

Are Children with Anxiety Disorders Privately Referred to a University Clinic Like Those Referred from the Public Mental Health System?

Michael A. Southam-Gerow · Bruce F. Chorpita ·
Lauren M. Miller · Alissa A. Gleacher

Published online: 29 November 2007
© Springer Science+Business Media, LLC 2007

Abstract Compared two groups of children with anxiety disorders served at a single mental health clinic whose referral source differed: private referrals (i.e., parent/legal guardian initiated) and public referrals (e.g., via state contracts—Departments of Health and Education, juvenile justice system). Comparisons were made across three domains of variables: (a) symptoms/diagnoses, (b) functioning, and (c) environments. Few symptom differences emerged. However, large differences were evident for contextual variables like family income and life stressors. Overall, the pattern of differences point to possible directions for adaptation of treatments for use with children with anxiety disorders served in public mental health systems.

Keywords Child anxiety · Dissemination research · Effectiveness research · Public mental health clinics · Services research

Child clinical researchers have accumulated impressive evidence in support of the efficacy of treatments for childhood mental health problems tested in research settings for both internalizing (e.g., Chorpita and Southam-

Gerow 2006; Weisz et al. 2006) and externalizing disorders (e.g., McMahon et al. 2006). Unfortunately, these therapies are not the ones typically used in “real-world” settings such as community mental health (e.g., Weiss et al. 2000; Weisz et al. 1992) or school-based programs (e.g., Burns and Hoagwood 2002; Kataoka et al. 2003), and it is only recently that efforts have been made to deploy evidence-based treatments (EBTs) in such settings (e.g., Chorpita et al. 2002). This gap between science and practice represents a critical public health issue, as indicated by the NIMH’s various initiatives and workgroups examining the gap and how to close it (e.g., Hoagwood and Olin 2002; NAMHC Workgroup 2001).

Scholarly and policy explanations have focused on differences between the clinics from which research evidence comes (i.e., research clinics) and community service clinics (e.g., public clinics). Research clinics typically treat a single problem area, like child anxiety, are usually university-based, and typically rely on a referral stream similar to that of the private practitioner—i.e., family-initiated referrals, often via advertisements or professional recommendation. Research clinics are typically grant-supported. On the other hand, community mental health programs are typically located in mental health agencies or schools, rely on referrals coming from multiple sources, though predominantly via other governmental agencies (e.g., schools, juvenile justice). Funding for these clinics come also from multiple sources, though again most funding comes from public sources (e.g., Medicaid). Given these differences, skepticism remains that research clinic samples are representative of the caseload seen by the typical public clinic therapist (e.g., Persons and Silberschatz 1998; Weisz 2000).

Researchers have offered an ecological model for considering the external validity of treatment research,

M. A. Southam-Gerow (✉) · L. M. Miller
Department of Psychology, Virginia Commonwealth University,
806 W. Franklin St, Box 842018, Richmond, VA 23284-2018,
USA
e-mail: masouthamger@vcu.edu

B. F. Chorpita · A. A. Gleacher
University of Hawai’i, Manoa, Honolulu, HI, USA

Present Address:
A. A. Gleacher
Center for Advancement of Children’s Mental Health,
Columbia University, New York, USA

suggesting that differences between children seen in research clinics and those seen in public clinics may differ across (at least) four levels: (a) client, (b) provider, (c) agency, and (d) system (see Chorpita et al. 2002; Schoenwald and Hoagwood 2001; Southam-Gerow et al. 2006). Differences at the level of the client represent the most frequently nominated reason for concern about the applicability of EBTs because most research on predictors of treatment outcome have emphasized how client-level variables influence the effects of treatments (e.g., Berman et al. 2000; Brent et al. 1998; Rohde et al. 2004; Southam-Gerow et al. 2001). A common claim is that the problems experienced by youth and families seen in “real world” settings are more numerous and more severe than those in the clinics from which most of the research evidence has come (e.g., Weisz et al. 1995). Another common assertion is that the context (e.g., ethnicity, parental psychopathology, significant life events) of the youth in research samples differs from that of youth in “real world” clinics (e.g., Hammen et al. 1999). Differences on such variables are a concern because these variables appear to influence treatment outcome (e.g., Griffith et al. 1998; Siqueland et al. 2002).

Relatedly, a critical external validity question faces the field as the emphasis moves toward dissemination of treatments: How different are the families who seek help at research clinics from those who seek help at public clinics? Because EBTs were primarily designed for and tested with clients from research clinics, any differences could necessitate changes (minor to major) to the treatment programs for use in community mental health program contexts. Thus, to facilitate effectiveness and dissemination research and further intervention innovations, differences between the two samples need to be identified.

A recent study by Southam-Gerow et al. (2003) provided preliminary evidence that supported the skeptics’ position: a research clinic sample did differ from a public clinic sample. In the study, public clinic youth were similar to the specialty clinic youth in terms of internalizing problems. However, the public clinic youth had much higher levels of externalizing behavior problems (effect sizes as high as Cohen’s $d = 1.42$). In addition, youth in public clinics were from families with lower household income and were more likely to come from single-parent households. Any of these differences could have an impact on the effects of intervention programs developed in research settings when transported to a service setting in the community. These findings are complemented by the recent finding of Southam-Gerow et al. (2006) suggesting that differences between child anxiety disorder research clinics are small, supporting the contention that research clinics recruit homogeneous samples (e.g., Weisz 2000). However, the Southam-Gerow et al. (2003) study was limited in that the authors (a) compared two samples from

different geographic areas, (b) used different diagnostic procedures, and (c) were not able to determine the primary diagnoses in the public clinic sample.

In the current study, we complement the findings of Southam-Gerow et al. (2003) by testing for differences between groups of children with anxiety disorders distinguished by their referral source. One group was family referred to a university clinic (hereafter, *privately referred*) whereas the second group was referred through state contracts with the public community mental health system (hereafter, *publicly referred*). Although the referral sources were different, all youth received services at the same clinic, the University of Hawai’i Center for Cognitive Behavior Therapy (hereafter CCBT). Further, although referral sources differed, we restricted our sample to youth whose primary diagnosis was an anxiety disorder. We chose to focus on anxiety disorders for several reasons. First, anxiety disorders represent the most common mental health problem in childhood. Further, the evidence base for treating youth with anxiety disorders is arguably the strongest (Chorpita and Southam-Gerow 2006). Because childhood anxiety disorders represent a reasonable first-choice focus for dissemination efforts, the external validity of the evidence base is a critical question. Third, the Southam-Gerow et al. (2003) findings suggest that youth with anxiety disorders seen in research clinics are different from those seen in public mental health clinics. Replication (or not) of this finding represents an important contribution.

Our *privately referred* sample was comprised of individuals contacting the CCBT on their own (e.g., via word of mouth, professional referrals, response to publicity). Such a method is similar to the one used in research clinics.

In contrast, children in the *publicly referred* sample were served at the CCBT as part of state contract designed to provide comprehensive emotional and behavioral assessment services to children between the ages of 5 and 18. The funding sources for the *publicly referred* sample were a mixture of state general funds from the departments of education and health and from the state Medicaid plan. In both groups, assessment referrals were accepted for the majority of disorders common among school-age children (e.g., anxiety, depression, oppositional behavior, ADHD, and conduct disorder). The CCBT is the largest contractor for assessments performed on children in the public mental health system; for some school districts, it is the exclusive outside contractor. Thus, children served through these public mental health contracts are believed to be representative of the typical children served in the state system at this level of care. Not surprisingly, the large majority of these referrals were initiated by a human services worker from one of the funding agencies (DOH or DOE) contacting the clinic to provide necessary background information and an accompanying referral packet.

Because these two different referral channels ultimately led to the same set of diagnostic procedures and the same clinical care, the current test of differences between populations seeking private versus public mental health programs was able to control for many more variables than previous research on this issue (e.g., geography, assessment protocol; Southam-Gerow et al. 2003).

We tested for differences by applying selected features of the assessment model proposed by Hoagwood et al. (1996). Their model identifies five concentric spheres of influence on child outcomes, and five related outcome domains that may be important to assess, including: (1) *symptoms and diagnoses*; (2) *functioning* (i.e., adaptation to demands of home, school, peer group, etc.); (3) *consumer perspectives* (i.e., subjective experience, satisfaction, etc., of participating child and/or family); (4) *environments* (i.e., modifications in the child's life settings, brought about by an intervention); and (5) *systems* (includes assessment of whether service use patterns change after a target intervention). For the present study, we focused on three of these domains, namely symptoms/diagnoses, functioning, and environments. We hypothesized that the privately referred sample would show lower levels of diagnostic comorbidity, fewer symptoms of disorders, and better functioning. We also expected that the publicly referred sample would be more ethnically diverse, based on past comparative research as well as reviews suggesting that diverse ethnic groups are poorly represented in the treatment literature (e.g., Safren et al. 2000). In addition, we expected that the publicly referred youth would be more economically disadvantaged and would include more single-parent families (cf. Hammen et al. 1999; Southam-Gerow et al. 2003; see also Capage et al. 2001). Finally, we expected that the publicly referred sample would have higher levels of parental symptoms (e.g., Hammen et al. 1999).

Method

Participants and Recruitment Information

Both the privately referred and publicly referred samples came from the Center for Cognitive Behavior Therapy (CCBT) at the University of Hawai'i. The Center for CBT at the University of Hawai'i provides assessment services to children in the public mental health system through contracts with the state departments of health and education. The majority of assessments are performed on children in residence on O'ahu, but children from Hawaii, Maui, Molokai, and Kauai are served as part of these contracts. The main purpose of these assessments is to determine eligibility for initiating or continuing mental health services through the state health plans, principally involving either Medicaid supported services or educationally related

services through the Individual with Disabilities Education Act or Section 504 of the Vocational Rehabilitation Act. The majority of assessments are performed in communities, either in public schools, homes, or regional 'family guidance centers.' Families who are publicly referred do not bear any of the costs of these services. The CCBT also serves clients outside of the public system through private referrals. Data for the current study include participants recruited at the clinic from 1998 to 2003. From a sample of 673, we restricted the sample for the present study to youth with a principal or co-principal anxiety disorder diagnosis by either child or parent report on the ADIS-C/P. All children participating in assessments spoke English. No other inclusion or exclusion criteria were applied to the sample.

The final sample size numbered 227 with a mean age of 12.45 years ($SD = 3.08$, range 5–18). There were 97 girls and 130 boys in the total sample. Regarding ethnicity, participants self-identified as follows: 49% multi-ethnic, 12% Japanese, 12% European-American, 8% "other," 8% Hawai'ian, 5% Filipino, 2% Samoan, 1% Chinese, 1% Korean, 1% African-American, and <1% Hispanic and Puerto Rican. The privately referred sample numbered 57 and the publicly referred sample 170. The privately referred sample had 49% boys whereas the publicly referred sample had 60% boys; the difference between the groups was not significant ($\chi^2 [1, n = 227] = 2.06, p = 0.15$). Regarding age, the privately referred sample's mean age was 12.5 and the publicly referred sample mean age was 12.4. The difference was nonsignificant ($t [225] = -0.25, p = 0.81$). Ethnic group differences are discussed later.

Diagnostic/Symptom Domain Measures

Child Behavior Checklist (CBCL; Achenbach 1991). The CBCL is a widely used 118-item scale that assesses parents' view of an array of behavioral problems and social competencies in their children. Psychometric characteristics of the measure are strong. For the present study, we focused on the eight narrow-band scales to parallel the Southam-Gerow et al. (2003) study.

Revised Child Anxiety and Depression Scales (RCADS; Chorpita et al. 2000, 2005). The RCADS is a 47-item self-report questionnaire, with scales corresponding to separation anxiety disorder (SAD), social phobia (SP), generalized anxiety disorder (GAD), panic disorder (PD), obsessive compulsive disorder (OCD), and major depressive disorder (MDD). The RCADS requires respondents to rate how often each item applies to them. Items are scored 0–3 corresponding to "never," "sometimes," "often," and "always." Several investigations have demonstrated support for the RCADS in

non-referred samples of youth (Chorpita et al. 2000, 2005; de Ross et al. 2002). In a clinical sample, the RCADS has been shown to have excellent reliability and validity, with discriminant validity superior to some of the most commonly used measures of anxiety and depression (Chorpita et al. 2005).

Anxiety Disorders Interview Schedule for Children, DSM-IV Child and Parent Version (ADIS-IV-C/P; Silverman and Albano 1996). The ADIS-IV-C/P is a pair of semi-structured diagnostic interviews administered to parents and children, respectively, to determine the presence of DSM-IV diagnoses in children and adolescents. Supportive reliability data have been reported for the ADIS-IV-C/P (Silverman et al. 2001) versions. Assessment involves two interviews, one with the parent(s) and one with the child. Separate diagnostic profiles are derived from each interview, and are then combined in a clinical supervision meeting to form a consensus diagnosis (see Silverman and Albano 1996). With all families, at least one caretaker was present for each of the parent interviews. For this study, child-report, parent-report, and clinician consensus diagnoses were recorded. For data analytic purposes, we used consensus diagnoses only.

Functioning Domain Measures

Child and Adolescent Functional Assessment Scale (CAFAS). The CAFAS is a clinician-scored, multidimensional rating scale designed to assess functional impairment as experienced by children and adolescents, ages 5–17, across eight domains of functioning (Hodges and Wong, 1996, 1997). Raters are provided with a list of behavioral descriptors on 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 and Wong 1996). Items within each subscale are grouped according to four degrees of impairment: severe, moderate, mild, and minimal or no impairment, yielding scores of 30, 20, 10, or 0 points, respectively. Data generated from two large-scale evaluation studies have indicated that the CAFAS possesses good internal consistency (ranging from 0.63 to 0.78) and high interrater reliability (above 0.92 for the Total CAFAS score, and above 0.83 for the individual scales; Hodges and Wong 1996). Content, concurrent, and predictive validity have also been examined with the CAFAS, suggesting that the CAFAS correlates significantly and positively with other indicators of impairment, including severity of psychiatric diagnosis and subsequent service utilization (Hodges and Wong 1996). We focused on the Youth scales, School/Work, Home, and Behavior toward Others, to tap child functioning.

Dimensional Ratings (Chorpita, Yim, et al. 2000). Following each diagnostic interview, ADIS-IV-C/P interviewers completed a separate measure that required them to assign severity ratings for 13 DSM-IV disorders, whether or not criteria were met for those disorders. Ratings were based on clinicians' assessments of the degree to which the dimension of each disorder was present in the child. Ratings were completed separately for parent and child interviews. These scores ranged from 0 to 8, with higher scores representing increased clinical severity. Thus, a child with severe panic disorder might get a rating of 7 or 8 for that disorder, whereas a child with mild apprehension about shortness of breath might receive a rating of 2 or 3. Dimensional ratings (DRs) were adapted from clinical severity ratings (CSRs), which are a standard part of ADIS administration (e.g., Silverman and Nelles 1988; Silverman and Albano 1996). DRs are assigned for each disorder regardless of whether diagnostic criteria are met or nearly met. A recent investigation (Francis and Chorpita 2004) found that DRs showed a greater range of variability than CSRs from standard administration of the ADIS, and that DRs more closely corresponded to child self- and parent report of behavioral and emotional symptoms, due to their improved distributional properties. Chorpita, Yim, et al. (2000) reported good reliability of these ratings in a child clinical sample (median inter rater reliability = 0.92), and ratings were found to significantly discriminate across diagnostic groups, suggesting favorable validity.

DSM Axis V, Global Adaptive Functioning (GAF). The GAF is a clinician-administered assessment of the global functioning and current level of impairment of the child using a 0–100 scale, with higher scores indicating higher levels of functioning (American Psychiatric Association 1994).

Environments Domain Measures

DSM Axis IV. The DSM-IV Axis IV is reserved for recording environmental or psychosocial factors that may impact the accurate diagnosis and/or treatment prognosis for the presenting mental health problem (American Psychiatric Association 1994). Stressors recorded on Axis IV can be independent of or a consequence of an individual's presenting problem. Severe environmental and psychosocial problems that are the main focus of intervention can be coded on Axis I. Following the entire interview procedure, interviewers rated each child for the presence of problems in any of 9 domains on Axis IV: Primary Support Group, Social Environment, Educational, Occupational, Housing, Economic, Access to Health Care, Interaction with Legal System/Crime, Other Psychosocial/Environmental.

Depression Anxiety Stress Scales, Short Version (DASS-21, Lovibond and Lovibond 1995). The DASS-21 is the short

form of the 42-item Depression Anxiety Stress Scales (Lovibond and Lovibond 1995). Antony et al. (1998) reported supportive psychometric data. The DASS was administered to the child's parent to complete about him/herself.

Sociodemographic information. In addition, sociodemographic information (e.g., family income, family composition) was collected from parent(s) of the privately referred and publicly referred samples. The primary variables of interest were ethnicity, family income, parent marital status, and parental education.

Results

Analytic Plan

We compared the publicly referred and privately referred samples on three domains of variables: (a) symptoms/diagnoses, including child (ADIS-C, RCADS) and parent (ADIS-C/P, CBCL) report; (b) functioning (CAFAS, GAF); and (c) environments (DSM Axis IV sociodemographic factors, DASS-21). Because we were conducting multiple tests, we adjusted our alpha level to minimize Type I errors using a modified Bonferroni procedure described by Holm (1979) as recommended by Jaccard and Guilamo-Ramos (2002) across each of the several broadly defined families of tests (see Table 1). The Holm procedure involves a step-down approach whereby the achieved p -values for each family of tests are ranked from lowest to highest. The lowest p -value is compared with the critical p -value of $0.05/k$, where k is the number of tests conducted in the family. If the result is deemed significant, then the next lowest p -value is compared with $0.05/(k - 1)$ and so on. The procedure is stopped once a non-significant result is found—all subsequent results in the family are deemed

non-significant. Given the different sizes of our families, the procedure resulted in our applying different significance levels for individual tests (i.e., per comparison), but we used the same *per-family* error rate across the tests, namely $\alpha = 0.05$. Table 1 summarizes our adjustments. For the ADIS-C/P data, we used a p -value of 0.05 for each test as each test represented a test of unique data (i.e., parent and child report for each ADIS-C/P diagnoses are independent of each other).

For the RCADS and CBCL, we also conducted clinical significance tests using normative comparisons (Kendall et al. 1999) to see if the proportion of youth in the clinical vs. non-clinical range on the scales was higher in the publicly referred or privately-referred groups.

In addition to these analyses, we also conducted equivalency tests (Rogers et al. 1993; Jaccard and Guilamo-Ramos 2002) for all nonsignificant results. We took this additional step because although we expected to find differences, past work (Southam-Gerow et al. 2003) has reported similarities. Equivalency tests involve calculating z -score based deviations of group differences from an equivalency point established by the researcher. In other words, the researcher selects a minimum difference (δ) between the groups that s/he will consider “important enough to make the groups nonequivalent” (Rogers et al. 1993, p. 554). The test involves determining if the two means differ by more or less than δ . The calculated z -scores represent a test of the null hypothesis that the means *do* differ by more than δ . Thus, significant results indicate statistical equivalence.

Symptom and Diagnostic Domain Analyses

Parent-report symptom measures. Table 2 reports the results for these analyses. A few statistically significant

Table 1 Alpha-levels applied for tests using Holm procedure

Domain Symptom	Measure	Test # 1	Test # 2	Test # 3	Test # 4	Test # 5	Test # 6	Test # 7	Test # 8
Symptoms	RCADS	0.008	0.010	0.013	0.017	0.025	0.050	NA	NA
	CBCL	0.006	0.007	0.008	0.010	0.013	0.017	0.025	0.050
	ADIS-C	0.050	NA						
	Dimensional ratings	0.013	0.017	0.025	0.050	NA	NA	NA	NA
Functioning	CAFAS	0.017	0.025	0.050	NA	NA	NA	NA	NA
	GAF	0.050	NA						
Environments	Sociodemographic	0.050	NA						
	Axis IV Stressors	0.125	0.017	0.025	0.050	NA	NA	NA	NA
	DASS-21	0.017	0.025	0.050	NA	NA	NA	NA	NA

Note: RCADS = Revised Child Anxiety and Depression Scale; CBCL = Child Behavior Checklist; ADIS-C = Anxiety Disorders Interview Schedule for Children; CAFAS = Child and Adolescent Functioning Scale; GAF = General Adaptive Functioning; DASS-21 = Depression and Anxiety Stress Scale-21 item version

Table 2 Symptoms/diagnoses domain results: parent report measures

	Public referrals		Private referrals		<i>t</i>	df	Cohen's <i>d</i>	Equivalence <i>z</i> -score
	Mean	SD	Mean	SD				
CBCL withdrawn	63.56	10.87	67.02	10.92	−1.95	200	0.32	1.83
CBCL somatic problems	61.53	9.93	64.60	12.06	−1.80	200	0.29	1.98
CBCL anxiety/depression	63.37	10.47	68.62	9.80	−3.12*	200	0.51	–
CBCL social problems	63.39	10.93	62.50	9.32	0.52	200	0.08	4.15*
CBCL thought problems	61.63	8.87	65.68	9.52	−2.75	200	0.45	1.71
CBCL attention problems	63.59	10.32	64.32	9.43	−0.45	200	0.07	3.46*
CBCL delinquent behavior	60.09	9.14	55.28	5.49	3.51*	200	0.57	–
CBCL aggressive behavior	58.71	9.53	58.80	8.01	−0.06	200	0.01	3.87*
DR-P anxiety	10.71	6.94	13.58	8.23	−2.36	179	0.39	1.24
DR-P behavior	8.56	6.63	6.46	6.02	1.96	179	0.32	1.36

Note: CBCL = Child Behavior Checklist; DR-P = Dimensional Ratings-Parent report

* $p < \text{appropriate alpha-level}$, as per the Holm procedure, see Table 1

group differences emerged from these analyses. CBCL Anxiety/Depression scores were higher for the privately referred youth ($d = 0.51$) whereas CBCL Delinquent Behavior scores were higher for the publicly referred youth ($d = 0.57$). These effect sizes were in the “medium” range using Cohen's standards and similar to the mean of the effect sizes found by Southam-Gerow et al. (2003). No other group differences emerged; effect sizes for the remaining six subscales ranged from 0.01 to 0.45. Analyses of parental Dimensional Ratings (DR) for anxiety disorders and behavior disorders also yielded no-significant differences between groups.

We followed up all eight (i.e., six of the CBCL scales, the two DR scales) nonsignificant t -tests with the equivalency testing procedure described by Rogers et al. We set δ at 10% of the privately referred sample mean, following Rogers et al.'s procedure. We adjusted alpha using the Holm procedure (see Table 1). Results from these are reported in Table 2. Briefly, three of the CBCL scales were statistically equivalent across the groups: Social Problems, Attention Problems, and Aggressive Behavior. Neither DR scale reached statistical equivalence.

In addition to these t -tests, a normative comparison procedure (cf. Kendall et al. 1999) was used to compare the proportions of youth in the two samples who were above/below a T-score of 70 on the eight CBCL scales. We applied the Holm procedure to adjust alpha-levels for these tests (see Table 1). Three of these tests supported significant group differences: Delinquent Behavior, Thought Problems, and Withdrawn. The results indicated that more publicly referred youth were above the clinical cutoff on the Delinquent Behavior scale whereas a greater proportion of privately referred youth were above the clinical-cut off on the Thought Problems and Withdrawn behaviors scales.

The remaining five tests indicated nonsignificant group differences.

Child-report symptom measures. Table 3 reports the results for these analyses. In brief, none of the tests was statistically significant. Effect sizes (d) for the RCADS subscales ranged from 0.01 to 0.18, all considered “small” using Cohen's (1988) standards. The DR tests were also non-significant with effect sizes of 0.14 for the anxiety DR and 0.24 for the behavior disorder DR. In addition, none of the normative comparison tests yielded statistically significant results. In other words, the two groups did not differ in terms of proportion of youth above/below the T-score of 70 on any of the RCADS scales.

We followed up these nonsignificant t -tests with the equivalency testing procedure described by Rogers et al. (1993; see also Jaccard and Guilamo-Ramos 2002). We set δ at 10% of the privately referred sample mean, following Rogers et al.'s procedure. We applied the Holm procedure in correcting out alpha-level to control for multiple tests (see Table 1). Results from these are reported in Table 3. Briefly, five of the six RCADS scales, Social Phobia, OCD, GAD, MDD, and Panic Disorder, were statistically equivalent between the groups. The equivalency test for the remaining RCADS scale, Separation Anxiety Disorder, was not statistically significant. Neither test of the DR scales achieved statistical equivalence.

Youth diagnoses. Tables 4 and 5 provide a summary of consensus diagnoses assigned for the sample by group—primary diagnoses are presented in Table 4 and any diagnoses are presented in Table 5. Chi-square statistics and p -values are also presented. In short, the privately referred sample primary diagnoses of specific phobia, GAD, and OCD were statistically significantly more likely than in the publicly referred sample. Concerning any diagnoses, youth

Table 3 Symptoms/diagnoses domain results: youth report measures

	Public referrals		Private referrals		<i>t</i>	df	Cohen's <i>d</i>	Equivalence <i>z</i> -score
	Mean	SD	Mean	SD				
RCADS SAD	57.40	14.70	54.81	14.09	1.11	213	0.18	1.24
RCADS social anxiety	48.20	11.57	47.89	12.24	0.17	213	0.03	2.71*
RCADS panic disorder	54.01	14.35	54.16	17.33	−0.06	211	0.01	2.28*
RCADS OCD	48.54	12.51	48.10	13.99	0.22	211	0.03	2.54*
RCADS GAD	48.93	12.36	49.85	13.25	−0.43	213	0.07	2.93*
RCADS MDD	54.68	14.98	55.16	15.50	−0.19	206	0.03	2.41*
DR-C anxiety	12.56	7.28	13.71	9.96	−0.85	177	0.14	0.16
DR-C behavior	5.76	4.57	4.65	4.90	1.41	178	0.24	0.82

Note: RCADS = Revised Child Anxiety and Depression Scale; DR-C = Dimensional Ratings-Child report

* *p* < appropriate alpha-level, as per the Holm procedure, see Table 1

Table 4 ADIS-C/P consensus primary diagnoses by group

	Public referrals (%)	Private referrals (%)	χ^2	<i>p</i> -value
Separation anxiety disorder	8.80	7.00	0.18	0.670
Social phobia	24.70	22.80	0.08	0.772
Specific phobia	1.80	7.00	3.94	0.047*
Generalized anxiety disorder	5.30	15.80	6.44	0.011*
Post traumatic stress disorder	7.10	3.50	0.93	0.335
OCD	4.70	17.50	9.64	0.002*
Anxiety NOS	4.70	0.00	2.78	0.095
Major depression, single episode	1.80	1.80	0.00	0.996
Major depression, recurrent	0.60	3.50	2.79	0.095
Dysthymia	0.60	3.50	2.73	0.093
ADHD	17.60	7.00	3.79	0.052
Oppositional defiant disorder	10.00	7.00	0.45	0.501
Conduct disorder	5.90	0.00	3.51	0.061

Note: Total primary diagnoses do not sum to 100% because some primary diagnoses not included in table

* *p* < 0.05

Table 5 ADIS-C/P diagnoses by group (primary and additional)

	Public referrals (%)	Private referrals (%)	χ^2	<i>p</i> -value
Separation anxiety disorder	14.10	14.00	0.00	0.988
Social phobia	40.00	36.80	0.18	0.673
Specific phobia	13.50	17.50	0.55	0.457
Generalized anxiety disorder	10.00	28.10	11.22*	0.001
Post traumatic stress disorder	10.60	3.50	2.66	0.103
OCD	7.10	21.10	8.84*	0.003
Anxiety NOS	4.70	0.00	2.78	0.095
Major depression, single episode	4.10	3.50	0.04	0.838
Major depression, recurrent	0.60	8.80	11.11*	0.001
Dysthymia	2.40	7.00	2.73	0.098
ADHD	26.50	14.00	3.69	0.055
Oppositional defiant disorder	21.20	19.30	0.09	0.762
Conduct disorder	8.20	0.00	5.00*	0.025

* *p* < 0.05

in the privately referred sample were more likely to have OCD, GAD, and MDD-recurrent diagnoses whereas youth in the publicly referred sample were more likely to have conduct disorder diagnoses. We also used to *t*-tests to compare the groups with regard to total number of comorbid diagnoses, $t(225) = -0.653$, $p = 0.515$, and total number of additional non-anxiety disorder diagnoses $t(225) = -0.402$, $p = 0.688$). Neither test was statistically significant ($p > 0.50$) and effect size estimates were low by Cohen's standards and compared to those found in the Southam-Gerow et al. (2003) study: 0.10 and 0.06 respectively.

Functioning Domain Analyses

Table 6 reports the results for the three CAFAS scales and the GAF. In brief, results for GAF and for two of the CAFAS scales (i.e., Home and Behavior toward Others) were not statistically significant. However, publicly referred youth had higher CAFAS School/Work scale scores ($d = 0.63$), indicating poorer role functioning for these youth compared to their privately referred peers.

For the follow-up equivalency tests, we set δ at 5 points for the two CAFAS tests, because that scale is rated in increments of 10 units. Using similar logic for the GAF, we set δ at 10 points for this test. We set our alpha at 0.025 (i.e., 0.05/2) for the CAFAS tests and at 0.05 for the GAF test. Results from these analyses are reported in Table 4. All three of these functioning measures (GAF and two

CAFAS scales) were statistically equivalent across the two groups.

Environments Domain Analyses

The environments domain measures included three DASS-21 (Anxiety, Depression, Stress) scales, the presence of observed DSM Axis IV stressors, and several sociodemographic variables (ethnicity, family income, parent marital status, and parental education). Table 7 reports the DASS-21 results and Table 8 reports the family income and parental education results.

There were no statistically significant differences for any of the three DASS-21 scales, with effect sizes ranging from 0.10 to 0.25. However, publicly referred referrals had significantly lower family income and significantly lower maternal and paternal education levels; the effect sizes for these differences ranged from 0.75 to 0.92, all considered large by Cohen's standards and comparable to the largest effects found in the Southam-Gerow et al. (2003) study.

For the equivalency testing procedure for the DASS-21 scales, we set δ at 10% of the privately referred sample's mean and set our alpha as indicated in Table 1. None of the three scales were statistically equivalent between the two groups.

Concerning marital status, publicly referred were significantly more likely to live with an unmarried parent (53.7% compared to 36.4% among private referrals); this difference was statistically significant $\chi^2(1, 219) = 4.93$

Table 6 Functioning domain results

	Public referrals		Private referrals		<i>t</i>	df	Cohen's <i>d</i>	Equivalence <i>z</i> -score
	Mean	SD	Mean	SD				
GAF	57.78	15.24	56.25	19.05	0.62	225	0.09	6.07*
CAFAS school/work	18.14	10.02	11.60	11.49	3.87*	204	0.63	–
CAFAS home	7.90	8.99	6.86	8.36	0.73	206	0.12	4.24*
CAFAS behavior towards others	8.28	7.00	6.67	6.64	1.29	207	0.23	5.93*

Note: GAF = General Adaptive Functioning; CAFAS = Child and Adolescent Functioning Assessment Scale

* $p <$ appropriate alpha-level, as per Holm procedure, see Table 1

Table 7 Environment domain results: non-sociodemographic continuous measures

Environments domain	Public referrals		Private referrals		<i>t</i>	df	Cohen's <i>d</i>	Equivalence <i>z</i> -score
	Mean	SD	Mean	SD				
Maternal DASS-21-Stress	7.47	7.92	9.02	9.26	–1.14	185	0.19	1.80
Maternal DASS-21-Anxiety	6.48	7.84	7.29	8.30	–0.62	186	0.10	1.18
Maternal DASS-21-Depression	5.77	7.22	7.69	8.36	–1.54	180	0.25	0.92

Note: DASS-21 = Depression and Anxiety Stress Scale-21 item version

Table 8 Environment domain results: continuous measures

Environments domain	Public referrals		Private referrals		<i>t</i>	df	Cohen's <i>d</i>	Equivalence <i>z</i> -score
	Mean	SD	Mean	SD				
Family income (in thousands US\$)	43.09	29.65	71.84	45.61	−4.65*	157	0.82	–
Mother education	3.14	1.33	4.39	1.30	−5.48*	202	0.94	–
Father education	3.03	1.40	4.10	1.49	−4.52*	181	0.75	–

* $p < 0.0001$ **Table 9** Environment domain: presence of specific axis IV stressors

	Public referrals (%)	Private referrals (%)	χ^2	<i>p</i> -value
At least one problem	78.20	40.40	28.50*	0.0009
Primary support group problems	40.00	17.50	9.54*	0.002
Social environment problems	25.90	21.10	0.54	0.464
Educational problems	59.40	24.60	20.74*	0.0009

* $p <$ appropriate alpha-level, as per Holm procedure, see Table 1

$p < 0.03$. Table 9 presents DSM Axis IV results. Our analyses for Axis IV focused on those problems that occurred in more than 5% of either sample. Consistent with expectations for a child anxiety sample accessing mental health services, very few participants from either referral source were coded as having problems with housing, finances, occupation, access to health care, or legal system involvement. Remaining problem areas were: primary support group, social environment, and educational. Rates for all three were moderate to high in both groups, with the publicly referred youth being significantly more likely to have had at least one stressor coded on DSM's Axis IV (78.2% compared with 40.4% among private referrals). Further, youth in the publicly referred sample were more likely to have primary support group (40 vs. 17.5%) and educational problems (59.6 vs. 24.6%). No differences emerged in terms of social environment problems (25.9 vs. 21.1%).

Considering the considerable diversity of Honolulu, we examined youth ethnicity as broadly as possible. Accordingly, we examined the five most common categories: Caucasian, Hawai'ian, Japanese, multi-ethnic, and "other." All but the last category were determined by participant report; the "other" category included participant reports of Chinese, Filipino, Korean, Samoan, African-American, and Hispanic youth. We conducted McNemar tests comparing the groups each to determine whether youth in the privately referred or publicly referred samples were more likely to be from one of these five ethnic minority groups; we set alpha according to the Holm procedure, as described above. From this perspective, we found one statistically significant difference, specifically that there were more Caucasian youth in the privately referred sample (see Table 10).

Table 10 Environment domain results: ethnicity

Ethnicity	Public referrals % (<i>n</i>)	Private referrals % (<i>n</i>)	Z_{un}
Multi-ethnic	48.8 (83)	47.4 (27)	0.19
Caucasian	9.4 (16)	21.1 (12)	−2.32*
Other	22.4 (38)	12.3 (7)	1.65
Japanese	10.6 (18)	14.0 (8)	−.71
Hawai'ian	8.8 (15)	5.3 (3)	0.86

* $p > 0.01$

Discussion

We compared youth referred from the public mental health system (i.e., public referrals) to youth privately referred to a university specialty clinic (i.e., private referrals). All youth received services at the same university-based mental health center. We examined for differences between these two groups of youth across three domains of variables, applying Hoagwood et al.'s (1996) model: (a) symptoms/diagnoses, (b) functioning, and (c) environments. Overall, results indicated differences, primarily related to functioning and family environment, and similarities, primarily related to symptoms and diagnoses. The pattern of these differences and similarities points to possible directions for treatment development and adaptation to guide dissemination of EBTs for childhood anxiety disorders.

In the symptom/diagnosis domain, similarities outnumbered differences; of 20 comparisons made, only two were statistically significant, and effect sizes for the differences were below 0.60 for all tests, with a mean of 0.20. In addition, nine of the comparisons suggested that symptoms

and diagnoses were statistically *equivalent* between the two groups; the majority of these were child-report variables. The significant differences between the groups concerning symptoms and diagnoses were as follows: publicly referred had higher scores on the CBCL Delinquent Behavior scale and were more likely to be in the clinical range on that scale whereas private referrals were more likely to have higher levels of anxiety/depression symptoms and more likely to be in the clinical range on CBCL Withdrawn and Thought Problem scales. Furthermore, youth in the publicly referred sample had more comorbid conduct disorder diagnoses than youth in the privately referred sample. All differences were by parent-report. The overall picture in terms of clinical symptoms suggested that (a) privately referred youth were as, if not more, severe than publicly referred youth on the target syndrome of anxiety and (b) publicly referred youth were somewhat more severe in terms of delinquent behavior—i.e., a non-target syndrome.

Concerning measures of functioning, public referrals functioned more poorly at school than privately referred youth. However, three other measures (two CAFAS scales and GAF) of functioning not only indicated no statistically significant differences but were all three statistically equivalent. The findings regarding school functioning are not completely surprising, given that privately referred youth are usually referred by their parents, whereas publicly referred youth are often identified and referred by child workers or educational professionals. It follows, then, that these children would likely have more problems at school, which is often the trigger for their referral by someone other than the parent.

The most notable differences were found in the environments domain. For example, family income and level of parental education were both much lower in publicly referred youth. Furthermore, life stressors, as tapped by DSM's Axis IV was also significantly higher in the publicly referred group: indeed, fully 78% of those youth experienced at least one life stressor coded on Axis IV compared to 40% of the privately referred youth, a nearly 2:1 ratio, with differences between groups emerging in the areas of education and primary support group. In addition, youth in the publicly referred sample were more likely to come from a single-parent family. The magnitude of these differences was quite large, with effect sizes exceeding 0.75 for these comparisons.

In addition, the privately referred sample contained more Caucasian youth compared to the publicly referred sample. The finding underscores criticism of the fact that Caucasian families are over-represented in the evidence base (e.g., Huey and Polo in press). Thus, recruitment and engagement of minorities in research remains a critical need for future work, especially given the relative paucity of published work with families with children (e.g., Huey

and Polo in press; Rodríguez et al. 2006; see also, US DHHS 2003). Although we were reasonably successful in our recruitment efforts in the publicly referred sample, there remains room for additional progress. No differences were extant between the groups concerning level of maternal depression and anxiety, another environments domain measure.

Before considering the implications of our findings, differences between our study and the Southam-Gerow et al. (2003) study merit discussion to clarify the independent variable under study. Our project compared youth from different *referral sources* and the Southam-Gerow et al. (2003) study compared youth from different *settings*. An argument can be made that setting and referral source are analogous. “Lab” studies conducted in research clinics typically employ similar recruitment methods to those used for our privately referred sample. As an example, the “lab” sample in the Southam-Gerow et al. (2003) study was drawn from a private university-based research clinic whose primary referral sources were advertisement and word-of-mouth. And the community sample in the Southam-Gerow et al. study represented youth from the public mental health system in Los Angeles, CA, comparable to our publicly referred sample. Thus, one may be tempted to generalize (if cautiously) from our present findings to differences between lab and community samples.

However, caution is necessary because setting and referral source are not necessarily the same. Discussion of lab/community differences typically assumes that lab (and community) samples are homogenous (e.g., Weisz et al. 1995). Although some data have supported the assumption (e.g., Southam-Gerow et al. 2006), future work needs to replicate the finding. Further, as our study indicates, one can see youth from both “populations” in the same setting. In other words, it is important to avoid conflation of setting and referral source. Accordingly, future lab- and community-based clinical research should carefully document recruitment strategies to help readers determine which population(s) is being studied (e.g., Weisz et al. 2003).

From the present findings, one could tentatively conclude that evidence-based treatments (EBTs) for anxiety tested with youth in settings that serve privately referred youth may be “ready” for the typical clinic anxiety disordered youth in the public mental health system. On the other hand, our findings regarding functioning and ecology suggest that the “readiness” may be limited. EBTs may be ready to treat severe anxiety, but less is known about how they will perform in the context of increased functional impairment, environmental challenges, and to a lesser extent, increased externalizing comorbidity. Environmental variables in particular (e.g., Hoagwood et al. 1996) have been associated with premature termination of treatment or attenuation of treatment effects in past research (see, e.g.,

Kazdin et al. 1992; Kendall and Sugarman 1997; Nock and Kazdin 2001) and are associated with lower levels of client engagement in treatment (e.g., Griffith et al. 1998; Siqueland et al. 2002). In fact, Kazdin's "barriers to treatment" model (Kazdin et al. 1997; Kazdin and Wassell 1999) suggests, in part, that what predicts early termination of treatment is not level of psychopathology but family ecology variables like income and motivation. Considering further that many of the publicly referred youth were referred through the public education or mental health systems, one could infer that parental awareness of the child's anxiety problem and/or parental motivation to seek a solution to that anxiety problem could also be lower, suggesting still more potential challenges to successful intervention.

Thus, it may be worth considering that although our intervention technology for anxiety disorders may be quite well designed for severe anxiety itself, supplemental components might be needed to address a host of additional challenges. For example, interventions that decrease barriers to treatment like lack of child care for siblings, lack of transportation to/from the clinic, or the need for evening hours to accommodate work schedules (e.g., Henggeler et al. 2002; Nock and Kazdin 2005) may help bolster outcomes for laboratory tested treatments when deployed in clinics serving clients like those in our publicly referred sample. Such supplemental intervention could for example make getting the child to the clinic an explicit therapeutic task for the provider and parent to address together, rather than a *fait accompli*. Further, engagement with the child and family would also become a therapeutic task. Evidence suggests that client engagement is related to outcomes for families with children experiencing externalizing behavior problems (e.g., Cunningham and Henggeler 1999; Patterson and Chamberlain 1994) and to mental health service utilization among children and families in urban inner cities (McKay et al. 1996).

Limitations should be noted in a few specific domains. First, the sample is confined to youth with primary anxiety disorders and the implications of these findings among children with other presenting problems are unknown. Because comorbidity was not an exclusionary criterion, the sample population includes youth with a variety of diagnoses. This alleviates some of the concern about generalizability, but a cautious generalization to youth with primary anxiety disorders is probably best. Second, though the ethnically diverse sample is representative of Hawai'i's population, it does not reflect the ethnic composition of the rest of the United States. However, as the United States becomes increasingly diverse, this sample's ethnic composition may foreshadow the desired research samples of

the future. Further, the sample may be more similar to public mental health clinics in terms of diversity than the composition of many laboratory research samples.

The breadth of measurement in this study could also have been improved. More measurement of non-anxiety related clinical concerns as well as factors related to family perceptions/attitudes toward therapy could improve the strength of the findings. In addition, although evaluation of group differences across three domains suggested by Hoagwood et al. (1996) (symptoms/diagnoses, functioning, and environment) speaks to the efforts made to assess a wide range of factors, we recognize that important outcome domains could have been measured better. As an example, the GAF as a measure of functioning, though common in clinical settings, may not be ideal, as other measures have been developed (e.g., Bird et al. 1993, 2005; Shaffer et al. 1983). Additionally, measurement of the number, nature, and severity of psychosocial stressors experienced by the youth and their families could have been improved. Future research could improve the measurement model used here as well as adding other measures to tap relevant constructs omitted here.

One final observation concerns the scientific "conversation" about treatment development and adaptation of current treatment models and manuals. A primary question is where to conduct the development and adaptation work. Considering the differences between "lab" and community clinic samples found in past work and the differences we report here between publicly and privately referred youth, research clinic "labs" may not represent the ideal treatment development platform for continued improvement and innovation in deployable treatment technologies. Indeed, our study suggests that the greatest barriers to implementation involve factors found more commonly in public mental health systems. Accordingly, we suggest a few options for future clinical research (a) alter "lab" clinics to include publicly referred youth or (b) seek partnerships with public mental health or community practice organizations and conduct treatment adaptation work there (e.g., Southam-Gerow 2005). Either direction affords an excellent opportunity to lead to the next generation of treatment innovation (Chorpita 2002; Southam-Gerow et al. in press).

Acknowledgments Preparation of this article was supported in part by National Institute of Mental Health Grant K23 MH69421 to Southam-Gerow and R03 MH60134 to Chorpita as well as an award from the University of Hawai'i Research Council, and awards from the Hawai'i Departments of Health and Education to Chorpita. We are grateful to Tina Donkervoet and John Weisz for their support and to A. Aukahi Austin, Sarah Francis, Jennifer Gray, Farrah Greene, Cassian Kimhan, Catherine Moffitt, Brad Nakamura, and John Young for their assistance with data collection.

References

- Achenbach, T. M. (1991). *Manual for the child behavior checklists/4-18 and 1991 profile*. Burlington: University of Vermont.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders*, 4th edn. Washington, DC: Author.
- Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment*, *10*, 176–181.
- Berman, S. L., Weems, C. F., Silverman, W. K., & Kurtines, W. M. (2000). Predictors of outcome in exposure-based cognitive and behavioral treatments for phobic and anxiety disorders in children. *Behavior Therapy*, *31*, 713–731.
- Bird, H. R., Canino, G. J., Davies, M., Ramirez, R., Chavez, L., Duarte, C., et al. (2005). The Brief Impairment Scale (BIS): A multidimensional scale of functional impairment for children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*, 699–707.
- Bird, H., Shaffer, D., Fisher, P., Gould, M., Staghezza, B., Chen, J., et al. (1993). The Columbia Impairment Scale (CIS): Pilot findings on a measure of global impairment for children and adolescents. *International Journal of Methods in Psychiatric Research*, *3*, 167–176.
- Brent, D. A., Kolko, D. J., Birmaher, B., Baugher, M., Bridge, J., Roth, C., et al. (1998). Predictors of treatment efficacy in a clinical trial of three psychosocial treatments for adolescent depression. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*, 906–914.
- Burns, B. J., & Hoagwood, K. (Eds.) (2002). *Community treatment for youth: Evidence-based interventions for severe emotional and behavioral disorders*. New York: Oxford.
- Capage, L. C., Bennett, G. M., & McNeil, C. B. (2001). A comparison between African American and Caucasian children referred for treatment of disruptive behavior disorders. *Child and Family Behavior Therapy*, *23*, 1–14.
- Chorpita, B. F. (2002). Treatment manuals for the real world: Where do we build them? *Clinical Psychology: Science and Practice*, *9*, 431–433.
- Chorpita, B. F., Moffitt, C. E., & Gray, J. (2005). Psychometric properties of the Revised Child Anxiety and Depression Scale in a clinical sample. *Behaviour Research and Therapy*, *43*, 309–322.
- Chorpita, B. F., Plummer, C., & Moffitt, C. E. (2000). Relations of tripartite dimensions of emotion to childhood anxiety and mood disorders. *Journal of Abnormal Child Psychology*, *28*, 299–310.
- Chorpita, B. F., & Southam-Gerow, M. A. (2006). Treatment of anxiety disorders in youth. In E. J. Mash & R. A. Barkley (Eds.), *Treatment of childhood disorders*, 3rd edn. New York: Guilford Press.
- Chorpita, B. F., Yim, L. M., Donkervoet, J. C., Arensdorf, A., Amundsen, M. J., McGee, C., et al. (2002). Toward large-scale implementation of empirically supported treatments for children: A review and observations by the Hawai'i Empirical Basis to Services Task Force. *Clinical Psychology: Science and Practice*, *9*, 165–190.
- Chorpita, B. F., Yim, L., Moffitt, C., Umemoto, L. A., & Francis, S. E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: A revised child anxiety and depression scale. *Behaviour Research and Therapy*, *38*, 835–855.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd edn. Hillsdale: Lawrence Erlbaum Associates, Inc.
- Cunningham, P. B., & Henggeler, S. W. (1999). Engaging multi-problem families in treatment: Lessons learned throughout the development of multisystemic therapy. *Family Process*, *38*, 265–286.
- de Ross, R., Gullone, E., & Chorpita, B. F. (2002). The Revised Child Anxiety and Depression Scale: A psychometric investigation with Australian youth. *Behaviour Change*, *19*, 90–101.
- Francis, S. E., & Chorpita, B. F. (2004). Behavioral assessment of children in outpatient settings. In S. N. Haynes & E. M. Heiby (Eds.), *Comprehensive handbook of psychological assessment: Vol. 3. Behavioral assessment* (pp. 291–319). Hoboken: John Wiley & Sons.
- Griffith, J. D., Knight, D. K., Joe, G. W., & Simpson, D. D. (1998). Implications of family and peer relations for treatment engagement and follow-up outcomes: An integrative model. *Psychology of Addictive Behaviors*, *12*, 113–126.
- Hammen, C., Rudolph, K., Weisz, J., Rao, U., & Burge, D. (1999). The context of depression in clinic-referred youth: Neglected areas in treatment. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*, 64–71.
- Henggeler, S. W., Schoenwald, S. K., Rowland, M. D., & Cunningham, P. B. (2002). *Serious emotional disturbance in children and adolescents: Multisystemic therapy*. New York: Guilford Publications, Inc.
- Hoagwood, K., Jensen, P. S., Petti, T., & Burns, B. J. (1996). Outcomes of mental health care for children and adolescents: I. A comprehensive conceptual model. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*, 1055–1063.
- Hoagwood, K., & Olin, S. S. (2002). The NIMH blueprint for change report: research priorities in child and adolescent mental health. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*, 760–767.
- Hodges, K., & Wong, M. M. (1996). Psychometric characteristics of a multidimensional measure to assess impairment: The child and adolescent functional assessment scale. *Journal of Child and Family Studies*, *5*, 445–467.
- Hodges, K., & Wong, M. M. (1997). Use of the Child and Adolescent Functional Assessment Scale to predict service utilization and cost. *Journal of Mental Health Administration*, *24*, 278–290.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*, 65–70.
- Huey, S. J., & Polo, A. J. (in press). Evidence-based psychosocial treatments for ethnic minority youth: A review and meta-analysis. *Journal of Clinical Child & Adolescent Psychology*.
- Jaccard, J., & Guilamo-Ramos, V. (2002). Analysis of variance frameworks in clinical child and adolescent psychology: Advanced issues and recommendations. *Journal of Clinical Child and Adolescent Psychology*, *31*, 278–294.
- Kataoka, S. H., Stein, B. D., Jaycox, L. H., Wong, M., Escudero, P., Tu, W., et al. (2003). A school-based mental health program for traumatized Latino immigrant children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*, 311–318.
- Kazdin, A. E., Holland, L., & Crowley, M. (1997). Family experience of barriers to treatment and premature termination from child therapy. *Journal of Clinical and Consulting Psychology*, *65*, 453–463.
- Kazdin, A. E., Siegel, T. C., & Bass, D. (1992). Cognitive problem-solving skills training and parent management training in the treatment of antisocial behavior in children. *Journal of Consulting and Clinical Psychology*, *60*, 733–740.
- Kazdin, A. E., & Wassell, G. (1999). Barriers to treatment participation and therapeutic change among children referred for conduct disorder. *Journal of Clinical Child Psychology*, *28*, 60–172.
- Kendall, P. C., Marrs-Garcia, A., Nath, S., & Sheldrick, R. C. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Consulting and Clinical Psychology*, *67*, 285–299.

- Kendall, P. C., & Sugarman, A. (1997). Attrition in the treatment of childhood anxiety disorders. *Journal of Consulting and Clinical Psychology, 65*, 883–888.
- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales*, 2nd edn. Sydney: Psychology Foundation of Australia.
- McKay, M. M., McCadam, K., & Gonzales, J. J. (1996). Addressing the barriers to mental health services for inner city children and their caretakers. *Community Mental Health Journal, 32*, 353–361.
- McMahon, R. J., Wells, K. C., & Kotler, J. S. (2006). Conduct problems. In E. J. Mash & R. A. Barkley (Eds.), *Treatment of childhood disorders*, 3rd edn. New York: Guilford Press.
- Nock, M. K., & Kazdin, A. E. (2001). Parent expectancies for child therapy: Assessment and relation to participation in treatment. *Journal of Child and Family Studies, 10*, 155–180.
- Nock, M. K., & Kazdin, A. E. (2005). Randomized controlled trial of a brief intervention for increasing participation in parent management training. *Journal of Consulting and Clinical Psychology, 73*, 872–879.
- Patterson, G. R., & Chamberlain, P. (1994). A functional analysis of resistance during parent training therapy. *Clinical Psychology: Science and Practice, 1*, 53–70.
- Persons, J. B., & Silberschatz, G. (1998). Are results of randomized controlled trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology, 66*, 126–135.
- Rodríguez, M. D., Rodríguez, J., & Davis, M. (2006). Recruitment of first-generation Latinos in a rural community: The essential nature of personal contact. *Family Process, 45*, 87–100.
- Rogers, J. L., Howard, K. I., & Vessey, J. T. (1993). Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin, 113*, 553–565.
- Rohde, P., Clarke, G. N., Mace, D. E., Jorgensen, J. S., & Seeley, J. R. (2004). An efficacy/effectiveness study of cognitive-behavioral treatment for adolescents with comorbid major depression and conduct disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 660–668.
- Safren, S. A., Gonzalez, R. E., Horner, K. J., Leung, A. W., Heimberg, R. G., & Juster, H. R. (2000). Anxiety in ethnic minority youth: Methodological and conceptual issues and review of the literature. *Behavior Modification, 24*, 147–183.
- Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination of interventions: What matters when. *Psychiatric Services, 52*, 1190–1197.
- 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.
- Silverman, W. K., & Albano, A. M. (1996). *Anxiety disorders interview schedule for DSM-IV: Child version*. San Antonio: Psychological Corporation.
- Silverman, W. K., & Nelles, W. B. (1988). The influence of gender on children's ratings of fear in self and same-aged peers. *Journal of Genetic Psychology, 149*, 17–21.
- Silverman, W. K., Saavedra, L. M., & Pina, A. A. (2001). Test-retest reliability of anxiety symptoms and diagnoses with anxiety disorders interview schedule for DSM-IV: Child and parent versions. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 937–944.
- Siqueland, L., Crits-Christoph, P., Gallop, B., Gastfriend, D., Lis, J., Frank, A., Griffin, M., Blaine, J., & Luborsky, L. (2002). Who starts treatment: Engagement in the NIDA collaborative cocaine treatment study. *American Journal on Addictions, 11*, 10–23.
- Southam-Gerow, M. A. (2005). Using partnerships to adapt evidence-based mental health treatments for use outside labs. *Report on Emotional and Behavioral Disorders in Youth, 5*, 58–60, 77–79.
- Southam-Gerow, M. A., Austin, A. A., & Marder, A. M. (in press). Transportability and dissemination of psychological treatments: Research models and methods. In D. McKay (Ed.), *Handbook of research methods in abnormal and clinical psychology*. Newbury Park: Sage.
- Southam-Gerow, M. A., Kendall, P. C., & Weersing, V. R. (2001). Examining outcome variability: Correlates of treatment response in a child and adolescent anxiety clinic. *Journal of Clinical Child Psychology, 30*, 422–436.
- Southam-Gerow, M. A., Ringeisen, H. L., & Sherrill, J. T. (2006). Introduction to Special Issue: Integrating interventions and services research: Progress and prospects. *Clinical Psychology: Science and Practice, 13*, 1–8.
- Southam-Gerow, M. A., Silverman, W. K., & Kendall, P. C. (2006). Client similarities and differences in two childhood anxiety disorders research clinics. *Journal of Clinical Child and Adolescent Psychology, 35*, 528–538.
- Southam-Gerow, M. A., Weisz, J. R., & Kendall, P. C. (2003). Youth with anxiety disorders in research and service clinics: Examining client differences and similarities. *Journal of Clinical Child and Adolescent Psychology, 32*, 375–385.
- U.S. Department of Health and Human Services. (2003). Outreach notebook for the inclusion, recruitment and retention of women and minority subjects in clinical research (NIH publication no. 03-7036). Retrieved 1 Nov 2007, from <http://orwh.od.nih.gov/pubs/outreach.pdf>
- Weiss, B., Catron, T., & Harris, V. (2000). A 2-year follow-up of the effectiveness of traditional child psychotherapy. *Journal of Consulting and Clinical Psychology, 68*, 1094–1101.
- Weisz, J. R. (2000). Lab-clinic differences and what we can do about them: I. The Clinic-Based Treatment Development Model. *Clinical Child Psychology Newsletter, 15*(1–3), 10.
- Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63*, 688–701.
- Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin, 132*, 132–149.
- Weisz, J. R., Southam-Gerow, M. A., Gordis, E. B., & Connor-Smith, J. K. (2003). Primary and Secondary Control Enhancement Training for youth depression: Applying the Deployment-Focused Model of treatment development and testing. In A. E. Kazdin & J. R. Weisz (Eds.), *Evidence-based treatments for children and adolescents* (pp. 165–183). New York: Guilford.
- Weisz, J. R., Weiss, B., & Donenberg, G. R. (1992). The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist, 47*, 1578–1585.