

## Client Similarities and Differences in Two Childhood Anxiety Disorders Research Clinics

**Michael A. Southam-Gerow**

*Virginia Commonwealth University*

**Wendy K. Silverman**

*Florida International University*

**Philip C. Kendall**

*Temple University*

*Some evidence suggests that research and service clinics differ on treatment-relevant dimensions, but no study has examined whether research clinics (RCs) themselves differ. We compared 2 samples of children and adolescents (ages 7 to 17 years) with anxiety disorders treated in 2 different university-based child anxiety RCs, one in Philadelphia ( $n = 184$ ) and one in Miami ( $n = 64$ ), on child symptom and diagnostic measures, family characteristics (e.g., income), and level of maternal depression. The samples were not significantly different on any youth symptom and diagnostic measures except 1 (parent-reported diagnoses); further, the 2 samples were statistically equivalent on 3 of 6 youth symptom measures. Although the Miami sample had significantly more parent-reported diagnoses than the Philadelphia sample, the effect size was small (Cohen's  $d = .44$ ). Statistically significant differences between the samples in ethnic background and family income were observed, though the former was not significant after controlling for population differences. Findings are discussed in terms of the implications of homogeneity of RCs for treatments tested there.*

Despite the growing empirical support for child and adolescent (hereafter child) mental health treatments, there remains skepticism about the generalizability of the evidence base (e.g., Persons & Silberschatz, 1998; Weisz, Southam-Gerow, Gordis, & Connor-Smith, 2003). This doubt persists in part because data suggest some differences between research clinics (RCs), the settings from which most of the evidence base has emerged, and service clinics, the settings at which many youth receive treatment (e.g., Leaf et al., 1996). For example, several studies report that youth who present to service clinics have significantly higher levels of externalizing behavior problems and diagnoses, lower family incomes, a greater likelihood of coming from a single-parent family, higher rates of stressful life events, and higher levels of parental psychopathology than youth who present to RCs (e.g., Hammen, Rudolph, Weisz, Rao, & Burge, 1999; Southam-

Gerow, Weisz, & Kendall, 2003; Taylor, Francis, Chorpita, Southam-Gerow, & Lam, 2003). Researchers and policymakers have argued that these differences between service and RCs complicate the transport and dissemination of evidence-based treatments (EBTs) into practice settings (e.g., Hoagwood, Burns, & Weisz, 2002; Schoenwald & Hoagwood, 2001; Southam-Gerow, 2004).

The differences highlighted would be even more problematic if, across different RCs, the children and families recruited and selected for participation were homogeneous with respect to outcome-relevant client characteristics. Evidence of such homogeneity would undermine efforts to generalize from the extant evidence base. To date, however, there is no evidence to support or controvert the notion that RCs are homogeneous with regard to client characteristics. Thus, claims about the homogeneity (or heterogeneity) of the evidence base are without empirical basis. This study, a comparison of two childhood anxiety clinic samples, represents an effort to fill this gap in the literature.

Such a comparison offers an opportunity to specify the extent of RC samples' homogeneity or heterogeneity. As an illustration of the importance of the comparison, consider a cognitive-behavioral treatment for youth anxiety disorders that has been found to have efficacy at two sites. If the samples at the two sites have different levels of comorbidity or differ with regard to family income, confidence in the generalizability of

---

Preparation of this article was supported in part by National Institute of Mental Health Grant R01 MH59087 to Philip C. Kendall, R01 MH63997 to Wendy K. Silverman, and K23 MH69421 to Michael A. Southam-Gerow. We are grateful to the staffs at the Child and Adolescent Anxiety Disorders Clinic and Child Anxiety and Phobia Program, especially Timothy Verduin and Armando Pina for their assistance with the data sets.

Correspondence should be addressed to Michael A. Southam-Gerow, Virginia Commonwealth University, Department of Psychology, 808 W. Franklin St. Box 842018, Richmond, VA 23284-2018. E-mail: masouthamger@vcu.edu

the treatment across these variables is enhanced. Of course, even in the face of this evidence, external validity limitations could remain so long as the setting remained a “research clinic.” As some have argued, the setting of the treatment can be an important consideration in the implementation of treatments (e.g., Schoenwald & Hoagwood, 2001; Southam-Gerow, Ringeisen, & Sherrill, 2006). However, a comparison of two RC samples affords a data-based method to determine the breadth of a treatment.

In a comparison between two RCs, relevant domains should include potential factors that either predict or moderate treatment outcome. As the field moves beyond evaluating outcome, questions relating to predictors and moderators of treatment outcome have become increasingly relevant (Hinshaw, 2002; Kazdin & Nock, 2003; Kraemer, Stice, Kazdin, Oford, & Kupfer, 2001). Although data on moderating factors of anxiety disorders treatment are scarce, research to date has suggested several predictors of child treatment outcome, including family characteristics (e.g., income, family stress), parental psychopathology, and child problem severity (e.g., comorbidity). Low family income, single-parent family status, and high levels of family stress and dysfunction are related to outcomes (e.g., Crawford & Manassis, 2001; Southam-Gerow, Kendall, & Weersing, 2001). Similar results have been found in the youth depression treatment literature (e.g., Hammen et al., 1999). Further, family characteristics such as low family income are associated with both premature termination and attenuation of treatment effects (Kazdin, Siegel, & Bass, 1992; Griffith, Knight, Joe, & Simpson, 1998; Siqueland et al., 2002; Webster-Stratton, 1990).

Another predictor of treatment outcomes for youth has been parental psychopathology. Maternal depression was the strongest predictor of poor treatment response in a study of youth treated for anxiety disorders (Berman, Weems, Silverman, & Kurtines, 2000). Similarly, Brent et al. (1998) found maternal depression predicted outcomes for youth depression. Again focused on youth with depression, Hammen et al. (1999) found that rates of maternal and paternal psychopathology (e.g., substance abuse, depression, antisocial personality disorder) were high in nonresearch clinic samples. However, evidence on the specificity of these findings is lacking, and it remains unclear whether maternal (vs. paternal) psychopathology are both important variables. Further, it is not clear which type(s) of parental psychopathology are most relevant, though maternal depression has received the most attention to date.

Concerning client symptom level, studies have generally found better outcomes related to lower levels of client symptoms at pretreatment (e.g., Berman et al., 2000; Southam-Gerow et al., 2001). Research on co-

morbidity as a predictor for treatment of childhood anxiety disorders treatment has suggested no effect (e.g., Flannery-Schroeder, Suveg, Safford, Kendall, & Webb, 2004; Verduin & Kendall, 2003). This stands in contrast to research on childhood depression treatment, which has found that comorbidity dampens treatment effects (Brent et al., 1998; Rohde, Clarke, Mace, Jorgensen, & Seeley, 2004). Although it is possible that childhood anxiety treatment is more robust to comorbidity than depression treatment, it is notable that the two anxiety samples (Flannery-Schroeder et al., 2004; Verduin & Kendall, 2003) had few youth with comorbid disruptive behavior disorders as compared to the sample of depressed youth (Rohde et al., 2004). Specifically, the Rohde et al. study included a sample of youth of whom 100% had comorbid depression and conduct disorder, whereas none of the youth in the Verduin and Kendall study had comorbid conduct disorder and less than 10% of the sample had oppositional defiant disorder. Given the Southam-Gerow et al. (2003) results suggesting higher rates of externalizing problems in service clinic samples of anxious youth, the question remains open regarding comorbidity with externalizing disorders as a potential correlate of treatment outcome for youth with anxiety disorders.

Although differences between RC samples are possible, there is reason to believe samples would be similar across multiple sites because of similar recruitment practices and inclusion criteria, at least with regard to the variables represented in these selection criteria (e.g., Weisz et al., 2003). As a result, although our study employed traditional tests for sample differences, we also used statistical procedures to test for sample equivalence for some variables (e.g., Jaccard & Guilamo-Ramos, 2002b; Rogers, Howard, & Vessey, 1993) with regard to these selection variables.

In short, this study compared two RCs on three domains of client variables: (a) level of client symptoms and degree of client comorbidity, (b) family characteristics (e.g., family income, number of parents or caregivers in the home), and (c) maternal depression. Because both samples were recruited for National Institute of Mental Health funded clinical trials for childhood anxiety disorders (and thus possessed similar inclusion and exclusion criteria), we hypothesized that they would be equivalent on client symptom severity and client comorbidity. For nonselection variables (i.e., family characteristics, maternal depression), we anticipated that sample differences would not be statistically significant. Because the two clinics were located in different U.S. cities with different socio-demographic characteristics, we recognized the need to control for those differences in our analyses using U.S. Census data (U.S. Department of Commerce, 1992, 1993; cf. Southam-Gerow et al., 2003).

## Method

### Participants

One sample came from a university-based RC (Child and Adolescent Anxiety Disorders Clinic) at Temple University in Philadelphia, Pennsylvania (RC-PA). A. Our second sample came from a university-based RC at Florida International University in Miami, Florida (RC-FL). For this study, we focused on youth from each clinic with primary diagnoses of the three most common *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) anxiety disorder diagnoses among youth: separation anxiety disorder, generalized anxiety disorder, and social anxiety disorder. Data for this study were based on participant data from the already published treatment outcome studies (i.e., Kendall, 1994; Kendall, Flannery-Schroeder, & Panichelli-Mindel, 1997; Silverman et al., 1999<sup>1</sup>; Silverman et al., 1999).

The RC-PA sample ( $n = 184$ ) had 41.8% girls and a mean age of 11.24 years ( $SD = 1.52$ , range = 8.4–15.7) and the RC-FL sample ( $n = 64$ ) had 37.5% girls and a mean age of 10.94 years ( $SD = 3.1$ , range = 7.0–18.0). Age ( $p < .31$ ) and gender ( $p < .54$ ) differences were not statistically significant.

### Procedure

Recruitment procedures at the two clinics were generally similar. Youth were referred by parents, other mental health and professionals, and school personnel. After obtaining signed informed consent from parents and signed informed assent from youth, youth and their parent(s) completed diagnostic interviews and questionnaires. Children who received a primary or principal diagnosis of a *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.; American Psychiatric Association, 1987) anxiety disorder (i.e., separation anxiety disorder, overanxious disorder or generalized anxiety disorder, social phobia, or avoidant disorder) as determined by the Anxiety Disorders Interview for Children (ADIS-C/P; Silverman, 1991; Silverman & Albano, 1996) were enrolled. All families participating in this project qualified for and were offered treatment through National Institute of Mental Health funded randomized clinical trials. Only the pretreatment assessment data were used in this study.

**Child Behavior Checklist (CBCL; Achenbach, 1991).** The CBCL is a widely used 118-item rating scale that assesses parents' views of an array of behavioral problems and social competencies in their chil-

dren. Psychometric characteristics of the measure are strong; for example, stability correlations ranging from .65 to .87 have been reported for periods as long as 2 years (Achenbach, 1991). For this study, we used the broadband Externalizing scale and the three narrowband Internalizing scales (Withdrawn, Somatic Problems, and Anxiety/Depression; all are reported in *T* scores with a mean of 50 and a standard deviation of 10). We chose to focus on the narrowband scales within the Internalizing scale because the samples were recruited to be high on internalizing symptoms. We also analyzed data from the Externalizing subscale because of the potential relevance for treatment outcomes if differences emerged on this scale (e.g., provide parents with child management skills if children are high in externalizing behavior problems) and because the Southam-Gerow et al. (2003) findings suggested the importance of the variable.

**Children's Depression Inventory (CDI; Kovacs, 1992).** The 27-item CDI is a widely used child self-rating scale of childhood depressive symptoms. In clinical samples, Cronbach's alphas for the CDI have ranged from .71 to .89 and retest reliability coefficients have ranged from .50 to .87 (see Kovacs, 1992). The CDI also correlates above .50 with several measures of child depression across numerous studies (e.g., McCauley, Mitchell, Burke, & Moss, 1988; see Kovacs, 1992, for review). We focused on the Total Depression scale for this study.

**Revised Children's Manifest Anxiety Scale (RCMAS; Reynolds & Richmond, 1985).** A widely used measure of a youth's chronic anxiety, this scale consists of 37 items, including a 9-item Lie scale, to which children respond "Yes" or "No." Reynolds (1981) reported a retest correlation of .68 for a study with elementary children retested after a 9-month interval. Concerning validity, Reynolds (1980) found that the Total Anxiety score was correlated .78 with the State-Trait Anxiety Inventory for Children (CITATION) Trait scale and only .08 with the State scale of that measure, a finding that supports the notion that the RCMAS measures chronic rather than situational anxiety. We focused on the Total Anxiety scale score for this study.

**ADIS-C/P (Silverman, 1991).** The ADIS-C/P is a structured diagnostic interview to assess for a *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.; American Psychiatric Association, 1987) diagnoses in children and adolescents. The following diagnostic modules were included: separation anxiety disorder, avoidant disorder, overanxious disorder, social phobia, generalized anxiety disorder, simple phobia, agoraphobia, panic disorder, major depressive disorder, dysthymic disorder, attention defi-

<sup>1</sup>This study had a large number of youth with simple phobias who were not included in this study; however, the trial also included youth with social phobia, and these youth were included in the study.

cit hyperactivity disorder, oppositional defiant disorder, and conduct disorder. Interrater reliability of project diagnostic interviewers was regularly assessed for both projects. In both samples, kappas across all diagnoses exceeded .80, as reported in the treatment outcome studies (Kendall, 1994; Kendall et al., 1997; Silverman, Kurtines, Ginsburg, Weems, Lumpkin, et al., 1999; Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999). Despite some difference in diagnostic procedures, in both clinics children with a primary diagnosis of separation anxiety disorder, social phobia, or overanxious disorder by either informant (i.e., parent or child) was enrolled in the clinical trial. For our diagnostic analyses, we compared the RC-FL sample's consensus diagnosis to an analogous index in the RC-PA sample, namely, combining child and parent reports such that a diagnosis endorsed by either reporter was considered present.

**Beck Depression Inventory (BDI; Beck & Beamesderfer, 1974).** A widely used 21-item rating scale, the BDI assesses adults' levels of depressive symptoms (Kendall, Hollon, Beck, Hammen, & Ingram, 1987). Coefficient alphas of the BDI average in the high .80s with clinical samples (Beck, Steer, & Garbin, 1988). Concurrent and construct validity evidence has been established using a variety of psychological measures, and the measure has been shown to discriminate between depressed and nondepressed clinical patients (Steer, Beck, Riskind, & Brown, 1986). We focused on the Total Depression scale score for this study.

**Sociodemographic information.** Data on family income, family structure (single-parent vs. two-parent), and child ethnicity as well as basic information about the child (e.g., age, gender) and family were collected from parents at the time of the initial interview using a brief form.

## Results

We compared the two samples across three domains: (a) level of client symptoms and degree of client comorbidity, as reported by the child (CDI, RCMAS, ADIS-C) and parent (CBCL, ADIS-P), (b) family characteristics (e.g., family income, number of parents or caregivers in the home), and (c) maternal depression. Because we were conducting multiple tests, we adjusted our alpha level to minimize Type I errors using the modified Bonferroni procedure described by Holm (1979), as recommended by Jaccard and Guilamo-Ramos (2002a), across each of the broadly defined families of tests. The Holm procedure involves a stepdown approach whereby the achieved  $p$  values for each family of tests are ranked

from least to greatest. The least  $p$  value is compared with the critical  $p$  value of  $.05/k$ , where  $k$  is the number of tests conducted in the family. If the result is deemed significant, then the next least  $p$  value is compared to  $.05/(k - 1)$  and so on. The procedure is stopped once a nonsignificant result is found; all subsequent results in the family are deemed nonsignificant. 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  $p < .05$ . Our per-comparison  $p$  values are reported later, concurrent with the report of the results.

For the four CBCL scales, we also conducted clinical significance tests using normative comparisons (e.g., Kendall, Marrs-Garcia, Nath, & Sheldrick, 1999) to see if the two samples differed in terms of the proportion of youth in the clinical and nonclinical ranges on the scales of the measures. Using  $T$  scores of 70 as the cutoff, we placed each participant into an above- or below-threshold group. We used chi-square tests to determine if the RC-PA or RC-FL sample had a higher proportion of youth with  $T$  scores either equal to or exceeding 70 on the three CBCL internalizing subscales (i.e., Withdrawn Behavior, Somatic Problems, and Anxious/Depressed) and the broadband externalizing scale.

In addition, we conducted equivalency tests (Jaccard & Guilamo-Ramos, 2002b; Rogers et al., 1993) for all nonsignificant results, an additional step needed because we hypothesized no differences between the samples. Equivalency tests involve calculating  $z$  scores based on deviations of group differences from a researcher-set equivalency point. The researcher selects a minimum difference (i.e.,  $\delta$ ) between the groups that is considered "important enough to make the groups nonequivalent" (Rogers et al., 1993, p. 554). The test then involves determining if the two means differ by more or less than  $\delta$  through the calculation of  $z$  scores (standard scores with a mean of zero and standard deviation of 1). The calculated  $z$  scores represent a test of the hypothesis that the means do differ by more than  $\delta$ , and significant results indicate statistical equivalence.

Finally, because we expected many nonsignificant results, we used two other procedures selected to provide additional confidence that lack of statistical power did not limit our findings. First, we reported confidence intervals for the differences between the samples for all continuous measures. These intervals provided a probable upper-bound for any group differences. We also conducted post hoc power analyses to determine (a) the power to detect an effect size of .50 (i.e., a medium-sized difference using Cohen's, 1988,  $d$  or greater) and (b) the achieved power to detect the effect size we obtained.

**Child Symptom and Diagnosis Analyses**

Means, standard deviations, *t* test results, effect sizes, and 95% confidence intervals for the mean differences for all continuous symptoms measures are presented in Table 1. Critical significance levels for these tests were as follows: (a) child report (two tests in the family: critical *p* < .025 for the most extreme result, .05 for the next most extreme, assuming the first was significant) and (b) parent report (four tests in the family: critical *p* < .0125 for the most extreme, .017, .025, .05). None of the differences on the four CBCL scales between the RC-FL and RC-PA samples were statistically significant. Effect sizes (*d*) for these four tests ranged from .00 to .31, all in the “small” range using Cohen’s (1988) effect size standards. In addition, comparing these results to those from the Southam-Gerow et al. (2003) study (effect sizes ranged from .35 to 1.42, mean effect size of .50), we found the effect sizes in this study were all smaller than the smallest effect size in the previous study.

We followed up these nonsignificant *t* tests with the equivalency testing procedure. We set  $\delta$  at 10% of the RC-PA mean, following Rogers et al. (1993) procedure. Results are reported in Table 1 under the heading “Equivalence *z* Score.” We used the Holm procedure as with all tests; for these tests, the absolute value of the *z* score representing a statistically significant (and thus statistically equivalent) depended on the number of tests run. For the parent-report measures, we used *p* values of .017, .025, and .05 (given three tests) and .025 and .05 for the child-report tests. The results indicated that statistical equivalence was found for the CBCL Withdrawn, Somatic Problems, and Anxious/Depressed scales but not for the CBCL Externalizing scale.

We used an alternative method for testing the differences between the samples on the CBCL by comparing

the number of youth from each group who were either above or below the clinical cutoff *T* score of 70 for the four CBCL scales. We used chi-square tests to evaluate the statistical significance of the group differences; we applied the modified Bonferroni procedure as discussed previously to these tests. Specifically, we used the critical *p* values of .0125, .017, .025, and .05 for these tests. Results are presented in Table 2. None of the tests yielded statistically significant results.

The two sample groups were not statistically significantly different on either the RCMAS or the CDI. Effect sizes obtained on these measures were .24 (RCMAS) and .15 (CDI). We followed up these nonsignificant *t* tests with the equivalency testing procedure described previously (see Jaccard & Guilamo-Ramos, 2002b; Rogers et al., 1993). The results indicated that statistical equivalence was not found for

**Table 2.** Normative Comparison Analyses

	Miami		Philadelphia	
	%	<i>n</i>	%	<i>n</i>
CBCL Withdrawn Behavior Scale				
Below 70 <i>T</i> score	64.1	41	60.9	112
At or above 70 <i>T</i> score	35.9	23	39.1	72
$\chi^2(1, 248) = .21 p < .65$				
CBCL Somatic Problems Scale				
Below 70 <i>T</i> score	39.1	25	57.1	105
At or above 70 <i>T</i> score	60.9	39	42.9	79
$\chi^2(1, 248) = 6.17 p < .014$				
CBCL Anxious/Depressed Scale				
Below 70 <i>T</i> score	40.6	26	46.2	85
At or above 70 <i>T</i> score	59.4	38	53.8	99
$\chi^2(1, 248) = .60 p < .44$				
CBCL Externalizing Scale				
Below 70 <i>T</i> score	89.1	57	94.0	173
At or above 70 <i>T</i> score	10.9	7	6.0	11
$\chi^2(1, 248) = 1.74 p < .19$				

Note: CBCL = Child Behavior Checklist

**Table 1.** Descriptive and Test Statistics for Primary Symptom and Diagnostic Analyses

	Miami		Philadelphia		<i>t</i>	<i>df</i>	Cohen’s <i>d</i>	Equivalence <i>z</i> Score	95% CI of Difference	Achieved Power
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>						
Child Symptoms										
RCMAS Total	13.48	7.19	12.13	4.78	1.70	241	0.24	0.17	-0.25-2.95	0.36
CDI Total	11.94	7.88	10.77	7.48	1.10	243	0.15	0.08	-1.01-3.34	0.19
CBCL Withdrawn	65.61	10.40	65.64	10.30	0.00	246	0.00	-4.40*	-2.99-2.92	0.00
CBCL Somatic	69.67	10.50	66.50	9.98	2.20	246	0.31	-2.37*	0.29-6.06	0.55
CBCL Anxious/Depressed	72.33	8.78	71.36	10.00	0.70	246	0.10	-4.37*	-1.82-3.74	0.10
CBCL Externalizing	56.38	10.00	53.28	10.10	2.10	246	0.31	-1.53	0.21-5.98	0.55
Child Diagnoses										
ADIS-C/P total diagnoses	2.28	1.09	1.84	0.96	3.05*	244	0.44	NA	0.16-0.73	0.83
ADIS-C/P nonanxiety diagnoses	0.45	0.78	0.32	0.58	1.50	244	0.21	-3.94*	-0.08-0.35	0.30
Parent symptoms										
BDI Total	8.68	7.11	8.26	6.87	0.40	214	0.10	-0.41	-1.67-2.51	0.04

Note: RCMAS = Revised Child Manifest Anxiety Scale; CDI = Children’s Depression Inventory; CBCL = Child Behavior Checklist; ADIS-C/P = Anxiety Disorders Interview Schedule for Children/Parent report; BDI = Beck Depression Inventory; CI = Confidence interval; NA = for any tests yielding significant differences between groups, equivalence testing procedures were not conducted.

\**p* < .025.

the RCMAS and CDI, as the  $z$  scores were both below 0.5.

We also used two other procedures to examine these comparisons: (a) examination of the confidence intervals for the differences between groups and (b) post hoc power analyses. Concerning the confidence intervals, all suggested that the sample differences were likely to be within the value of the standard deviation of the measure. However, results indicated that the RC-FL sample scores for the Somatic Problems and Externalizing Behavior Problems scales could have been as high 6 points more than the RC-PA sample. Still, even at this extreme end of the confidence interval, such differences would only be in the moderate range, using Cohen's (1988) guide, and would still not approach the more extreme results found by Southam-Gerow et al. (2003). Sometimes nonsignificant results can be explained by lack of statistical power. For all of our tests, to detect an effect size of .50 or greater, achieved statistical power was greater than .90. Achieved power to detect the effect sizes that we obtained (i.e., all lower than .50) was lower, ranging from .00 to .55 (see Table 1).

Finally, using ADIS-C/P data, we compared the two samples on two diagnostic indexes: total diagnoses and total nonanxiety diagnoses (two tests in the family: critical  $p < .025$  for the most extreme result, .05 for the next most extreme, assuming the first was significant). Results indicated that the RC-FL sample had more total diagnoses than the RC-PA sample but that there were no statistically significant differences regarding total nonanxiety diagnoses. Results are presented in Table 1. Effect sizes were small for both comparisons ( $d = .44$  and  $.21$ , respectively). Using the equivalency testing procedure, we followed up the nonsignificant difference between the groups on nonanxiety disorder diagnoses, setting  $\delta$  at .5 diagnoses (see Table 1). We found statistical equivalence for the number of parent-reported nonanxiety diagnoses. As with the tests for child symptoms, we report confidence intervals and achieved power. It is worth noting that our power to detect effect sizes greater than .50 was at or above .90 for all tests.

### Contextual Factors: Maternal Depression and Sociodemographic Characteristics

Table 1 contains the results for maternal depression on the BDI, and Table 3 contains the sociodemographic data. Data on maternal-reported depression on the BDI was compared using a  $t$  test (see Table 1). The results of the  $t$  test indicated that the groups were not statistically significantly different, and the equivalence tests indicate that the groups were not statistically equivalent.

**Table 3.** Results for Sociodemographic Analyses

	Miami		Philadelphia	
	%	<i>n</i>	%	<i>n</i>
Ethnicity—two group				
European American	46.9	30	79.3	146
Non-European American	53.1	34	20.7	38
$\chi^2(1, 248) = 24.30 p < .0009$				
Ethnicity—four group				
European American	46.9	30	79.3	146
African American	3.1	2	13.0	24
Hispanic/Latino	45.3	29	3.3	6
Other	4.7	3	4.3	8
$\chi^2(3, 248) = 71.02 p < .0009$				
Family Income				
\$0–\$39,999	65.6	42	41.8	77
\$40,000–up	34.4	22	58.2	107
$\chi^2(1, 248) = 10.76 p < .002$				
Family structure				
Single parent	28.3	17	22.0	37
Two parent	71.7	43	78.0	131
$\chi^2(1, 228) = .37 p < .32$				

Concerning the sociodemographic data, we analyzed client ethnicity data in three ways. First, we compared the samples by dichotomizing cases as either European American or not. We did this because Census data suggested that although both populations were relatively diverse, each city was more likely to have one or another ethnic group. Using this dichotomous variable, we found that RC-FL youth were more likely to come from an ethnic minority group (primarily Hispanic) than the RC-PA youth; 53% of the RC-FL youth identified themselves as non-European American, whereas the proportion of RC-PA youth identifying themselves as non-European American was 20.7%,  $\chi^2(1, 248) = 24.30, p < .009$ . Second, we examined the data using a four-group variable for ethnicity: (a) European American, (b) African American, (c) Hispanic, and (d) other. This chi-square test was significant,  $\chi^2(3, 248) = 71.02, p < .009$ . Follow-up  $z$  tests indicated that the RC-PA sample had more European Americans ( $z = 4.93, p < .005$ ) and more African Americans ( $z = 2.23, p < .01$ ) than the RC-FL sample, whereas there were more Hispanics in the RC-FL sample compared to the RC-PA sample ( $z = -8.32, p < .005$ ).

Finally, because the two samples were drawn from geographically different areas, we considered whether group differences we did find were due to the geographical difference. To investigate this possibility, we examined 1990 Census data (U.S. Department of Commerce, 1992, 1993) to determine if the differences found would have been expected based on the geographical difference alone. Census data from 1990 were used for primary analyses instead of 2000 data because the two studies from which the data were drawn occurred in the 1990s. When we reconducted these analyses using 2000 Census data, our conclusions remained

the same.<sup>2</sup> Because we were conducting multiple tests, we applied the Holm procedure. For these tests, we used the following three significance levels: .017, .025, and .05. Using Census data, we determined that the differences concerning European Americans ( $z_{un} = 1.93, p < .027$ )<sup>3</sup> and Hispanics ( $z_{un} = .37, p < .36$ ) were not significantly different from what would be expected given the ethnic differences in the two metropolitan areas. However, that the RC-PA sample had more African American youth was not expected based on Census data ( $z_{un} = 6.61, p < .001$ ). Because the two cities have relatively equal proportions of African Americans, the RC-FL sample had an underrepresentative number of clients from this ethnic group. However, this finding is not directly pertinent to differences between the two RCs.

The RC-FL youth were more likely to come from families with low income levels. Almost 66% of the families of the RC-FL youth reported incomes below \$40,000, whereas the proportion of RC-PA youth reporting income below that level was 41.8%,  $\chi^2(1, 248) = 10.76, p < .002$ . We also determined the veracity of our income finding by comparing the two metropolitan areas on household income using 1990 Census data (U.S. Department of Commerce, 1993). We determined that the difference we found, though in the direction suggested by the census, was more extreme than would be anticipated ( $z_{un} = 2.58, p < .005$ ). Specifically, the Census data indicated that in the Philadelphia, Pennsylvania–New Jersey Principal Metropolitan Statistical Area, 56% of families reported income below \$40,000 compared to 67% of families in the Miami–Hialeah, Florida Principal Metropolitan Statistical Area.

Given this difference, we sought to determine if the Miami sample was different on family income compared to the Los Angeles sample from the Southam-Gerow et al. (2003) study. We observed that the proportion of lower income families in the Miami clinic, though larger than the Philadelphia clinic, was still not comparable to what we found in Los Angeles. Specifically, 84% of the families in the Los Angeles clinics had family incomes below \$40,000, compared to 65% in the Miami clinic (and 42% in the Philadelphia clinic). Statistical tests found this difference to be statistically significant,  $\chi^2(1, 138) = 6.095, p < .02$ , meaning that youth from Los Angeles clinics were poorer than those in the Miami clinic. This finding held even when controlling for Census data.

<sup>2</sup>Details are available from the first author.

<sup>3</sup>To test the hypothesis that the difference between the RC-PA and RC-FL samples were (or were not) different from the expected difference based on Census data on the populations, we adapted the  $z_{un}$  test for proportion differences for two independent samples. Interested readers can refer to our earlier article for a complete description (Southam-Gerow et al., 2003).

Concerning family structure (i.e., single-parent vs. two-parent), results indicated no significant differences between the two samples,  $\chi^2(1, 228) = .37, p < .32$ . We did not follow this nonsignificant test result with the U.S. Census analyses.

### Post Hoc Analysis

Because sociodemographic differences emerged between the two samples on two of three variables, we followed a procedure used by Southam-Gerow et al. (2003) to ascertain if the single significant difference on the diagnostic measure could be attributed to these sociodemographic differences. Thus, after completing the primary analyses, we conducted a hierarchical regression analysis to control for the impact of the sociodemographic factors to determine if the RC-PA–RC-FL difference on total ADIS–C/P diagnoses remained significant. In the regression analysis, we entered the ethnicity, family income, and family living arrangement and composition variables in Step 1 and the RC-PA–RC-FL status in Step 2 (see Table 4). The results indicated that site differences on the ADIS–C/P were still statistically significant ( $p > .05$ ) once we controlled for sociodemographic differences. Thus, the RC-FL sample included children with more diagnoses than those in RC-PA sample even when sociodemographic factors were controlled.

### Discussion

We compared two child anxiety RC samples across family characteristics, maternal depression, and child comorbidity. Few significant differences emerged, and those that did were small in magnitude. Further, using equivalence tests, there were several variables on which the two samples were statistically equivalent. Overall, the two samples were homogeneous.

As expected, the samples were not significantly different on any youth symptom and diagnostic measures except one. Further, the samples were statistically equivalent on several measures, specifically all three CBCL Internalizing subscales (i.e., Withdrawn, So-

**Table 4.** Summary of Hierarchical Multiple Regression Analyses for Total ADIS–C/P Diagnoses

Variable	B	SE B	$\beta$
Step 1			
Ethnicity	-.036	.161	-.016
Family Income	-.337	.142	-.170
Family structure	.224	.163	.097
Step 2			
Site (i.e., Miami vs. Philadelphia)	-.333	.157	-.149

Note: ADIS–C/P = Anxiety Disorders Interview Schedule for Children, Parent Version.  $R^2 = .063$  for Step 1 ( $p = .002$ );  $\Delta R^2 = .019$  for Step 2 ( $p = .035$ ).

matic Complaints, and Anxious/Depressed). As tapped by Cohen's (1988) *d*, an index of effect size, the differences on symptom and diagnostic measures between the two samples ranged from .00 to .44, with a mean of .20. Placing these findings in the context of similar research, none of the effect sizes approached the magnitude found in a previous investigation comparing a RC sample to a public mental health service clinic sample (Southam-Gerow et al., 2003); those differences had Cohen's *d* estimates ranging from .35 to 1.42 ( $M = .50$ ). Post hoc analyses indicated that our power to detect an effect of .50 or larger exceeded .80. Thus, although a larger sample could be used to replicate the findings, the results suggest that differences between these two RCs are very few and likely to be small.

We did find significant differences between the clinics on total number of child diagnoses, though the effect size was small (.44). Although the difference suggests that the Miami sample was somewhat more impaired than the Philadelphia sample, this finding was not corroborated by any other measure, indicating that any severity difference is small and limited. Further, the difference between the two samples for non-anxiety disorder diagnoses were in the opposite direction (i.e., the Philadelphia sample had more nonanxiety disorder diagnoses), though the difference was non-significant. One possible reason for the difference is that the Miami clinic also served youth with primary diagnoses of specific phobia (the Philadelphia clinic did not). Thus, although we excluded youth with primary specific phobia diagnoses from the study, the clientele of the clinic may have differed along these lines.

Sample differences were initially evident on two of three sociodemographic variables: ethnicity and family income. However, when geographic differences were controlled using Census data, most of the ethnicity differences were no longer significant. Specifically, differences between the two samples regarding the proportions of European Americans and Hispanics were expected based on Census data. However, the difference between the samples for African Americans was larger than would be expected based on Census data, indicating that the Philadelphia sample had more African American participants than did the Miami sample. Whereas the proportion of African American youth in Philadelphia was consistent with the Philadelphia metro area, the Miami clinic had an underrepresentation of African American youth (compared to the Miami metro area).

Family income difference was more extreme than would be expected given the population differences, based on U.S. Census data. Youth in the Miami sample had lower family incomes than those in the Philadelphia sample. In part, this reflects differences between the two cities—Census data indicated that Philadelphia families have higher incomes. However, controlling for these population differences, we still found that the

family incomes in the Miami sample were lower than those in Philadelphia sample. Finally, no differences emerged between the two samples on maternal depression, as tapped by the BDI.

Because our findings support the notion that specialty RCs tend to have relatively homogenous samples, several implications follow. As an overview, the results suggest that clinical research may benefit from diversification with regard to several of the client characteristics we studied. Such a conclusion is in concordance with the urgings of some clinical scientists who have suggested that we transform RCs into real-world labs in which treatments can be tested under progressively more genuine circumstances (e.g., complex cases, diverse samples, and so on; see Chorpita & Nakamura, 2004; Weisz et al., 2003). We briefly discuss areas for diversification next.

Our findings suggest that increasing the diversity of client samples is an important step. Here, diversity refers to a variety of client characteristics, including sociodemographic and clinical characteristics. For example, recruiting greater numbers of youth from lower income families or from ethnically diverse families would greatly help the field assess the viability of EBTs for diverse populations. Although our findings suggest that the cognitive-behavioral treatment programs used in the two clinics appear to apply across two economically different groups of children with anxiety disorders, some rather strong caveats are in order. Past findings suggest that the income levels in both clinics is still rather high compared to those found in community clinics (e.g., Southam-Gerow et al., 2003; Taylor et al., 2003). Indeed, our own post hoc analyses suggested that although the income levels differed between these two RC samples, neither approximated the low levels of income found in service clinics described in the literature (e.g., Southam-Gerow et al., 2003). Because research has suggested a link between income and treatment outcome and between income and psychopathology (e.g., Atkins, Graczyk, & Frazier, 2003; Attar, Guerra, & Tolan, 1994; Costello, Compton, Keeler, & Angold, 2003; Kim-Cohen, Moffitt, & Caspi, 2004), identifying ways of increasing representation of low-income families in clinical trials research represents an important direction.

A similar argument could be made concerning increasing ethnic minority participation in clinical research studies, a long-term problem in the child anxiety literature (e.g., Safren et al., 2000; Southam-Gerow & Chorpita, in press). On the positive side, the participation of Hispanic families at the Miami clinic resembled the population of Miami. However, participation of African American youth was not representative of the larger metropolitan communities at either clinic. Thus, there remains room for improvement for these (and perhaps other) RCs in terms of minority family participation. Our conclusions, thus, are consistent

with Huey's (in press) recent review: Although there is reason for optimism concerning the positive effects of treatments for ethnic minority youth, work remains with the goal to make treatments (and therapists) more culturally competent (e.g., Cunningham, Foster, & Henggeler, 2002; Sue, 1998).

Diversification with regard to parental psychopathology may also represent an important direction. Our findings suggest that these two anxiety-disorder RCs were homogenous with regard to maternal depression. In general, the BDI scores of mothers in this study were below clinical levels ( $ms < 9$ ). This suggests that parents who attend RCs may be different from those found in studies of service clinics (Hammen et al., 1999; Swartz et al., 2005). For example, Swartz et al. found that 61% of mothers of youth in a pediatric mental health clinic had a diagnosable disorder themselves, most commonly depression and anxiety. Thus, future studies should seek to include more parents with psychological problems of their own to examine the ramifications of maternal psychopathology on treatment development and adaptation as the field moves toward dissemination research.

In addition, increasing the number of youth with diverse comorbidities in research studies represents another key way to decrease homogeneity of current RC samples. Although some have argued that comorbidity is common in child anxiety RCs and does not affect outcomes (e.g., Verduin & Kendall, 2003), this study and others (e.g., Southam-Gerow et al., 2003; Taylor et al., 2003) suggest that the presence of certain comorbid disorders (e.g., oppositional defiant disorder) is rare in anxiety-disorder RCs. This is unfortunate because recent studies with depressed youth have demonstrated the importance of including youth with comorbid internalizing and externalizing disorders (e.g., Rohde et al., 2004; Weiss et al., 2003). For example, in the Rohde et al. study, a sample of youth diagnosed with comorbid major depressive disorder and conduct disorder were treated with an EBT for major depressive disorder (Coping With Depression program; Lewinsohn, Clarke, & Hops, 1990). The treatment was effective at reducing depressive symptoms (compared to a control, life-skills intervention) but not for conduct disorder symptoms. If youth commonly experience high levels of both internalizing and externalizing psychopathology, we need to test treatments with those youth.

In sum, because client differences may influence treatment outcomes, the client homogeneity we found in these RCs suggests that EBTs may require adaptation when tested with increasingly diverse client samples (cf. Schoenwald & Hoagwood, 2001; Southam-Gerow, 2005; Weisz, 2000). Further, differences at the client level may not be the only reason for possible adaptation of EBTs. Many researchers have highlighted possible differences at the therapist, agency, and sys-

tem level between RCs and service clinics. Because discussion of such adaptations is beyond the scope of this article, the interested reader is referred to other sources for extended discussion (e.g., Chorpita & Nakamura, 2004; Schoenwald & Hoagwood, 2001; Southam-Gerow, Austin, & Hershberger, in press).

Limitations of this study include a lack of assessment of other potentially relevant domains of difference between the samples, such as broader measurement of parental psychopathology (e.g., paternal psychopathology measures, measuring other problem areas). The study would have been strengthened if the measurement in the two trials had been completely uniform. Finally, this study was limited to an examination of client-level differences. The two clinics would likely be similar across other potentially relevant variables (e.g., provider training, agency financing). Still, these other domains of potential differences may be relevant to expanding the external validity of the treatment literature and merit exploration in future studies (e.g., Schoenwald & Hoagwood, 2001; Southam-Gerow et al., in press).

## References

- Achenbach, T. M. (1991). *Manual for the Child Behavior Checklists/4-18 and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Atkins, M. S., Graczyk, P. A., & Frazier, S. L. (2003). Toward a new model for promoting urban children's mental health: Accessible, effective, and sustainable school-based mental health services. *School Psychology Review, 32*, 503-514.
- Attar, B. K., Guerra, N. G., & Tolan, P. H. (1994). Neighborhood disadvantage, stressful life events, and adjustment in urban elementary school children. *Journal of Clinical Child Psychology, 23*, 391-400.
- Beck, A. T., & Beamesderfer, A. (1974). Assessment of depression: The Depression Inventory. In P. Pichot (Ed.), *Psychological measurements in psychopharmacology* (pp. 151-169). Paris: Karger.
- Beck, A. T., Steer, R. A., & Garbin, M.G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*, 77-100.
- 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.
- 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 & Adolescent Psychiatry, 37*, 906-914.
- Chorpita, B. F., & Nakamura, B. J. (2004). Four considerations for dissemination of intervention innovations. *Clinical Psychology: Science & Practice, 11*, 364-367.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

- Costello, E. J., Compton, S. N., Keeler, G., & Angold, A. (2003). Relationships between poverty and psychopathology: A natural experiment. *Journal of the American Medical Association, 290*, 2023–2029.
- Crawford, A. M., & Manassis, K. (2001). Familial predictors of treatment outcome in childhood anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*, 1182–1189.
- Cunningham, P. B., Foster, S. L., & Henggeler, S. W. (2002). The elusive concept of cultural competence. *Children's Services: Social Policy, Research, & Practice, 5*, 231–243.
- Flannery-Schroeder, E., Suveg, C., Safford, S., Kendall, P. C., & Webb, A. (2004). Comorbid externalizing disorders and child anxiety treatment outcomes. *Behaviour Change, 21*, 14–25.
- 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 & Adolescent Psychiatry, 38*, 64–71.
- Hinshaw, S. P. (2002). Intervention research, theoretical mechanisms and causal processes related to externalizing behavior patterns. *Development and Psychopathology, 14*, 789–818.
- Hoagwood, K., Burns, B. J., & Weisz, J. R. (2002). A profitable conjunction: From science to service in children's mental health. In B. J. Burns & K. Hoagwood (Eds.), *Community treatment for youth: Evidence-based interventions for severe emotional and behavioral disorders* (pp. 327–338). New York: Oxford University Press.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65–70.
- Huey, S. J. (in press). Empirically supported psychosocial treatments for ethnic minority youth. *Journal of Clinical Child and Adolescent Psychology*.
- Jaccard, J., & Guilamo-Ramos, V. (2002a). Analysis of variance frameworks in clinical child and adolescent psychology: Advanced issues and recommendations. *Journal of Clinical Child and Adolescent Psychology, 31*, 278–294.
- Jaccard, J., & Guilamo-Ramos, V. (2002b). Analysis of variance frameworks in clinical child and adolescent psychology: Issues and recommendations. *Journal of Clinical Child and Adolescent Psychology, 31*, 130–146.
- Kazdin, A. E., & Nock, M. K. (2003). Delineating mechanisms of change in child and adolescent therapy: Methodological issues and research recommendations. *Journal of Child Psychology & Psychiatry, 44*, 1116–1129.
- 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 & Clinical Psychology, 60*, 733–740.
- Kendall, P. C. (1994). Treating anxiety disorders in children: Results of a randomized clinical trial. *Journal of Consulting & Clinical Psychology, 62*, 100–110.
- Kendall, P. C., Flannery-Schroeder, E., & Panichelli-Mindel, S. M. (1997). Therapy for youths with anxiety disorders: A second randomized clinical trial. *Journal of Consulting & Clinical Psychology, 65*, 366–380.
- Kendall, P. C., Hollon, S., Beck, A., Hammen, C., & Ingram, R. (1987). Recommendations regarding the Beck Depression Inventory. *Cognitive Therapy & Research, 11*, 289–299.
- Kendall, P. C., Marrs-Garcia, A., Nath, S., & Sheldrick, R. C. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Consulting & Clinical Psychology, 67*, 285–299.
- Kim-Cohen, J., Moffitt, T. E., & Caspi, A. (2004). Genetic and environmental processes in young children's resilience and vulnerability to socioeconomic deprivation. *Child Development, 75*, 651–668.
- Kovacs, M. (1992). *Children's Depression Inventory manual*. North Tonawanda, NY: Multi-Health Systems.
- Kraemer, H. C., Stice, E., Kazdin, A., Offord, D., & Kupfer, D. (2001). How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *American Journal of Psychiatry, 158*, 848–856.
- Leaf, P. J., Alegria, M., Cohen, P., Goodman, S. H., Horwitz, S. M., Hoven, C. W., et al. (1996). Mental health service use in the community and schools: Results from the four-community MECA Study. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*, 889–897.
- Lewinsohn, P. M., Clarke, G. N., & Hops, H. (1990). Cognitive-behavioral treatment for depressed adolescents. *Behavior Therapy, 21*, 385–401.
- McCauley, E., Mitchell, J. R., Burke, P., & Moss, S. (1988). Cognitive attributes of depression in children and adolescents. *Journal of Consulting & Clinical Psychology, 56*, 903–909.
- Persons, J. B., & Silberschatz, G. (1998). Are results of randomized controlled trials useful to psychotherapists? *Journal of Consulting & Clinical Psychology, 66*, 126–135.
- Reynolds, C. R. (1980). Concurrent validity of What I Think and Feel: The Revised Children's Manifest Anxiety Scale. *Journal of Consulting & Clinical Psychology, 48*, 774–775.
- Reynolds, C. R. (1981). Long-term stability of scores on the Revised Children's Manifest Anxiety Scale. *Perceptual and Motor Skills, 53*, 702.
- Reynolds, C. R., & Richmond, B. O. (1985). *Revised Children's Manifest Anxiety Scale (RCMAS): Manual*. Los Angeles: Western Psychological Services.
- Rogers, J. L., Howard, K. I., & Vessey, J. T. (1993). Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin, 113*, 553–566.
- 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 & 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.
- Silverman, W. K. (1991). *Anxiety Disorders Interview Schedule for Children*. Albany, NY: Graywind.
- Silverman, W. K., & Albano, A. M. (1996). *Anxiety Disorders Interview Schedule for DSM-IV: Child version*. San Antonio, TX: Psychological Corporation.
- Silverman, W. K., Kurtines, W. M., Ginsburg, G. S., Weems, C. F., Lumpkin, P. W., & Carmichael, D. H. (1999). Treating anxiety disorders in children with group cognitive-behavior therapy: A randomized clinical trial. *Journal of Consulting & Clinical Psychology, 67*, 995–1003.
- Silverman, W. K., Kurtines, W. M., Ginsburg, G. S., Weems, C. F., Rabian, B., & Serafini, L. T. (1999). Contingency management, self-control, and education support in the treatment of childhood phobic disorders: A randomized clinical trial. *Journal of Consulting & Clinical Psychology, 67*, 675–687.
- Siqueland, L., Crits-Christoph, P., Gallop, B., Gastfriend, D., Lis, J., Frank, A., et al. (2002). Who starts treatment: Engagement in the NIDA collaborative cocaine treatment study. *American Journal on Addictions, 11*, 10–23.
- Southam-Gerow, M. A. (2004). Some reasons that mental health treatments are not technologies: Toward treatment development

- and adaptation outside labs. *Clinical Psychology: Science and Practice*, 11, 186–189.
- Southam-Gerow, M. A. (2005). Using partnerships to adapt evidence-based mental health treatments for use outside labs. *Report on Emotional & Behavioral Disorders in Youth*, 5, 58–60, 77–79.
- Southam-Gerow, M. A., Austin, A. A., & Hershberger, 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, CA: Sage.
- Southam-Gerow, M. A., & Chorpita, B. F. (in press). Anxiety disorders. In E. J. Mash & R. A. Barkley (Eds.), *Assessment of childhood disorders* (4th ed.). New York: Guilford.
- 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., Weisz, J. R., & Kendall, P. C. (2003). Youth with anxiety disorder in research and service clinics: Examining client differences and similarities. *Journal of Clinical Child and Adolescent Psychology*, 32, 375–385.
- Steer, R. A., Beck, A. T., Riskind, J. H., & Brown, G. (1986). Differentiation of depressive disorders from generalized anxiety by the Beck Depression Inventory. *Journal of Clinical Psychology*, 42, 475–478.
- Sue, S. (1998). In search of cultural competence in psychotherapy and counseling. *American Psychologist*, 53, 440–448.
- Swartz, H. A., Shear, M. K., Wren, F. J., Greeno, C. G., Sales, E., Sullivan, B. K., et al. (2005). Depression and anxiety among mothers who bring their children to a pediatric mental health clinic. *Psychiatric Services*, 56, 1077–1083.
- Taylor, A. A., Francis, S. E., Chorpita, B. F., Southam-Gerow, M. A., & Lam, C. (2003, November). *Examining differences between publicly and privately referred youth at a university-based clinic*. Poster presented at the Association for the Advancement of Behavior Therapy Convention, Boston, MA.
- U.S. Department of Commerce. (1992). *1990 Census of population: General population characteristics: Metropolitan areas*. Washington DC: Author.
- U.S. Department of Commerce. (1993). *1990 Census of population: Social and economic characteristics: Metropolitan areas*. Washington DC: Author.
- Verduin, T. L., & Kendall, P. C. (2003). Differential occurrence of comorbidity within childhood anxiety disorders. *Journal of Clinical Child & Adolescent Psychology*, 32, 290–295.
- Webster-Stratton, C. (1990). Enhancing the effectiveness of self-administered videotape parent training for families with conduct-problem children. *Journal of Abnormal Child Psychology*, 18, 479–492.
- 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., 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.

Received April 14, 2005

Accepted July 17, 2006