guidelines with respect to the rates of seriously impaired functioning that are to be expected in clinical samples. Likewise, an additional limitation of these findings is the absence of an established cutoff score on the GAF with which to identify SED status among youth, thus precluding a comparison of the present results with previous findings. Although the CGAS (used here as a guide with respect to appropriate cutoff scores) is a parallel measure to the GAF, which has demonstrated superior reliability relative to the CGAS when assessing functional status amongst youth, this area of research would benefit from addressing the specific question of cutoff scores to use with the GAF. Similarly, although two different GAF cutoff scores were tested in the present investigation to assess rates of SED among youth, empirically defined cutoff scores for the GAF would serve to advance research in this area. In addition, reliability estimates for the GAF when used with children and adolescents have been moderate (ranging from .54 to .72), suggesting that the present results be interpreted in light of the GAF’s demonstrated reliability with this population. Moreover,
although both are measures of functional impairment, the CAFAS and GAF each specify different time periods for rating. Specifically, the CAFAS indicates that the assessor consider the past 1 month, whereas the GAF instructs one to rate “present” difficulties. The present findings must therefore be interpreted in the context of the potential limitation created by the discrepancy in the time frames assessed by each instrument. However, it must also be noted that previous studies have not found significant differences between functional impairment ratings using unidimensional scales (e.g., the CGAS and the GAF) when raters are asked to consider different time periods (e.g., current vs. previous year functioning; Rey et al., 1995).

It must be noted that an additional limitation of this study is the potential bias introduced by having the same intake assessor complete both ratings of functional impairment as well as the diagnostic assessment. This lack of independence of ratings might have resulted in inflated associations between the CAFAS and the GAF as well as between ratings of impairment and diagnosis. However, given that it is routine clinical practice to use the information obtained during a detailed diagnostic interview to inform ratings of functional impairment, employing the same rater across measures might represent a degree of ecological validity. Still, one area of further research might include examining any potential changes in the relationship between diagnostic and impairment measures when having the same and different clinicians provide ratings on each set of measures. Similarly, although each intake assessor was trained to reliability using each of the key measures employed here (e.g., the ADIS-IV-C/P, the CAFAS, and the GAF), standardized follow-up evaluations of reliability were not conducted throughout the study. As such, potential respondent bias might have influenced the findings reported here. It is also apparent that the measures of diagnosis and impairment employed in this study are all completed from the perspective of the clinician. One particularly interesting area of further research could be to include and assess the contribution of a measure of functional impairment from the perspective of the family. Finally, the diagnostic interview employed in this investigation (the ADIS-IV-C/P) was designed primarily to assess for the presence of anxiety disorders, with a secondary function being to identify comorbid behavioral disorders. As such, the use of this interview to assess all participants might have introduced a potential bias with respect to the emotional and behavioral functioning of those youth diagnosed with externalizing disorders.

This study is intended to serve as an initial step in future determinations of the appropriate measures of functional impairment to employ with youth diagnosed with internalizing or externalizing disorders and the optimal way to use such measures. Future investigations in this area should first attempt to discern whether the differential rates of SED identification observed here are due to inherent differences in functional impairment experienced by youth with internalizing and externalizing disorders or whether these differences result from the structure and scoring procedures of the CAFAS and GAF. Future investigations might also examine whether other commonly used measures of functional impairment yield differential rates of clinically significant functional impairment across different groups, such as across diagnostic, gender, age, or ethnic groups. In addition, should measurement issues be determined to cause the differential identification rates observed here, future studies might also strive toward identifying optimal measurement strategies for assessing functional impairment amongst children and adolescents. Changing the way in which total CAFAS scores are derived might result in comparable case identification across both youth with internalizing and externalizing diagnoses. Similarly, defining functional impairment by a minimum score on a global measure of impairment (i.e., the GAF) with at least one elevated scale of a multidimensional measure (i.e., the CAFAS) might also result in more comparable rates of SED identification across diagnostic groups. Indeed, strategies that rely on multiple methods of assessing and defining functional impairment might maximize the unique features of different measures of functional impairment, allowing unidimensional measures to identify the presence of impaired functioning and multidimensional measures to pinpoint the specific domain in which deficits are present.

In summary, the intent of the present investigation was not to advocate for or against any given measure of functional impairment. Rather, the aim was to examine how two commonly employed measures of functional impairment (the CAFAS and the GAF), each representative of the two types of functional impairment measures currently available (multidimensional and unidimensional, respectively), performed with respect to identifying youth as functionally impaired across those diagnosed with externalizing or internalizing disorders. Differential rates of clinically significant functional impairment, and hence SED identification, were observed with the CAFAS eight-scale summed total score, but not with the GAF, such that proportionally more youth with externalizing than internalizing disorders were identified as functionally impaired by the CAFAS. However, the extent to which the observed findings are related to actual differences in the experience of functional impairment between youth with internalizing and externalizing diagnoses as opposed to measurement biases remains unclear. Specifically, these results demonstrate that the GAF and CAFAS each work differently with respect to the proportion of youth identified as impaired within each broad diagnostic category (internalizing vs.
externalizing). What remains to be determined, however, is whether the rates of actual impairment associated with these two diagnostic categories are similar (as indicated by the GAF) or different (as suggested by the CAFAS). Findings from this study thus demonstrate that caution is needed when relying on a single measure of functional impairment to make decisions with respect to service allocation and that more studies are needed to better understand the ways in which functional impairment is measured by unidimensional and multidimensional scales across youths with internalizing and externalizing disorders.

Note

1. Nearly all consensus diagnoses incorporated information from all five sources, as standard protocol for CCBT assessors included gathering information from child–parent interviews, child–parent assessment measures, and relevant collateral sources. There were a few instances, however, whereby teachers were unable to be contacted, or the child or parent was unavailable to complete the interview.

References


**About the Authors**

*Sarah E. Francis*, PhD, is an assistant professor of psychology at Memorial University of Newfoundland. Her research interests include parental contributions to child anxiety, anxiety sensitivity, and anxiety and related disorders in children.

*Chad Ebesutani*, MA, is a clinical psychology resident in the University of Mississippi Medical Center’s Residency Training Program in the Department of Psychiatry and Human Behavior.

*Bruce F. Chorpita*, PhD, is a professor of psychology at the University of California, Los Angeles. His research interests include advancing the effectiveness of current mental health practice technologies for children and adolescents.