

School-Based Mental Health Programs for Students Who Have Emotional Disturbances: Academic and Social-Emotional Outcomes

Krista Kutash · Albert J. Duchnowski ·
Amy L. Green

Published online: 21 July 2011
© Springer Science + Business Media, LLC 2011

Abstract Most schools offer some type of school-based mental health service to students, and there is a growing body of empirically rigorous studies examining the effects on academic and mental health outcomes for students. However, students classified as having emotional disturbances are under-represented in these studies. Using a convenience sample of four different types of school-based mental health programs, changes in achievement levels and social and emotional functioning in youth with emotional disturbances served in special education ($n = 148$) were examined. Longitudinal results reveal there was improvement in either the emotional or social functioning of these youth in all four programs, while results for improvement in achievement levels were less consistent. Results reveal that in the targeted sample of programs with intensive or multifaceted services, there was improvement in academic as well as social and emotional functioning in these youth. Implications for school-based mental health services are discussed.

Keywords School-based mental health services · Special education · Emotional disturbance

Introduction

There continues to be increasing concern about the need to improve the achievement and social-emotional functioning of youth who are identified as having emotional disturbances (ED) and served in school district special education programs. The federal definition of ED refers to one of the several disability categories within special education. Specifically, the current federal educational definition of ED is as follows: Emotional disturbance refers to a condition exhibiting one or more of the following characteristics over a long period of time and to a marked degree that adversely affects a child's educational performance: (a) An inability to learn that cannot be explained by intellectual, sensory, or health factors, (b) An inability to build or maintain satisfactory interpersonal relationships with peers and teachers, (c) Inappropriate types of behavior or feelings under normal circumstances, (d) A general pervasive mood of unhappiness or depression, or (e) A tendency to develop physical symptoms or fears associated with personal or school problems. Emotional disturbance includes schizophrenia. The term does not apply to children who are socially maladjusted, unless it is determined that they have an emotional disturbance under paragraph (c)(4)(i) of this section. (U.S. Government, 2004, Code of Federal Regulation, Title 34, Section 300.8(c)(4)(i) and (ii)). Approximately 131,000 students are in special education under the category of ED, representing less than 1% of the current student population but 7.7% of students in special education (US Department of Education, Annual Report to Congress, 2007).

K. Kutash (✉)
Department of Child and Family Studies, University of South
Florida, 13301 Bruce B. Downs Blvd. MHC 2335, Tampa, FL
33612, USA
e-mail: Kutash@usf.edu

A. J. Duchnowski
Department of Child and Family Studies, University of South
Florida, 13301 Bruce B. Downs Blvd. MHC 2334, Tampa, FL
33612, USA
e-mail: duchnows@usf.edu

A. L. Green
Department of Child and Family Studies, University of South
Florida, 13301 Bruce B. Downs Blvd. MHC 2336, Tampa, FL
33612, USA
e-mail: algreen@usf.edu

The concern for this population of students is well documented. For example, in a series of studies beginning in the 1990s and continuing to the present, Blackorby and Wagner (1996), Wagner, Kutash, Duchnowski, Epstein, and Sumi (2005b), and Wagner et al. (2006) concluded the outcomes for students who have ED were the poorest compared with other disability groups. Results from Wagner's studies revealed the average academic achievement for these students was below the 25th percentile; they had the highest dropout rate compared with all disability groups; and half of these students were involved with the justice system 2 years after separating from school. The poor academic functioning of students who have ED is further documented through the results of a meta-analysis in which an overall effect size of $-.64$ indicated significant deficits in the academic achievement of these students (Reid, Gonzalez, Nordness, Trout, & Epstein, 2004). As they age and transition into adulthood, youth with ED have the lowest paying jobs, the fewest instances of full-time employment, and the highest risk of entering the adult mental health system (Kessler, Chiu, Demler, & Walters, 2005; Wagner, Kutash, Duchnowski, & Epstein, 2005a). Even with these documented negative outcomes, while in school, these youths receive minimal amounts of mental health services and very little support for their families (Wagner et al., 2006).

In contrast, the potential for school-based mental health (SBMH) interventions to improve the functioning of youth who have ED and need mental health services has been long recognized (Kutash, Duchnowski, & Lynn, 2006). The Surgeon General's Report (Department of Health and Human Services [DHHS], 1999) and the President's New Freedom Commission on Mental Health (2003) identified school as the ideal location to meet the mental health needs of the nation's youth and to increase the capacity of the children's mental health service system. The No Child Left Behind act [NCLB] (2002) also promotes the provision of mental health services in schools, especially through collaboration between schools and community agencies. Policies and practices aimed at increasing the capacity to provide mental health services to children and youth are important given the estimate that only one-third of children who need mental health services actually receive them (US Public Health Service, 2000). Epidemiological studies assessing mental health needs and receipt of services in youth demonstrate the school system is the *de facto* mental health system for youth, with more than 70% of youth who receive mental health services receiving them at their schools (Burns et al., 1995; Leaf et al., 1996). Consequently, efforts to enhance SBMH programs are a sound strategy for policy makers and providers.

While there has been a strong emphasis at the federal policy level on the development of effective mental health

interventions that could be implemented in schools, progress in the implementation of SBMH services is mixed. For example, at present, no clear model of best practice has been established to guide the development and implementation of a comprehensive SBMH program (Kutash et al., 2006; Paternite, 2005). In addition, while there has been considerable activity over the last two decades by researchers and program developers aimed at increasing SBMH services in the nation's schools, the resulting research base is slim. In a recent review of this literature, Hoagwood et al. (2007) identified over 2,000 articles dealing with SBMH. However, only 64 of these 2,000 articles (3%) reported using rigorous empirical designs that allowed for an evaluation of the efficacy of the interventions. While the results from these studies indicate there are some efficacious treatments available, especially aimed at improving social competence and decreasing aggression and problem behavior, only 24 of the 64 studies reviewed (37%) had even a rudimentary measure of academic achievement. Further, only 15 programs were found to be dually effective at meeting both the academic and behavioral needs of youth. These programs were highly intensive and complex, targeted children at risk for antisocial behavior, and involved interventions at multiple levels across multiple contexts (i.e., home, classroom, and school) over an extended period of time (at least a year). However, none of the studies reviewed by Hoagwood et al. had samples that included children identified as having ED and served in special education programs, although one study did include participants who had ADHD. For students with ED and served in special education, who may be most in need of effective services, there is a critical gap in the knowledge base examining academic and mental health interventions and their outcomes (Hoagwood et al., 2007; Rones & Hoagwood, 2000; Simpson, 2004).

Focus on Youth Who Have ED and Who are in Special Education Programs

The over-arching goal of this paper is to contribute to the knowledge base on SBMH services for those children who are identified as having ED and who are placed in school district special education programs for ED. We are encouraged in this effort by the recent recommendation of some researchers in the SBMH community to re-focus mental health interventions in schools on the core function of schools, which is learning (e.g., Atkins, Hoagwood, Kutash, & Seidman, 2010; Cappella, Frazier, Atkins, Schoenwald, & Glisson, 2008). Historically, mental health services have focused more on general behavior than on behavior as it relates to classroom behavior or learning. Atkins and colleagues suggest "A comprehensive approach is needed in which interventions address multiple

dimensions of the child's behavior such as relationships with peers, teachers, and parents, and children who have ED and who are in special education programs are included" (Atkins et al., 2010, p. 43). They further recommend that support for teachers, the primary change agent, in classroom management and instruction should be the major vehicle for implementing effective interventions. In addition, rigorous evaluation of multiple domains of outcomes that include emotional functioning and academic functioning is also called for (Adelman & Taylor, 2006; Hoagwood et al., 2007). It is only through this comprehensive approach that outcomes for this group of youth will be improved.

While the efforts to re-focus SBMH interventions face formidable challenges, the plan to rigorously evaluate these comprehensive services poses some specific problems for researchers and evaluators. Over 15 years ago, Knapp (1995) identified similar problems facing researchers in their quest to evaluate the emerging integrated educational service programs aimed at children at risk for academic failure, children in poverty, children with serious health conditions, and children in the child welfare system. He proposed that before rigorous evaluations were attempted, much work was needed, which included strong conceptualization of program models, rich, comparative description, and a collaborative "bottoms up" approach. Since these integrated programs were multi-disciplinary, he cited different language across disciplines and lack of agreement on the nature of the independent and dependent variables as barriers that needed to be resolved (Knapp, 1995).

We propose that a similar case can be made for the examination and evaluation of SBMH services for students who have ED and who are educated in special education programs. In addition to the lack of consensus on a conceptual model describing best practices in SBMH, the descriptive information available for SBMH service programs is weak. While information about this group of students who have disabilities is emerging from large scale studies such as the Special Education Elementary Longitudinal Study (SEELS) and the National Longitudinal Transition Study-2 (NLTS2) (Newman, Wagner, Cameto, Knokey, & Shaver, 2010), the field has little systematic, empirical information about the nature of ED in students, how the complex and multiple characteristics of the students moderate and mediate outcomes, and what program components are causally related to improved outcomes in these students (Adelman & Taylor, 2006; Evans & Weist, 2004; Hoagwood et al., 2007; Kratochwill, Clements, & Kalyon, 2007). The failure to adequately describe current practice in SBMH (usual care) limits the ability of researchers to design rigorous studies that use experimental designs (Bickman, 2000; Garland et al., 2010; Hoagwood & Kolko, 2010).

What We Know About SBMH Program Structures

While our specific knowledge about mental health services for children who have ED and who are served in special education may be limited, there is information about the implementation of SBMH to students in general. Two surveys have reported on the organization of SBMH services using a nationally representative sample of schools and districts or states. These two surveys are the School Health Policies and Program Study 2000 (Brener, Martindale, & Weist, 2001) and School Mental Health Services in the United States 2002–2003 (Foster et al., 2005). From these surveys, we know that a majority of schools provide individual counseling (76%), case management (71%), and group counseling (68%). Providing support to parents has been found to be very difficult and is rarely provided (Wagner et al., 2006).

We also know that there are several organizational structures that schools use to provide SBMH services to their students. These structures vary from the commonly used method in which school districts hire their own staff to provide mental health services, to formal contracts with community mental health centers for services, to comprehensive and integrated systems in which districts collaborate with multiple community agencies to provide mental health services to students. A continuum of organizational arrangements to provide SBMH services has been developed (Policy Leadership Cadre for Mental Health in Schools, 2001; Weist, 1997) and includes the following configurations: (1) school districts exclusively use district-financed professional staff to provide traditional mental health services; (2) school districts only use outside agencies for the provision of mental health services or they contract with an outside agency to augment mental health services through a combination of district staff and staff from community agencies; (3) school districts operate and finance their own mental health units or clinics that provide services, training, and/or consultation to schools; (4) schools use curriculum-based programs to enhance social and emotional functioning and reduce barriers to learning, interventions tend to be teacher-led and prevention-oriented; and (5) comprehensive, multifaceted, and integrated approaches are used in which districts bring multiple partners (e.g., community-based organizations) together to provide a full spectrum of services for children and youth with mental health needs. This approach would include such models as Systems of Care (Stroul & Friedman, 1994), in which an array of mental health and wraparound services are provided to children with mental health problems and their families via partnerships among various child-serving systems.

Through an analysis of the data from the two national surveys of SBMH (Brener et al., 2001; Foster et al., 2005),

we can determine the approximate number of school districts that use each of the organizational arrangements. About 33% of school districts use the first option, exclusive use of school personnel to provide mental health services, while 25% exclusively use staff from community agencies. Additionally, about 55% of schools report having contracts with outside agencies to provide staff to augment mental health services provision through a combination of district staff and staff from community agencies. Only about 2% of districts report using district operated clinics, while 59% of districts report using curriculum-based programs to enhance social and emotional learning. While there is no accurate estimate of how many schools have comprehensive and integrated programs, there are currently over 160 communities and tribal nations that have received grants from the Substance Abuse and Mental Health Services Administration (SAMHSA) to establish such programs (Blau, Huang, & Mallery, 2010). The purpose of this study is to begin to examine the outcomes for youth served in these various models.

Purpose of the Current Study

Even with the current efforts to build and expand SBMH models and programs, youth who have emotional and behavioral challenges and who are served in special education remain a critically understudied group. In addition, there are few empirically supported integrative models available to overcome the poor outcomes of students who have ED. The field faces a major challenge in developing a unified research agenda on effective school models that will support the learning and behavioral health for this group of students (Atkins et al., 2010). The purpose of the current study is to begin the conversation on building a research agenda for this group of students, and to provide a rich description of procedures schools use to provide SBMH services to youths who have ED and who are educated in special education programs, and to examine changes over time in emotional and academic functioning. In this study, we examined a purposive sample of four different SBMH programs that were representative of the most frequent types of structures used by school districts to provide mental health services to students. The description of the youths served, the program features, the mental health services provided, and longitudinal mental health and academic outcomes are presented. This information is intended to inform the field and contribute to future research aimed at rigorous evaluation of the effectiveness of SBMH for children who qualify for special education programs. The study will establish some initial benchmarks on the changes to be expected in the emotional and behavioral and academic functioning over time when youth in special education receive mental health services.

The specific questions driving this study are as follows: (1) what are the major structural and procedural aspects of four distinct SBMH programs for youth served in special education?; (2) What changes occur in emotional, behavioral, and academic functioning over time?; and (3) what types and amounts of mental health services are delivered?

Method

SBMH Programs and Participants

Program Inclusion

The sample of programs included in the current study is a purposive sample selected for convenience, in that programs were not selected from a pool of eligible sites. Programs selected for the current study were known to the authors through the literature or conference presentations, and all of them served youth who: (1) met criteria that identified them as having an emotional disturbance, (2) had been placed in a special education setting due to ED, (3) had completed standardized measures indicating level of emotional disturbance and impairment upon entry into the program, and (4) had completed these measures again at either 9- or 12-months after entry. The authors toured each of the programs, collected written information or reports, and conducted informal interviews with program staff. Program administrators from all of the selected programs agreed to share with researchers existing data sets that described the number of youth served by the program and any standardized measures collected longitudinally on youth in their programs.

The four programs selected for inclusion in the current study served youth who had an individual educational plan (IEP) and were educated in public schools where school-based mental health services were provided as specified on the IEP. These programs were selected to reflect the various structures or models of how school systems provide school-based mental health services to youth in special education. The four programs selected include: (1) a comprehensive and multifaceted approach to delivering mental health services that involves the school working with several community agencies to provide support and services—referred to in the current study as the Integrated Program; (2) a program in which the school district hires their own professional staff to provide mental health services, along with classroom-based curricula and programs to reduce psycho-social barriers to learning—referred to in the current study as the Milieu Program; (3) a program in which mental health services are supplied by school personnel who provide counseling to youth—referred to in the current study as Pull-Out Program 1; and (4) a program in

which mental health services are provided through contracts with mental health agency staff who come into the school to provide individualized and group therapy—referred to in this study as Pull-Out Program 2. A description of the four programs and study participants from each of these programs follows and is summarized in Table 1.

Integrated Program

This program is an example of the most comprehensive school-based mental health service model described in the literature (Policy Leadership Cadre for Mental Health in Schools, 2001; Weist, 1997). Supported by its state's initiative to develop multi-agency systems of care to serve children in need of mental health services and their families and a SAMHSA grant, the Integrated Program is a partnership of schools and community provider agencies, including a community mental health center. The goal of this program is to establish a school-based continuum of mental health services and supports that employs a three-tiered model of prevention, targeted intervention, and intensive intervention. The participants in the current study

are at the third tier (i.e., identified as having serious emotional disturbances and receiving special education and intensive mental health services).

Delivery of school-based mental health services is accomplished through a partnership with school personnel, the family, and the Student Service Team (SST). The SST is composed of a service coordinator who acts as a case manager, a family liaison, and an intervention specialist, all of whom are employees of a community mental health center but are housed on the school campus. The service coordinator, a bachelor's level professional who received training in service coordination and team facilitation strategies, facilitates wraparound meetings linking the family with natural supports and formal resources in the community. The family liaison is a parent of a child with an emotional disability who serves as a peer mentor to family members and assists in building local and regional family support networks. They received certification by the state office for family leadership. The intervention specialist is a master's level mental health clinician who received additional training in functional behavioral assessment and the development of behavior intervention plans.

Table 1 Overview of four programs and participant characteristics

	Integrated (<i>n</i> = 50)	Milieu (<i>n</i> = 51)	Pull-Out 1 (<i>n</i> = 23)	Pull-Out 2 (<i>n</i> = 24)
Model	Comprehensive and integrated across child-service agencies	District operated curriculum enhanced	District operated and contracts with MH clinics	District operated with contracts with MH clinics and local county agency
Mental health (MH) service delivery	Wrap-around approach with treatment planning within a school setting	Milieu therapy with two MH providers and two Special Educators in each classroom	“Pull-Out” Program for traditional delivery of MH talk therapy	“Pull-Out” Program for delivery of traditional MH talk therapy
Funding	IDEA Medicaid Federal Grant	IDEA Medicaid	IDEA	IDEA County MH funds
Grades	K–12th	1st–12th	6th–12th	6th–12th
Mean age (SD)	10.90 (3.29)	11.02 (2.82)	14.43 (1.90)	14.96 (1.68)
Age range	4–17 years	6–16 years	10–17 years	12–19 years
% Male	80.0	80.4	73.9	70.8
% Black	2.0	11.8	47.8	91.7
% White	90.0	74.5	30.4	8.3
% Hispanic ^a	0.0	3.9	17.4	0.0
% Multiracial	0.0	7.8	4.3	0.0
% Other	8.0	2.0	0.0	0.0
% Poverty at baseline ^b	60.0	23.0 ^c	26.9	48.3

^a For the Integrated Program, the reported percentage of Hispanic participants is likely underestimated. While 4% (*n* = 2) of participants did indicate they were of Hispanic ethnicity, “Hispanic” was not an available response option when asked to report race

^b Family income was compared with poverty thresholds for the baseline year for the Integrated, Pull-Out 1, and Pull-Out 2 Programs

^c This percentage is an estimate for poverty based on the average number of all children from schools in the Milieu program eligible for free/reduced lunch

For each child receiving intensive intervention, a wraparound planning meeting is held (Eber, Osuch, & Reddit, 1996). The wraparound team consists of the SST joined by the child and his or her family, school personnel, service providers, and any natural supports that the family feels could have a positive impact on the child or family (e.g., extended family members, clergy, and other peers). The wraparound meeting follows an eight step process (Robbins & Armstrong, 2005) that includes a review of child and family strengths, identification of needs, development of actions and assignment of tasks to members, and culminates in a summary document prepared by the team facilitator. This document serves as the foundation for future team meetings, continued planning, and monitoring outcomes. The wraparound plan addresses all the important domains in the child's life such as school, family, and community. A range of mental health services for the children are available that include individual and group therapy, case management, medication management, crisis management, and family therapy. Individual behavioral aides are also available.

Students in the Integrated Program

An administrative data set was provided to researchers that reflected youth enrolled in the Integrated Program between September 1999 and June 2004. A review of this data set revealed a total of 632 youth were enrolled in the program during this time period. Of these students, 93 had either an IEP upon admission or had an IEP instituted during the first 6 months of the program and were therefore eligible for inclusion in the current study. An analysis of the data revealed no statistically significant differences between students with an IEP upon admission ($n = 65$) and students with an IEP instituted during the first 6 months of the program ($n = 28$) with regard to gender, race, or ethnicity. Of the 93 youth with an IEP, 50 (54%) had complete data at baseline and follow-up and were included in the current study. In regard to differences in race, gender, and age, students with complete data differed from those without complete data ($n = 43$) in that they were more likely to be male ($\chi^2 = 4.29$, $df = 1$, $P < .05$). Further, these students differed from the total population of students enrolled in the Integrated Program ($n = 632$) in that they were more likely to be non-white ($\chi^2 = 21.57$, $df = 4$, $P < .001$) and male ($\chi^2 = 4.33$, $df = 1$, $P < .05$). Data for Integrated Program participants were collected at baseline and again 6- and 12-months after entry into the program.

Milieu Program

This program is an example of a school financed program in which school districts hire their own professional staff to

provide mental health services, along with classroom-based curricula and programs to reduce psycho-social barriers to learning. In this case, 13 school districts formed an intermediate unit that could provide services for students with serious ED more efficiently than an individual district. The program is viewed as the most restrictive placement for the most seriously disturbed youth in the intermediate unit catchment area in this region. This program uses a curriculum-based approach which may be considered part of a psycho-educational model; however, the admission policies and funding are driven by a medical model. The mental health treatment component is funded by the state's version of Medicaid, and the educational component of the program is paid for by each referring district using local and federal special education funds. All children in the program are eligible for medical assistance under the state's Medicaid guidelines. The program serves all age levels and is implemented at 19 schools (seven high schools, seven middle schools, and five elementary schools) across a three county area.

Admission to this program is determined by an Inter-agency Team composed of representatives from the county MH/MR agency, the school district, the family, and the intermediate unit Interagency Coordinator and requires a psychiatric evaluation. The Milieu Program is intended to be an intensive intervention that will stabilize a student and transition him to a less restrictive setting. The intensity is illustrated by the presence of four adults in each of the self-contained classrooms, located in regular schools, which serve about 12 students who have ED. The lead person is a master's level mental health professional (M.S. in psychology or counseling, M.S.W.). The other staff includes a bachelor's level mental health technician, a certified special education teacher, and an education para-professional (high school graduate). An integrated treatment plan is implemented by the staff. They are supported by school psychologists, psychiatric consultants, program supervisors, doctoral level clinical psychologists, and a case manager who maintains contact with the home and school. All of these professionals are school employees.

The activities that comprise the mental health component of the Milieu Program are influenced by the regulations required for reimbursement from the state Medicaid Office by documenting billable hours. Typically, between 3 and 6 h of service are billed each day for every student in the program. Students receive "lump services" while in the program that include group and individual psychotherapy, and Milieu therapy that includes academic activities. The treatment plan developed for each student is goal oriented with extensive daily written progress notes for which the mental health counselor is responsible. These are reported daily for billing purposes as well as for program monitoring. There are frequent staffing meetings for each student

as well as formal evaluative summaries that correspond to the quarters of the school year. The IEP includes behavioral and academic goals and the curriculum and instruction are coordinated by the special education teacher. Extensive use is made of curriculum-based programs to enhance social and emotional learning.

Students in the Milieu Program

An administrative data set was provided to the researchers and reflected youth who were enrolled in the Milieu Program between 2003 and 2008. Altogether there were 333 youth enrolled in the program during this time period. Of these youth, 230 (69%) had an IEP upon admission and had an emotional disorder, making them eligible for the study. Fifty-one of these students (22%) had complete data at baseline and follow-up and were included in the current study. In regard to race, gender, and age, students with complete data differed from those without complete data ($n = 179$), in that they were more likely to be male ($\chi^2 = 4.451$, $df = 1$, $P = .035$) and younger ($t(228) = 4.077$, $P < .001$). Similarly, these students differed from the total population of students enrolled in the Milieu Program ($n = 230$) in that they were more likely to be male ($\chi^2 = 8.257$, $df = 1$, $P = .004$) and younger ($t(331) = 3.782$, $P < .001$). Data for participants were collected at baseline and again between 9- and 12-months after enrollment, depending upon the time of the school year the student was admitted.

Pull-Out Program 1

This program is a school district special education center that exclusively serves students with ED. It has approximately 120 students in grades K–12; however, only students in grades 6–12 participated in the study. A principal, two special education specialists, and twenty teachers and classroom aides provide the educational component of the program. The center is located on the campus of the county children's services agency, which affords an opportunity for some collaboration and co-location of mental health services. Pull-Out Program 1 is an example of a school financed program in which school support service staff are augmented by staff from a community agency. The school district employs a school psychologist, a social worker, and a case manager and provides stipends to four clinical psychology interns who provide 6 h of counseling service per week. The county agency provides three mental health counselors, a nurse, and a consulting psychiatrist to augment the mental health services component of the program. Mental health services are provided in offices in the school and include individual and group therapy, case management, medication management, and family therapy.

Psycho-education activities are provided in classrooms. Informal interviews with staff revealed both the school district and the community agency staff deliver therapy that is essentially similar, with staff members describing the type of therapy they provide as eclectic.

Students in Pull-Out Program 1

Researchers utilized an administrative data set from a previously conducted randomized controlled trial that included students from Pull-Out Program 1 who entered the program between August 2005 and October 2005. Initially, 56 youth were enrolled in this trial, all of whom had a current IEP. Randomization for the RCT resulted in 30 youth assigned to the experimental condition and 26 youth assigned to the comparison condition. Youth in the experimental condition received an experimental intervention in addition to usual program services. Participants included in the current study are youth from the comparison, or services as usual, condition. Youth in this condition received usual Pull-Out Program services, but did not receive the experimental intervention. Analyses indicated no significant differences between youth in the experimental and comparison conditions with regard to gender, race, or age. Of the 26 youth in the comparison condition, 23 had complete data at baseline and follow-up and were included in the current study. There were no significant differences between those with complete data and those without complete data ($n = 3$) with regard to gender, race, or age. Data for participants from Pull-Out Program 1 were collected at baseline and again 9 months later.

Pull-Out Program 2

This program is a school district exceptional student center that exclusively serves students who have ED. This program serves approximately 120 students in grades 6–12. The academic program is administered by a principal and an assistant principal, 16 special education teachers and teacher aides, two behavioral coaches, and an exceptional student education specialist who works with teachers on classroom management. Mental health services are provided to students through a contract with a community mental health center. Two master's level mental health counselors each provide 16 h per week of individual or group therapy in offices in the school. There is some family therapy offered and the IEP Team determines how much, if any, mental health counseling a student will receive.

Students in Pull-Out Program 2

Researchers utilized an administrative data set from a previously conducted randomized controlled trial that

included students from Pull-Out Program 2 who entered the program between August 2006 and October 2006. Initially, 59 youth were enrolled in this trial, all of whom had a current IEP upon admission. Randomization for the RCT resulted in 30 youth assigned to the experimental condition and 29 youth assigned to the comparison condition. Youth in the experimental condition received an experimental intervention in addition to usual program services. Participants included in the current study are youth from the comparison, or services as usual, condition. Youth in this condition received usual Pull-Out Program services, but did not receive the experimental intervention. Youth in the experimental and comparison conditions did not differ with regard to gender, race, or age. Of the 29 students in the comparison group, 24 students had complete data at baseline and follow-up and were included in this study. A comparison of students with complete data and those without complete data revealed significant differences in race ($\chi^2 = 6.440$, $df = 3$, $P = .04$). Youth with complete data were 91% Black and 0% Hispanic, whereas those without complete data were 75% Black and 25% Hispanic. Data for participants from this program were collected at baseline and 9-months.

Measures

Socio-Demographic Variables

Caregivers for youth from all four programs provided socio-demographic information including the student's age, gender, race, and family income. The ethnic/racial categories used to classify participants varied by program. Therefore, the racial/ethnic categories for the four programs were recoded to allow for comparison between groups. This recoding was accomplished by doing a frequency analysis of race/ethnicity for participants in each of the four programs; based on this analysis, five race/ethnicity categories were created: Black, White, Hispanic, Multi-racial, and Other (e.g., American Indian, Alaskan Native, Asian, Pacific Islander, and "other").

The method for determining participants' poverty status differed depending on program. Participants in the Integrated Program, Pull-Out Program 1, and Pull-Out Program 2 reported their annual household income. To determine whether participants from these programs fell at or below the poverty level, annual household income was compared with national poverty thresholds for the baseline year of each program. If the reported annual income was less than or equal to the national poverty threshold for that year, the participant was said to be at or below the poverty level.

Participants from the Milieu Program did not provide information on annual household income. For these participants, the percentage of students eligible for

free/reduced lunch in each program school was used as a proxy variable for poverty level. Using data from the National Center for Education Statistics, the percentage of students eligible for free/reduced lunch in each program school was averaged to represent the percentage of students from this sample at or below the poverty level.

Emotional Functioning

Depending on the program, emotional functioning was measured using the *Child Behavior Checklist (CBCL)* or the *Strengths and Difficulties Questionnaire (SDQ)*. The Child Behavior Checklist (CBCL; Achenbach, 1991) is an individually administered instrument designed to measure children's competencies and behavior problems compared with a national representative sample. The parent/caregiver provides information on 118 problem behaviors using a rating scale for how true the item is of the child now or within the past 6 months. The CBCL yields normalized T scores ($M = 50$, $SD = 10$) and percentiles on numerous scales. The Externalizing Behavior scale includes antisocial and aggressive behaviors such as stealing, truancy, fighting, and running away; the Internalizing Behavior scale includes withdrawn and anxious behaviors such as fearfulness, worrying, crying, and feelings of worthlessness. The Total Problem Behavior scale includes social problems and attention problems in addition to items from the Externalizing and Internalizing Behavior scales. A T-score above 63 is considered to be in the clinical range, and a score between 60 and 63 is considered borderline. The psychometric properties concerning reliability and validity of the CBCL have been well established and reported in several studies (Achenbach, 1991; Dedrick, Greenbaum, Friedman, Wetherington, & Knoff 1997).

The Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001) assesses behavioral problems and competencies of youth. Parents respond to items that assess their youth's behavior over the past 6 months. The SDQ is a 25-item measure that yields five domain behavior problem scores (Emotional Symptoms, Conduct Problems, Hyperactivity, Peer Problems, and Prosocial Functioning) and a total difficulties score. Each of the five domain scores is based on five items that are rated on a 3-point scale (0 = not true; 1 = somewhat true; 2 = certainly true). The total difficulties score, reported in the current study, is derived by summing all of the domain scores, except the prosocial functioning score. The total difficulties score can range from 0 to 40, with higher scores indicating greater problems. Total difficulties scores can be interpreted as follows: "normal" (0–13), "borderline" (14–16), and "abnormal" (17–40). The SDQ is a popular brief measure of psychopathology in youth with extensive documentation of adequate reliability and validity (Goodman, 2001; Mellor, 2004).

Functional Impairment

Depending on the program, functional impairment was measured using the *Brief Impairment Scale (BIS)* or the *Child and Adolescent Functional Assessment Scale (CAFAS)*. The Child and Adolescent Functional Assessment Scale (CAFAS; Hodges, 1990) is comprised of items that describe behaviors organized into eight domains of functioning (i.e., role performance in school, home and community, behavior toward others, moods/emotions, self-harmful behavior, substance use, and thinking). Parent responses to these items yield a score that describes the severity of the youth's behavior on a particular domain: 30 for severe (severe disruption or incapacitation), 20 for moderate (persistent disruption or major occasional disruption of functioning), 10 for mild (significant problems or distress), and 0 for minimal or no impairment (no disruption of functioning). The total score refers to the sum of the five subscales with a range from 0 to 150, with a higher score reflecting greater impairment. A total score of 140 or more is in the functional limitations range. Hodges and Wong (1996) have reported evidence for internal consistency, test–retest reliability, and inter-rater reliability of the CAFAS. Numerous studies have demonstrated concurrent criterion-related validity (Hodges, Doucette-Gates, & Liao, 1999; Hodges & Wong, 1996; Manteuffel, Stephens, & Santiago, 2002) and predictive validity (Hodges & Wong, 1997; Hodges, Doucette-Gates, & Kim, 2000; Hodges & Kim, 2000; Quist & Matshazi, 2000) of the CAFAS.

The Brief Impairment Scale (BIS; Bird et al., 2005) is a multidimensional scale of functional impairment containing 23 questions. Parents are asked to rate their youth on three functional domains: interpersonal relations, school/work, and self-fulfillment, based on the last 6 months. The BIS Total Score ranges from 0 to 69, with higher scores indicating greater impairment. The interpersonal and school/work subscales range from 0 to 24, and the self-fulfillment subscale ranges from 0 to 21. Scores of 14 or greater are considered to be in the clinical range. The BIS has demonstrated acceptable reliability and concurrent validity (Bird et al., 2005).

Academic Outcomes

Depending on the program, academic outcomes were measured using the *Wechsler Individual Achievement Test (WIAT)*, the *Educational Questionnaire—Revised (EQ-R)*, or the *Wide Range Achievement Test 3 (WRAT3)*. The Wechsler Individual Achievement Test (WIAT; The Psychological Corporation, 1992) is an individually administered multiple-subject comprehensive test covering reading, mathematics, language skills, and writing. It is designed for children and young adults between 5 and

19 years of age and includes eight subtests that form six composite scores. Raw scores are converted to standard scores ($M = 100$, $SD = 15$). The WIAT has demonstrated acceptable reliability and validity (Sattler, 2001).

Caregivers who completed the Educational Questionnaire-Revised (EQ-R) were asked “Which of the following best describes (your child’s) grades or school performance in the past 6 months?” To convert data from this variable to continuous, letter grades were recoded into intervals of equal size such that $A = 100$, $B = 80$, $C = 60$, $D = 40$, and $F = 20$. Responses of either “failing all or most classes” or “failing about half of his/her classes” were considered as grade average F . Data on this item were not collected from caregivers of youth who attended a school or program that did assign grades to students.

The Wide Range Achievement Test 3 (WRAT3; Wilkinson, 1993) is a norm-based standardized assessment of academic functioning and is one of the most commonly used academic assessments in research with students with emotional disturbances (Reid et al., 2004). The WRAT3 has three subtests: the reading/word decoding subtest, the spelling/written encoding subtest, and the arithmetic subtest. Scoring of the WRAT3 yields a raw score that can be converted into age-based standard scores ($M = 100$, $SD = 15$), absolute scores, percentiles, and grade equivalent scores; standard scores on the arithmetic (i.e., math) and reading/decoding words (i.e., reading) subtests are reported for the current study. Interpretation of standard scores is as follows: 130 and above (very superior), 120–129 (superior), 110–119 (high average), 90–109 (average), 80–89 (low average), 70–79 (borderline), and 69 and below (deficient). Reliability has been extensively documented, and the correlations between the WRAT3 and other achievement tests support the validity of this instrument (Burns & Kutash, 2000).

School Participation

Depending on the program, school participation was assessed using school records or caregiver report of typical absenteeism patterns. For the Milieu and Pull-Out Programs (1 and 2), school participation was measured by school records describing the number of days present at school. School participation for the Integrated Program was measured by caregiver report of the student's typical absenteeism pattern. Parents indicated how often their child was absent on average during the last 6 months; response options ranged from “absent three or more days per week” to “absent less than 1 day per month.” To convert the six absenteeism categories into continuous variables, each category was recoded into a defined attendance count. For example, a response of “absent about 2 days per week” was recoded into an integer denoting presence at school

3 days per week. Subsequently, these values were used to calculate an approximate number of school days present during a 180 day school year.

Mental Health Services Received/Delivered

For participants in the Integrated Program, receipt of mental health services was assessed during the follow-up interview using the Multi-Sector Service Contacts—Revised (MSSC-R) questionnaire (see Gyamfi et al., 2010). Caregivers were asked to indicate which of a list of 23 specific services was received by the child or the child's family in the last 6 months. Of these 23 services listed, 7 were considered for analysis in the present study: crisis stabilization services, medication treatment/monitoring services, group therapy, individual therapy, behavioral/therapeutic aide services, case management services, and family therapy. Results reflect the number of mental health services participants received over the course of 1 year in the program by combining services reported during two 6-month interviews.

The nature of the Milieu Program was such that core staff provided 6 h of integrated mental health/psycho-educational service to youth each day. Using 6 h as the length of one school day, the number of minutes of services received by participants was estimated by multiplying the number of days present at school by 360 min (or 6 h).

For Pull-Out Program 1 and Pull-Out Program 2, records of mental health service contact dates and duration were maintained by counselors throughout the school year. These service contact logs were maintained by counselors who delivered mental health services to the students in each program. Each time a counselor provided services to a student, he/she indicated on the log the type of service delivered (i.e., individual therapy, small group therapy, classroom psycho-education, individual therapy with parent, family therapy, and case management or medication management) and the number of minutes the service was delivered.

Analysis Procedures

Descriptive statistics included frequencies and percentages, means and standard deviations. The statistical significance of differences between percentages was evaluated with chi-squared tests. For significance tests, unless adjustments for multiple comparisons or violation of test assumptions were made, data were evaluated using an alpha level of 0.05. Based on the recommendations by the APA (Wilkinson & Task Force on Statistical Inference, 1999), as well as numerous researchers (e.g.,

Cumming & Fidler, 2009; Thompson, 2007; Capraro, 2005; Cumming & Finch, 2001), the difference between means at baseline and follow-up were analyzed by calculating standardized effect size estimates and corresponding confidence intervals.

Analyses were conducted using the Exploratory Software for Confidence Intervals program (ESCI; Cumming, 2001). ESCI is a set of interactive simulations that runs under Microsoft Excel. Case 2 (for two independent groups) from the CIDelta simulation was used for the current study. Using this simulation, Hedge's *g* and corresponding confidence intervals were calculated using noncentral *t* distributions and an iterative algorithm (Cumming & Fidler, 2009). Cohen (1988) suggested guidelines for interpretation of Hedge's *g* such that 0.20 reflects a small effect size; 0.50 reflects a medium effect size; and 0.80 reflects a large effect size. Confidence intervals around effect sizes are included to allow for comparisons across studies and to provide an estimate of how big or small the effect might be in the population (Aberson, 2002).

Results

Socio-Demographic Characteristics

Overall, participants from the four programs were similar in terms of gender; most of the participants were male (between 70 and 80%, depending on the program). Youth enrolled in the Integrated and Milieu Programs were generally around 10–11 years old, and youth in Pull-Out Programs 1 and 2 were generally around 14–15 years old. The majority of participants from the Integrated Program were white (90.0%), as were the majority of participants from the Milieu Program (74.5%). In contrast, the majority of the participants from Pull-Out Program 1 were either black (47.8%) or white (30.4%); and the majority of participants from Pull-Out Program 2 were black (91.7%). The Integrated Program had the highest percentage of participants at or below the poverty level (60.0%), and the Milieu Program had the lowest percentage of participants at or below the poverty level (23.0%); see Table 1. It is important to note that the youth served in these programs were similar in their emotional functioning at admission or when data was initially collected. As will be seen below, the majority of youth served in each of the four programs were in the "clinical range" on standardized measures of emotional functioning (96, 71, 65, and 88%, respectively) at the beginning of the study. Clearly, all four programs were serving students with high-levels of mental health need.

Emotional Functioning

Integrated Program

The mean Total Problems CBCL t-score for participants from the Integrated Program was 76.32 (SD = 7.54) at baseline and 70.42 (SD = 11.54) at follow-up. The resulting effect size for change from baseline to follow-up was in the medium to large range ($g = 0.61$, 95% CI = 0.20, 1.00). The percentage of youth in the clinical range on the CBCL Total Problems scale decreased from 96% at baseline to 70% at follow-up.

Milieu Program

The mean Total Problems CBCL t-score for participants from the Milieu Program was 67.41 (SD = 9.09) at baseline and 64.39 (SD = 10.19) at follow-up. The resulting effect size for change from baseline to follow-up was small ($g = 0.31$, 95% CI = -0.08 , 0.70). The percentage of youth in the clinical range on the CBCL Total Problems scale decreased from 70.6% at baseline to 58.8% at follow-up.

Pull-Out Program 1

The mean SDQ total difficulties score for participants from Pull-Out Program 1 was 20.91 (SD = 8.44) at baseline and 18.17 (SD = 8.02) at follow-up. The resulting effect size for change from baseline to follow-up was in the small to medium range ($g = 0.33$, 95% CI = -0.25 , 0.91). The percentage of participants in the clinical range decreased from 65.2% at baseline to 56.5% at follow-up.

Pull-Out Program 2

For Pull-Out Program 2 participants, the mean SDQ total difficulties score was 21.62 (SD = 5.25) at baseline and 17.83 (SD = 6.59) at follow-up. The resulting effect size for change from baseline to follow-up was in the medium to large range ($g = 0.64$, 95% CI = 0.05, 1.21). The percentage of participants in the clinical range decreased from 87.5% at baseline to 50% at follow-up.

Overall, for all four programs, results indicated an improvement in emotional functioning for participants from baseline to follow-up, with effect sizes for this change ranging from small to large, depending on the program. This improvement in emotional functioning was further evidenced by a decrease in the percentage of participants in the clinical range (on either the CBCL or SDQ) from baseline to follow-up; see Table 2.

Functional Impairment

Integrated Program

The average total CAFAS score for participants in the Integrated Program was 108.88 (SD = 40.69) at baseline and 86.00 (SD = 43.40) at follow-up. The resulting effect size for this decrease in mean total score on the CAFAS was medium ($g = 0.54$, 95% CI = 0.14, 0.94). The percentage of participants in the functional limitations range decreased slightly, from 26.7% at baseline to 24.0% at follow-up.

Milieu Program

Youth in the Milieu Program had a mean total CAFAS score of 87.84 (SD = 38.12) at baseline and 80.20 (SD = 34.26) at follow-up. The effect size for this change was small ($g = 0.21$, 95% CI = -0.18 , 0.60). There was a slight decrease in the percentage of participants with total CAFAS scores in the functional limitations range from 9.8% at baseline to 5.9% at follow-up.

Pull-Out Program 1

Total scores on the BIS for participants in Pull-Out Program 1 decreased from 30.72 (SD = 10.29) at baseline to 22.69 (SD = 9.96) at follow-up. The resulting effect size for this decrease was large ($g = 0.79$, 95% CI = 0.19, 1.39). All the participants in this program (100%) scored in the clinical range for the BIS at baseline. This percentage decreased to 91.3% at follow-up.

Pull-Out Program 2

Total scores on the BIS for participants in Pull-Out Program 2 increased slightly from 24.4 (SD = 9.72) at baseline to 25.64 (SD = 12.82) at follow-up. The effect size for this change was less than small ($g = -0.11$, 95% CI = -0.67 , 0.46). At baseline, 100% of youth scored in the clinical range on the BIS. At follow-up, the percentage of youth in the clinical range decreased to 78.3%. This suggests that although the total percentage of students in the clinical range decreased, the sample may have had some outliers (i.e., participants with very high scores, meaning very low functioning) that contributed to the mean increase in scores on the BIS.

Overall, the general trend observed for change in functional impairment from baseline to follow-up was similar to that observed for emotional functioning, such that participants tended to improve with regard to their functional impairment. Participants from three of the four programs (Integrated, Milieu, and Pull-Out 1) showed decreases in

Table 2 Change in emotional functioning and functional impairment from baseline to follow-up

Program	Baseline <i>M</i> (SD)	Follow-up <i>M</i> (SD)	Hedge's <i>g</i> (95% CI) ^a
<i>Emotional functioning</i>			
Integrated	76.32 (7.54)	70.42 (11.54)	0.61 (0.20, 1.00)
Milieu	67.41 (9.10)	64.39 (10.19)	0.31 (−0.08, 0.70)
Pull-Out 1	20.91 (8.44)	18.17 (8.02)	0.33 (−0.25, 0.91)
Pull-Out 2	21.62 (5.25)	17.83 (6.59)	0.64 (0.05, 1.21)
<i>Functional impairment</i>			
Integrated	108.88 (40.69)	86.00 (43.40)	0.54 (0.14, 0.94)
Milieu	87.84 (38.12)	80.20 (34.26)	0.21 (−0.18, 0.60)
Pull-Out 1	30.72 (10.29)	22.69 (9.96)	0.79 (0.19, 1.39)
Pull-Out 2	24.40 (9.72)	25.64 (12.83)	−0.11 (−0.68, 0.46)

For the Integrated ($n = 50$) and Milieu ($n = 51$) Programs, emotional functioning was measured using the CBCL and functional impairment was measured using the CAFAS. For Pull-Out 1 ($n = 23$) and Pull-Out 2 ($n = 24$), emotional functioning was measured using the SDQ and functional impairment was measured using the BIS

^a Interpretation of hedge's g as suggested by Cohen (1988): 0.20 = small; 0.50 = medium; 0.80 = large. Confidence intervals provide an estimate of how big or small the effect might be in the population. For example, for $g = 0.30$ (CI = −0.10 to 0.70), the population effect size is unlikely to be larger than 0.10 in favor of worsened functioning over time or larger than 0.70 in favor of improved functioning over time

standardized scores of functional impairment, indicating a decrease in impairment. It was only for participants in Pull-Out Program 2 that this trend did not hold; scores on a standardized measure of impairment actually increased slightly from baseline to follow-up for this group of participants, though the percentage of participants in the clinical range on this measure decreased over time; see Table 2.

Academic Achievement

Integrated Program

At baseline, grades reported by caregivers for participants in the Integrated Program were around 46.06 (SD = 25.24), reflecting grades just above the D level. At follow-up, the average grades reported by caregivers increased to 53.33 (SD = 22.17), reflecting grades near the C level. This change resulted in a small to medium effect size ($g = 0.31$, 95% CI −0.18, 0.79).

Milieu Program

The mean standard score on the WIAT reading subtest for participants in the Milieu Program was 85.15 (SD = 18.63) at baseline. Because WIAT scores are standardized so that an average score is 100, it is clear that these students were performing below average academically. At follow-up, the mean reading score increased slightly to 87.87 (SD = 20.30), resulting in a small effect size ($g = 0.14$, 95% CI = −0.31, 0.58). On the WIAT math subtest, the mean standard score at baseline for participants in this program was 78.00 (SD = 12.31). The mean math subtest

score increased to 83.72 (SD = 14.8) at follow-up. The net effect size for this change was in the small to medium range ($g = 0.42$, 95% CI = −0.03, 0.87).

Pull-Out Program 1

The mean standard score on the WRAT3 reading subtest for participants in Pull-Out Program 1 was 87.25 (SD = 17.94) at baseline. These students were below average in academic functioning at the beginning of the study as WRAT3 scores are standardized and scores between 90 and 109 are average functioning. At follow-up, the mean reading score decreased to 86.55 (SD = 17.6). This decrease resulted in a less than small effect size ($g = -0.04$, 95% CI = −0.66, 0.58). A similar trend was seen for the WRAT3 math subtest for these participants. The mean standard score at baseline on this subtest was 77.95 (SD = 10.78), and the mean score at follow-up was 76.7 (SD = 10.13). This change also resulted in a less than small effect size ($g = -0.12$, 95% CI = −0.74, 0.50). These standardized scores reveal these youth were unable to keep up with their non-handicapped peers in reading and math, and as they age they may continue to fall further behind.

Pull-Out Program 2

For participants in Pull-Out Program 2, the mean standard score on the WRAT3 reading subtest at baseline was 87.18 (SD = 22.57) and the mean score at follow-up was 80.75 (SD = 17.21). The effect size for this change was in the small to medium range ($g = -0.32$, 95% CI = −0.94, 0.31). Similarly, the mean standard score on the WRAT3

math subtest decreased from baseline ($M = 78.20, SD = 10.39$) to follow-up ($M = 76.30, SD = 8.24$), resulting in a small effect size ($g = -0.20, 95\% CI = -0.82, 0.42$); see Table 3. Again, these youth were unable to keep up with their non-handicapped peers in reading and math achievement.

Overall, results for academic achievement were somewhat inconsistent across programs. For the Integrated and Milieu Programs, results suggest some improvement (small to medium effect sizes) in academic achievement from baseline to follow-up. Note that the measure for academic achievement used for the Integrated Program was based on parent report of their child’s school performance, which may not be as reliable as either standardized academic measures or grades obtained from a child’s academic school record. Results for the Pull-Out Programs were quite different from the other two programs, with participants from these programs showing a slight decrease on standardized measures of academic achievement over time. This decrease in scores on both reading and math subtests was greater for participants in Pull-Out Program 2 than for those in Pull-Out Program 1; see Fig. 1.

School Participation

The percentage of days present in school was highest for participants from the Integrated Program (91.07%, $M = 163.94$ days), followed by participants from the Milieu Program (80.41%, $M = 144.75$), participants from Pull-Out 1 (62.44%, $M = 112.40$), and participants from Pull-Out 2 (48.43%, $M = 87.17$). It is important to note

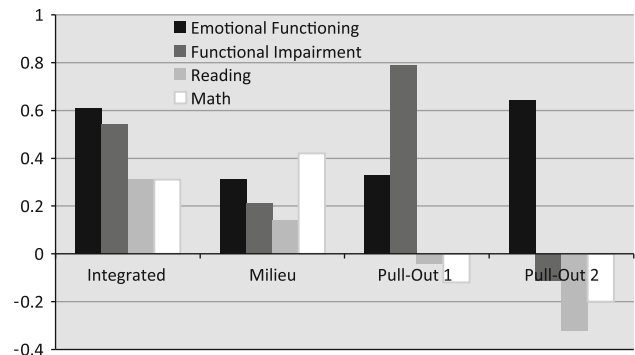


Fig. 1 Effect sizes for emotional functioning, functional impairment, and achievement. Emotional functioning was assessed using the CBCL for the Integrated and Milieu Programs, and the SDQ for the Pull-Out 1 and Pull-Out 2 Programs. Functional impairment was assessed with the CAFAS for the Integrated and Milieu Programs, and the BIS for the Pull-Out 1 and Pull-Out 2 Programs. Achievement was measured by average class grades according to parent report for the Integrated Program; the Milieu Program used the WIAT; and the Pull-Out Programs (1 and 2) used the WRAT3

that attendance data for participants from the Integrated and Milieu Programs were obtained from parent report, whereas attendance data for participants from the Pull-Out Programs were obtained from records provided by the schools.

Mental Health Services Received

Integrated Program

Results for the Integrated Program indicated that, according to caregiver report, 100% of youth who

Table 3 Change in academic achievement from baseline to follow-up

Program	Baseline M (SD)	Follow-up M (SD)	Hedge’s g (95% CI) ^a
<i>Reading achievement</i>			
Milieu	85.15 (18.63)	87.87 (20.30)	0.14 (−0.31, 0.58)
Pull-Out 1	87.25 (17.94)	86.55 (17.60)	−0.04 (−0.66, 0.58)
Pull-Out 2	87.18 (22.57)	80.75 (17.21)	−0.32 (−0.94, 0.31)
<i>Math achievement</i>			
Milieu	78.00 (12.31)	83.72 (14.80)	0.42 (−0.03, 0.87)
Pull-Out 1	77.95 (10.78)	76.70 (10.13)	−0.12 (−0.74, 0.50)
Pull-Out 2	78.20 (10.39)	76.30 (8.24)	−0.20 (−0.82, 0.42)
<i>Academic grades</i>			
Integrated	46.06 (25.24)	53.33 (22.17)	0.31 (−0.18, 0.79)

For the Milieu program ($n = 39$), reading and math achievement were assessed using the WIAT reading and math subtests. For Pull-Out 1 ($n = 20$) and Pull-Out 2 ($n = 20$), the WRAT3 reading and math subtests were used to assess reading and math achievement. For the Integrated Program ($n = 33$), caregiver report of academic grades were used as a measure of academic achievement

^a Interpretation of hedge’s g as suggested by Cohen (1988): 0.20 = small; 0.50 = medium; 0.80 = large. Confidence intervals provide an estimate of how big or small the effect might be in the population. For example, for $g = 0.30$ (CI = −0.10 to 0.70), the population effect size is unlikely to be larger than 0.10 in favor of lower reading achievement over time or larger than 0.70 in favor of greater reading achievement over time

participated in the program received individual therapy and 95.8% of the youth received case management services. More than half of the participants received medication management (69.8%) and group therapy (54.5%). Around one quarter or less of youth received a behavioral aide (26.2%), crisis management services (20.4%), and family therapy (16.7%). Out of seven possible services, the average number of services received over 12 months was 3.67 (SD = 1.26).

Milieu Program

Because mental health services for the Milieu Program were delivered throughout the entire school day, the number of minutes present at school was used as a proxy variable to indicate number of mental health service minutes received. Results indicated that participants received an average of 52,108 min (SD = 11,998) or 868.47 h (SD = 199.96 h) of mental health services over a 12-month time period.

Pull-Out Program 1

The most frequently received services for participants in Pull-Out Program 1, according to program counselor report, were individual therapy (82.6%) and case management or medication management (73.9%). Approximately half (52.2%) of youth received group therapy, and 39.1% received individual therapy with a parent. Fewer youth received group family therapy (34.8%) and classroom psycho-education (21.7%). The average number of these different services received by youth in this program was 3.04 (SD = 1.72). Youth received an average of 1,357 min (SD = 1,139 min) or 22.6 h (SD = 19.0 h) of mental health services over a 9-month time period.

Pull-Out Program 2

The most frequently received service for participants in Pull-Out Program 2, according to program counselor report, was individual therapy (54.2%). About one-third of the participants of this program received case management or medication management (33.3%). One quarter or less of youth received group therapy (25.0%), family therapy (20.8%), and individual therapy with a parent (16.7%). No participants from this program received classroom psycho-education services (0.0%). The average number of services received by youth in this program was 1.50 (SD = 1.69). The average number of minutes of any service received was 514 (SD = 351 min) or 8.55 h (SD = 5.8 h) of mental health services over a 9-month period of time.

Discussion

In this study, we have examined four different organizational models of SBMH that are representative of the majority of service delivery models found in the nation's schools (Foster et al., 2005; Policy Leadership Cadre for Mental Health in Schools, 2001). The rationale for conducting a descriptive study of SBMH is based on the observations of several researchers in the children's mental health services research field (e.g., Bickman, 2000; Garland et al., 2010; Hoagwood & Kolko, 2009; Weisz, Jensen-Doss & Hawley, 2006) that in order to improve the quality of care we need to know about the nature of the care that is being delivered, particularly in what is described as usual care. In addition, it has been noted that before rigorous evaluations of complex systems such as multi-agency SBMH programs are conducted, rich comparative descriptions are needed to guide the development of conceptual models for investigation (Knapp, 1995).

The four SBMH programs participating in the study differed in how they provided mental health services as follows: the Integrated Program had a comprehensive service delivery system in place that integrated services from the school system and a variety of child-serving community agencies; the Milieu Program used professionals who were school employees; Pull-Out Program 1 used school employed staff and providers contracted from a community agency; and Pull-Out Program 2 contracted for services from a community agency. While there was variability in demographic characteristics across programs, participating students, in general, were predominantly boys, from low income families, and black students were over-represented, primarily in the Pull-Out Programs. These findings are similar to demographic characteristics found in nationally representative samples of students who have ED (Wagner et al., 2005b).

As a group, students in the study exhibited very poor emotional functioning, with 79% scoring in the clinical range at baseline. An encouraging finding was that emotional functioning improved over time in all four programs with effect sizes ranging from .31 to .61. Only three of the programs demonstrated improvement in a measure of functional impairment. The range of effect sizes was larger, .21–.79, and one program reported a negative effect size indicating that functioning deteriorated over time. It has been suggested that functional impairment may be a critical indicator of improved functioning in youth who have ED (Greenbaum et al., 1998) and interventions that focus on functional impairment should be a high priority for program administrators.

Results from measures of academic functioning support calls in the literature for mental health interventions to be more focused on learning (Atkins et al., 2010; Cappella

et al., 2008). The National Center for Education Statistics reports an average attendance rate of 92% for all students in the U.S.; however, in the four programs in this study, the average attendance rate was 70%, with a range of 48–91%. Of the seven effect sizes calculated for academic achievement measures, three were positive reflecting small gains and several reveal negative effects. These negative effects reveal that on these standardized achievement measures, these youth are falling behind their non-handicapped peers. Given the high correlation between attendance and achievement (Lamdin, 1996; Roby, 2004), supporting students to attend school is an important goal for SBMH programs along with a need to improve instruction for students who have ED. The overall findings from measures of academic achievement in this study were similar to findings from national studies that indicate deficits in the academic performance of students who have ED (Reid et al., 2004; Wagner et al., 2005b).

Perhaps the most difficult challenge in this study was the attempt to describe the mental health services in each of the programs. Different funding mechanisms, organizational procedures, and program capacity made a comparable description of each program complicated. In the Integrated Program, which had the most developed system of services delivery, there were seven different services available to participating youth and they received an average of 3.67 of these services. There was no mechanism in the Integrated Program data collection procedure to record minutes of mental health service received. Pull-Out Program 1 was a type of hybrid of school employed staff joined by staff from community agencies, and six different types of services were available in this program. Participants received an average of 3.04 services for an average of 2.5 h/month.

The Milieu Program had the most extensive capacity for service delivery, having Medicaid reimbursement available for mental health services and special education funds supporting the educational component of the program. However, this program designed an integrated psycho-educational model delivered in a self-contained classroom that enabled a blending of funding sources to produce an intensive intervention that encompassed the entire school day for students. There were 3.5 h of mental health services provided each day for an approximate total of 72 h/month. Specific mental health services included individual and group therapy and medication management.

In Pull-Out Program 2, mental health counselors came into the school to provide individual and group counseling to students. A contract was issued by the school district for these services. A psychiatric consultant was also available through the contract for medication management issues. There were five different services in this program, and on average, a student received 1.5 services per month for approximately 1 h per month. The contracted mental health

counselors did not take part in any psycho-educational activities in the classrooms.

These findings indicate an extensive range and diversity in how mental health services were provided in these programs. When school staff and community staff collaborate in the service delivery model (Integrated Program and Pull-Out 1), there appeared to be more types of services available and students received more of these different services on average. In Pull-Out Program 2, the community agency staff worked alone and appeared to focus more on providing individual therapy. Students served in this model received the fewest number of minutes of service/month. The Milieu Program capitalized on the state Medicaid regulations, became an eligible provider for direct Medicaid reimbursement, and developed a very intensive psychosocial program that is influenced by the medical model underlying the funding mechanism.

While each program had some area of relative success, no model proved to be highly effective across both mental health and school/academic outcomes. As Hoagwood and her colleagues noted in their review of empirically based interventions (Hoagwood et al., 2007), the few interventions dually effective in both the mental health and academic domains were intensive and complex. They involved the home and the school, requiring teachers and parents to be actively involved in the interventions. In addition, they included school-wide reform activities such as reorganization of discipline policies and procedures, changes in classroom management techniques, and individual level methods that involved behavior management. Such interventions are not characteristic of programs that are considered to be care as usual at this time.

Limitations

The current study reports a descriptive account of components in four SBMH programs from a convenience sample. While there are longitudinal data in the study obtained over a school year, there are no controls (i.e., control groups) enabling any causal interpretations, and the measures used by each program differed. No satisfactory method of describing the mental health services in each program that would adequately compare each was developed. Likewise, we could not compare the instructional programs in detail. This made it impossible to determine whether the effects observed over time were due to the educational component or the mental health component of each program. Many of the differences across programs can be attributed to unique community needs, state and local funding mechanisms, and the advocacy level of local program administrators for this population of students who have disabilities. Additionally, costs associated with serving these students were not

obtained. However, the school-based mental health literature reviewed indicated that the four programs that participated in the present study are representative of the models used by the majority of school districts nationally. A rich description of usual care has been provided and may inform the developing research agenda that targets a needy group of children and youth who have ED.

Conclusions

The research agenda for SBMH needs to be more focused on developing interventions that are dually effective in the emotional and academic domains, especially for youths who have ED and are educated in special education programs. This will involve the development of better measures of what constitutes mental health services provided in schools. While there are criticisms of standardized achievement tests, these measures are clearly viewed as the gold standard and SBMH program research will have to be in concordance with the field.

There are several conceptual challenges facing research in this field. By its nature, SBMH is a multi-disciplinary entity and faces barriers posed by different language, theoretical foundations, and emphasis in training of different professions (Kutash et al., 2006). Consequently, the research agenda continues to need conceptual refinement to bring about consensus on some very basic issues such as identifying the independent variable in empirical studies of SBMH and clearly describing usual care for comparison purposes (Garland et al., 2010).

Over 20% of the nation's youth meet criteria for a diagnosable emotional disorder, and in 2006, \$8.6 billion was spent to treat them, the highest cost for treating any childhood illness (Blau et al., 2010). Schools have been identified as the ideal location to reach these children. The development of effective, intensive, collaborative SBMH programs has the potential to meet the emotional as well as the academic needs of this group of the nation's youth and ameliorate the current poor life outcomes experienced by many of these youth.

Acknowledgments This research was partially supported by Grant H133B90004 from the Center for Mental Health Services, Substance Abuse and Mental Health Administration, and the National Institute for Disability and Rehabilitation Research. The opinions contained in this manuscript are those of the authors and do not necessarily reflect those of either the US Department of Education or the Center for Mental Health Services, SAMSHA.

References

- Achenbach, T. (1991). *Manual for the child behavior Checklist/4–18 and 1991 profile*. Burlington, VT: University of Vermont Department of Psychiatry.
- Adelman, H. S., & Taylor, L. (2006). Mental health in schools and public health. *Public Health Reports*, 121, 294–298.
- Atkins, M. S., Hoagwood, K. E., Kutash, K., & Seidman, E. (2010). Toward the integration of education and mental health in schools. *Administration and Policy in Mental Health and Mental Health Services Research*, 37, 40–47.
- Bickman, L. (2000). The most dangerous and difficult research question in mental health services research. *Mental Health Services Research*, 2, 71–72.
- Bird, H., Canino, G., Davies, M., Ramirez, R., Chavez, L., & Duarte, C. (2005). The brief impairment scale (BIS): A multidimensional scale of functional impairment for children and adolescents. *Journal of American Academy Child and Adolescent Psychiatry*, 44(7), 699–707.
- Blackorby, J., & Wagner, M. (1996). Longitudinal postschool outcomes of youth with disabilities: Findings from the national longitudinal transition study. *Exceptional Children*, 62(5), 399–414.
- Blau, G. M., Huang, L. N., & Mallery, C. J. (2010). Advancing efforts to improve children's mental health in America: A commentary. *Administration and Policy in Mental Health and Mental Health Services Research*, 37, 140–144.
- Brener, N. O., Martindale, J., & Weist, M. D. (2001). Mental health and social services: results from the school health policies and programs study. *Journal of School Health*, 71(7), 305–312.
- Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E. M., et al. (1995). Children's mental health service use across service sectors. *Health Affairs (Project Hope)*, 14(3), 147–159.
- Burns, B. J., & Kutash, K. (2000). Child and adolescent measures of functional status. In *Handbook of psychiatric measures and outcomes* (pp. 357–392). Washington, D.C.: American Psychiatric Association.
- Cappella, E., Frazier, S. L., Atkins, M. S., Schoenwald, S. K., & Glisson, C. (2008). Enhancing schools' capacity to support children in poverty: An ecological model of school-based mental health services. *Administration and Policy in Mental Health*, 35(5), 395–409.
- Capraro, M. M. (2005). An introduction to confidence intervals to intervals for both statistical estimates and effect sizes. *Research in the Schools*, 11(2), 22–33.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.
- Cumming, G. (2001). Exploratory software for confidence intervals (CIdelta) [Computer software]. Bundoora, AU: La Trobe University. Retrieved from <http://www.latrobe.edu.au/psy/esci/>.
- Cumming, G., & Fidler, F. (2009). Confidence intervals: Better answers to better questions. *Journal of Psychology*, 217(1), 15–26.
- Cumming, G., & Finch, S. (2001). A primer on the understanding, use, and calculation of confidence intervals that are based on central and noncentral distributions. *Educational and Psychological Measurement*, 61(4), 532–574.
- Dedrick, R. F., Greenbaum, P. E., Friedman, R. M., Wetherington, C. M., & Knoff, H. M. (1997). Testing the structure of the child behavior checklist/4–18 using confirmatory factor analysis. *Educational and Psychological Measurement*, 57(2), 306–313.
- Department of Health and Human Services. (1999). The surgeon general's report—Chapter 3: Children and mental health. *Mental health: A report of the Surgeon General* (pp. 124–220). Atlanta, GA: Department of Health and Human Services.
- Eber, L., Osuch, T., & Redditt, C. A. (1996). School-based applications of the wraparound process: Early results on service

- provision and student outcomes. *Journal of Child and Family Studies*, 5(1), 83–99.
- Evans, S. W., & Weist, M. D. (2004). Implementing empirically supported treatments in the schools: What are we asking? *Clinical Child and Family Psychology Review*, 7(4), 263–267.
- Foster, S., Rollefson, M., Doksum, T., Noonan, D., Robinson, G., & Teich, J. (2005). *School mental health services in the United States, 2002–2003*. DHHS Pub. No. (SMA) 05–4068. Rockville, MD: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration.
- Garland, A. F., Brookman-Frazee, L., Hurlburt, M. S., Accurso, E. C., Zoffness, R. J., Haine-Schlagel, R., et al. (2010). Mental health care for children with disruptive behavior problems: A view inside therapists' offices. *Psychiatric Services*, 61(8), 788–795.
- Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(11), 1337–1345.
- Greenbaum, P. E., Dedrick, R. F., Friedman, R. M., Kutash, K., Brown, E. C., Lardieri, S. P., et al. (1998). National adolescent and child treatment study (NACTS): Outcomes for children with serious emotional disturbances. In M. H. Epstein, K. Kutash, & A. J. Duchnowski (Eds.), *Outcomes for children and youth with behavioral and emotional disorders: Programs and evaluation best practices* (pp. 21–54). Austin, TX: Pro-Ed.
- Gyamfi, P., Walrath, C., Burns, B. J., Stephens, R. L., Geng, Y., & Stambaugh, L. (2010). Family education and support services in systems of care. *Journal of Emotional and Behavioral Disorders*, 18(1), 14–26.
- Hoagwood, K., & Kolko, D. J. (2009). Introduction to the special section on practice contexts: A glimpse into the nether world of public mental health services for children and families. *Administration and Policy in Mental Health and Mental Health Services Research*, 36, 35–36.
- Hoagwood, K., Olin, S. S., Kerker, B. D., Kratochwill, T. R., Crowe, M., & Saka, N. (2007). Empirically based school interventions targeted at academic and mental health functioning. *Journal of Emotional and Behavioral Disorders*, 15(2), 66–92.
- Hodges, K. (1990). *Child and adolescent functional assessment scale (1994th ed.)*. Ypsilanti, MI: Eastern Michigan University, Department of Psychology.
- Hodges, K., Doucette-Gates, A., & Kim, C. S. (2000). Predicting service utilization with the child and adolescent functional assessment scale in a sample of youths with serious emotional disturbance served by center for mental health services-funded demonstrations. *The Journal of Behavioral Health Services and Research*, 27, 47–59.
- Hodges, K., Doucette-Gates, A., & Liao, Q. (1999). The relationship between the child and adolescent functional assessment scale (CAFAS) and indicators of functioning. *Journal of Child and Family Studies*, 8, 109–122.
- Hodges, K., & Kim, C. S. (2000). Psychometric study of the child and adolescent functional assessment scale: Prediction of contact with the law and poor school attendance. *Journal of Abnormal Child Psychology*, 28, 287–297.
- 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. *The Journal of Mental Health Administration*, 24, 278–290.
- Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, 62(6), 617–627.
- Knapp, M. S. (1995). How shall we study comprehensive, collaborative services for children and families? *Educational Researcher*, 24(4), 5–16.
- Kratochwill, T. R., Clements, M. A., & Kalymon, K. M. (2007). Response to intervention: Conceptual and methodological issues in implementation. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention* (pp. 25–52). New York: Springer.
- Kutash, K., Duchnowski, A. J., & Lynn, N. (2006). *School-based mental health: An empirical guide for decision-makers*. Tampa: The Research and Training Center for Children's Mental Health; Louis de la Parte Florida Mental Health Institute. University of South Florida.
- Lamdin, D. J. (1996). Evidence of student attendance as an independent variable in education production functions. *The Journal of Educational Research*, 89, 155–162.
- Leaf, P. J., Alegria, M., Cohen, P., Goodman, S. H., Horowitz, 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 and Adolescent Psychiatry*, 35(7), 889–897.
- Manteuffel, B., Stephens, R., & Santiago, R. (2002). Overview of the national evaluation of the comprehensive community mental health services for children and their families program and summary of current findings. *Children's Services: Social Policy, Research, and Practice*, 5, 3–20.
- Mellor, D. (2004). Furthering the use of the strengths and difficulties questionnaire: Reliability with younger child respondents. *Psychological Assessment*, 16(4), 396–401.
- New Freedom Commission on Mental Health. (2003). *Achieving the promise: Transforming mental health care in America* (Final Report ed.). Rockville, MD: DHHS Pub. No SMA-03-3832.
- Newman, L., Wagner, M., Cameto, R., Knokey, A. M., & Shaver, D. (2010). *Comparisons across time of the outcomes of youth with disabilities up to four years after high school: A report of findings from the national longitudinal transition study (NLTS) and the national longitudinal transition study-2 (NLTS2)* (NCSE 2010–3008). Washington, DC: Institute of Education Sciences. Retrieved from http://policyweb.sri.com/cehs/publications/nlts2_report_2010_09.pdf.
- No Child Left Behind (NCLB). (2002). Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425.
- Paternite, C. E. (2005). School-based mental health programs and services: Overview and introduction to the special issue. *Journal of Abnormal Child Psychology*, 33(6), 657–663.
- Policy Leadership Cadre for Mental Health in Schools. (2001). *Mental health in schools: Guidelines, models, resources, and policy considerations*. Los Angeles, CA: School Mental Health Project, Department of Psychology, UCLA.
- Quist, R., & Matshazi, D. (2000). The child and adolescent functional assessment scale (CAFAS): A dynamic predictor of juvenile recidivism. *Adolescence*, 35, 181–192.
- Reid, R., Gonzalez, J. E., Nordness, P., Trout, A., & Epstein, M. H. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of Special Education*, 38(3), 130–143.
- Robbins, V., & Armstrong, B. J. (2005). The Bridges Project: Description and evaluation of a school-based mental health program in eastern Kentucky. In M. H. Epstein, K. Kutash, & A. J. Duchnowski (Eds.), *Outcomes for children and youth with emotional disorders and their families: Programs and evaluation best practices* (2nd ed.; pp. 355–373). Austin, TX: Pro-Ed.
- Roby, D. E. (2004). *Research on school attendance and student achievement: A study of Ohio schools*. Behavioral Research Press, LLC, Grambling, LA. ERIC EJ714746.

- Rones, M., & Hoagwood, K. E. (2000). School-based mental health services: A research review. *Clinical Child and Family Psychology Review*, 3(4), 223–241.
- Sattler, J. M. (2001). *Assessment of children: Cognitive applications* (4th ed.). La Mesa, CA: Jerome M. Sattler, Publisher, Inc.
- Simpson, R. L. (2004). Inclusion of students with behavior disorders in general education settings: Research and measurements issues. *Behavior Disorders*, 30(1), 19–31.
- Stroul, B. A. & Friedman, R. M. (1994). *A system of care for children and youth with severe emotional disturbances*. Tampa: The Research and Training Center for Children's Mental Health; Louis de la Parte Florida Mental Health Institute. University of South Florida.
- The Psychological Corporation. (1992). *Wechsler individual achievement test*. San Antonio, TX: The Psychological Corporation.
- Thompson, B. (2007). Effect sizes, confidence intervals, and confidence intervals for effect sizes. *Psychology in the Schools*, 44(5), 423–432.
- US Department of Education, Office of Special Education and Rehabilitative Services. (2007). *29th Annual report to congress on the implementation of the individuals with disabilities education act*. Washington, DC.
- US Government. (2004). Code of federal regulation, title 34, Section 300.8.
- US Public Health Service. (2000). *Report of the surgeon general's conference on children's mental health: A national action agenda*. Washington, DC: Department of Health and Human Services.
- Wagner, M., Friend, M., Bursuck, W., Kutash, K., Duchnowski, A. J., Sumi, W. C., et al. (2006). Educating students with emotional disturbances: A national perspective on school programs and services. *Journal of Emotional and Behavioral Disorders*, 14(1), 12–30.
- Wagner, M., Kutash, K., Duchnowski, A. J., & Epstein, M. H. (2005a). The special education elementary longitudinal study and the national longitudinal transition study: Study designs and implications for children and youth with emotional disturbance. *Journal of Emotional and Behavioral Disorders*, 13(1), 25–41.
- Wagner, M., Kutash, K., Duchnowski, A. J., Epstein, M. H., & Sumi, W. C. (2005b). The children and youth we serve: A national picture of the characteristics of students with emotional disturbances receiving special education. *Journal of Emotional and Behavioral Disorders*, 13(2), 79–96.
- Weist, M. (1997). Expanded school mental health services: A national movement in progress. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 19, pp. 319–352). New York: Plenum Press.
- Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual care: A meta-analysis of direct comparisons. *American Psychologist*, 61, 671–689.
- Wilkinson, G. S. (1993). *Wide range achievement test* (3rd ed.). Wilmington, DE: Wide Range.
- Wilkinson, L., & Task Force on Statistical Inference, American Psychological Association, Science Directorate. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54(8), 594–604.