Chapter Ten

Measurements and Instrumentation

A 'Special Feature' for America's Children—Key National Indicators of Well-Being, 2005: Parental Reports of Emotional and Behavioral Difficulties

Gloria Simpson

Introduction

In 1994, the leaders of seven federal agencies or offices[1] met for the purpose of improving the data collection, reporting and dissemination of information about U. S. children and families. This group, the Federal Interagency Forum on Child and Family Statistics, was formalized by Executive Order of the President in 1997. One of the Forum's first initiatives was to publish a volume of key indicators of child well-being. The first issue of *America's Children: Key National Indicators of Well-Being* was released in July 1997 with 25 indicators.

This report, which calls attention to the well-being of children and families, receives widespread media attention. In 2004, it was estimated that media coverage (including print, radio, television, and the Internet) for *America's Children* reached 37 million people in the U. S. International coverage was known to include Canada, South Africa, and Great Britain.

Since the first report, an indicator measuring children's mental health has been among the measures identified in the "Data needs" section of *America's Children*. In 1999, the Forum's Reporting Committee created a work group to explore closing this data gap. This work group includes staff from the National Institute of Mental Health (NIMH), the Substance Abuse and Mental Health Services Administration (SAMHSA), the National Center on Birth Defects and Developmental Disabilities, CDC, and the National Center for Health Statistics (NCHS, CDC).

Criteria for indicators in *America's Children* include that they:

- be easy to understand by the general public;
- be based on substantial research relating to child well-being;
- be representative of large segments of the national population;
- use data from a federal survey; and
- be measured regularly, (if not annually, at least every two or three years).

The following summary presents the process and rationale behind selection of a children' mental health indicator, preliminary findings from its administration, and predictive ability related to service use and diagnoses of mental health issues in children.

Methods

Finding a mental health indicator that would meet these criteria presented a real challenge. Complicated measures and scales may not be understood by the general public and could be misinterpreted by the media. There are few federal surveys that include questions on children's mental health, and most of these were not done with a large enough sample or on a regular basis. The Mental Health Work Group examined data from several different federal surveys and consulted with an outside expert panel in order to find an appropriate and acceptable indicator. Because no one available mental health scale or measure met the report criteria, the Mental Health Work Group and the outside consultant panel agreed on a simple direct indicator derived from a question used in the Strengths and Difficulties Questionnaire (SDQ) developed by Robert Goodman in London, England (Goodman, 1997).

---

[1] The founding agencies/offices of the Forum were the Bureau of the Census, the Bureau of Labor Statistics, the National Center for Education Statistics, the National Center for Health Statistics, the National Institute for Child Health and Human Development, the Office of Management of Budget, and the Office of the Assistant Secretary for Planning and Evaluation, DHHS. The forum has now grown to represent 20 federal agencies. These agencies are listed on the forum's Web site at: www.childstats.gov.

The items from the SDQ were first introduced to the annual National Health Interview Survey (NHIS) in 2001 (NCHS, 2002). The SDQ questions were also included in the NHIS in 2003 and 2004 (NCHS, 2004). A shortened set of six SDQ questions were included in the NHIS in 2002 (NCHS (2003). In the NHIS, the SDQ is asked of parents (or someone in the household who knows the most about the children's health) of a sample child in the household, including approximately 9,000 to 10,000 children ages 4 to 17. The SDQ includes an overall question before the questions on impairment: "Overall, do you think that your child has difficulties in one or more of the following areas: emotions, concentration, behavior or being able to get along with other people?" Response choices are: *No; Yes, minor difficulties; Yes, definite difficulties; or Yes, severe difficulties*. Responses of **definite** or **severe** difficulty are considered an indicator of emotional and behavioral difficulties.

The rationale for this approach was largely based on the fact that responses of a *definite* or *severe* difficulty to this question had been a reasonable predictor of DSM-IV diagnostic status among 10,201 children in a British sample (Goodman, Ford, Richards, Gatward, & Melzer, 2000). In addition, a three-year follow-up of the British sample revealed responses of *definite* or *severe* to this question to be predictive of future mental health service use among children (Goodman, 2004). In the NHIS, this indicator was significantly related to parent reports of children having mental retardation, ADHD, a learning disorder, and autism. Responses of *definite* or *severe* to this question were also significantly related to contact with mental health services in the NHIS. In the 2003 NHIS, over 60% of these children used some type of mental health service (Bourdon, Goodman, Rae, Simpson & Koretz, 2005; Federal Interagency Forum on Child and Family Statistics, 2005).

Results

Data from the 2003 administration of the NHIS indicate that 5% of U. S. children ages 4 to 17 had definite or severe difficulties in emotions, concentration, behavior, or being able to get along with other people. The percentage of children with definite or severe difficulties varied by sex, age, family structure and poverty level. Six percent of boys had definite or severe difficulties compared to 3% of girls. Younger children were less likely to have definite or severe difficulties compared to older children. These percentages ranged from 3.3% for children 4-7 years to 6.1% for children 15-17 years. Children living with two parents were less likely to have definite or severe difficulties (4%) compared to children living with a single mother (7%). Poor children were more likely to have definite or severe difficulties compared to other children (8% versus 5-6%).

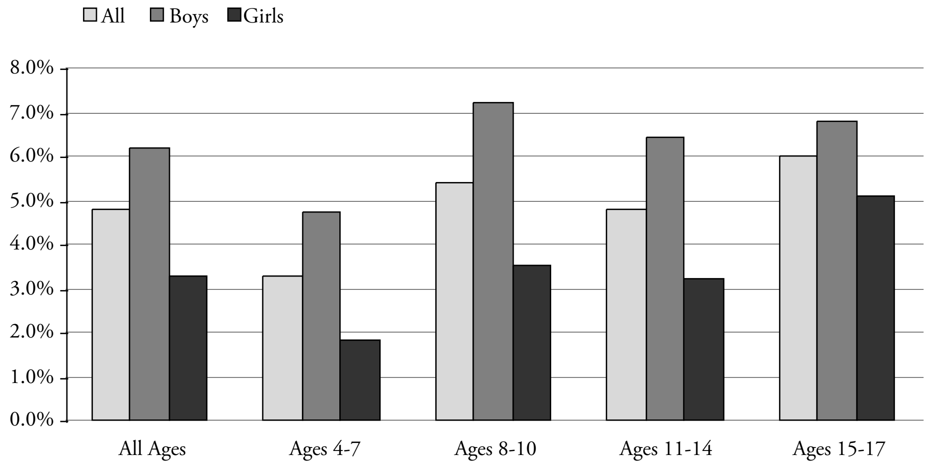
Conclusion

The draft chart proposed to report 2003 data for *America's Children, 2005* is provided below. The final 2005 *America's Children* report has an anticipated release date of July 2005. Validity work on this question and other questions in the SDQ is on-going. Further information on *America's Children* may be found at: www.childstats.gov. Information on the SDQ may be found at: www. sdqinfo.com.

When the *America's Children* report is released each July, it receives a great deal of national media attention. It is hoped that including a mental health measure in *America's Children* will call attention to mental health as a critical aspect of children's overall well-being.

A final note. The NHIS data files which include data on the SDQ, provide a wealth of information that can be used for analysis of mental health and mental health services in relation to other health conditions, insurance, socio-democratic and other variables. Furthermore, data from additional questions on children's use of mental health services will be available sometime in 2006. These data may be accessed on the NHCS website: www.cdc.gov/nchs/ and used by researchers to further our knowledge in this area.

Figure 1
Percentage of Children Ages 4 to 17 Reported by a Parent to have Definite or Severe Emotional or Behavioral Difficulties, by Age and Sex, 2003



Notes. Children with definite or severe emotional or behavioral difficulties are defined as those whose parent responded “yes, definite” or “yes, severe” to the following question on the Strengths and Difficulties Questionnaire (SDQ): “Overall, do you think that (child) has any difficulties in one or more of the following areas: emotions, concentration, behavior, or being able to get along with other people?” Response choices were: (1) No; (2) Yes, minor difficulties; (3) Yes, definite difficulties; (4) Yes, severe difficulties.
Source: Centers for Disease Control and Prevention, National Center for Health Statistics, National Health Interview Survey

References

Bourdon, K. H., Goodman, R., Rae D. S., Simpson, G., & Koretz, D. S. (2005). The Strengths and Difficulties Questionnaire: U.S. normative data and psychometric properties. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*(6), 557-564.

Federal Interagency Forum on Child and Family Statistics. *America's children: Key national indicators of well-being: 2005*. Author: Washington DC, U. S. Gov. Printing Office. Tables Special 3.A, 165-166.

Goodman R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586.

Goodman, R. (2004). Proposal accepted by the Reporting Committee. Special tabulations. Electronic memo.

Goodman, R., Ford, T., Richards, H. Gatward, R. & Melzer, H. (2000). The Development and Well-being Assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal Child Psychol. Psychiatry, 41*, 645-655.

National Center for Health Statistics (2002). *2001 NHIS Survey Description Document*. Retrieved from http://www.cdc.gov/nchs/about/major/nhis/

National Center for Health Statistics (2003). *2002 NHIS Survey Description Document*. Retrieved from http://www.cdc.gov/nchs/about/major/nhis/

National Center for Health Statistics (2004). *2003 NHIS Survey Description Document*. Retrieved from http://www.cdc.gov/nchs/about/major/nhis/

CONTRIBUTING AUTHORS

Gloria Simpson, M.A.
National Center for Health Statistics, Toledo Road, Hyattsville, Maryland 20782,
301-458-4502, fax: 301-458-4035, e-mail: gas4@cdc.gov

Level of Care Determination in Child Welfare: Evidence from the Child Program Outcome Review Team (CPORT) Review

Andres J. Pumariega
Pat Wade
Udema Millsaps
Michele Moser
Terrence Clark

Introduction

Level of care determination is an important process in treatment and service. This process has always been fraught with unreliability across clinicians, given the multiple factors involved in arriving at this determination (Lyons and Abraham, 2001). In previous decades, prevailing orientations led clinicians to treat children and adolescents in facilities away from the child's home, making level of care determination a simpler task. Such practices have changed with the advent of evidence-based psychosocial and pharmacological treatment modalities, the community-based systems of care model (supporting intensive treatment services within the child's home and community), and resource and financial pressures increasingly placed on care delivery systems (Pumariega, et al., 1997; Pumariega & Winters, 2003). The child welfare system faces similar changes and challenges as those faced by the mental health system around service delivery, planning, and funding. Although many children are shared across child serving agencies, the translation of approaches from the mental health sector to the child welfare sector is not certain. There are added challenges for children in state custody whose families are either not available or functional, and overall greater levels of stressors faced by both child and family. Tools and approaches used within child mental health require formal evaluation with the population of children served by child welfare to ensure their applicability and need for adaptation.

As a response to this evolution in practice, and to develop an open and objective level of care decision support tool, the American Academy of Child and Adolescent Psychiatry (AACAP) Work Group on Community-based Systems of Care, with the assistance of the American Association of Community Psychiatrists (AACP) developed the Child and Adolescent Service Intensity Instrument (CASII; American Academy of Child & Adolescent Psychiatry, 2004). The CASII is based on key dimensions in the determination of level of care (risk of harm, level of function, stability of the child's environment, presence of co-morbidity, resiliency and recovery potential, and engagement by the child and family in the care process) that are relevant to services delivered by child welfare. The levels of care defined by CASII were constructed from an inter-agency, community-based systems of care perspective, and outline levels of intensity of care/treatment that can be translated to child welfare levels of care. This summary outlines results from evaluation of the CASII as a level of care decision support tool for children served in child welfare.

Methods

Participants

The annual Tennessee Child Program Outcome Review Team (CPORT) Review provided the opportunity to evaluate the reliability, validity, and applicability of a level of care tool to a population of children served in child welfare. We report on data collected in the 2003 CPORT Review, which include 437 children in custody ages 6 to 19 years of age, from dependent/neglected populations. They constitute a stratified probability sample 95% representative of the over 11,000 children in custody in the state. These children were in levels of care ranging from residential treatment, therapeutic and regular foster care, to family placements.

Instruments/Ratings

The 49 raters of the CPORT Review (bachelors and masters level social workers) were trained on the CASII in six hour workshops by trainers trained by one of the original developers of the instrument. They also underwent similar training for the CAFAS and the rest of the CPORT evaluation protocol and were evaluated for inter-rater reliability, reaching the 90% to 95% level of reliability for overall level of care recommendations in training vignettes with trained and experienced reviewers/supervisors. The children and youth were rated using the CASII, the Child and Adolescent Functional Assessment Scale (CAFAS, Hodges & Wong, 1996), the Child Behavior Checklist/Youth Self Report/Teacher Report Form (CBCL/ YSR/ TRF; Achenbach, 1991), and child and family and systems indicators defined by the CPORT review.

We examined the inter-rater reliability of these raters in the overall CASII total score and level of care as well as its dimensions. We analyzed CASII total scores and level of care recommendations for these children across age, gender, racial (primarily Black-White), and mean CASII total scores diagnostic status (whether or not the child has a psychiatric diagnosis). Additionally, we compared the CASII dimensional ratings and total score to the Child and Adolescent Functional Assessment Scale (CAFAS; Hodges & Wong, 1996) composite scores, and the CASII total scores and levels of care to CBCL, YSR, and TRF sub-scales, internalizing and total T scores. We also analyzed the correlations between CASII total scores and level of care scores to both actual level of care placement as well as to child welfare outcomes (child and family outcomes and system outcomes) defined by the CPORT Review.

Results

The overall inter-rater reliability (ICC 2,1) of raters were CASII total scores (0.916) and level of care recommendations (0.918), with dimension scores ranging from 0.885 to 0.619, except for the Environmental Stress dimension (0.034). There were no significant differences across level of education (bachelors versus masters) or level of experience (less than the 16 years' median versus 16 or more years). CASII levels of care were not significantly correlated to gender, race, or diagnosis, but were correlated to age $\chi^2(df = 12, N = 437) = 58.5, p < 001$, suggestive of higher levels of care in older children.

The Pearson correlation coefficients between CAFAS composite scores and CASII dimensions ranged from 0.708 to 0.381, with a correlation to CASII total score of 0.710 (all $p < 0.001$). The Pearson correlations to CBCL sub-scales ranged from 0.456 to 0.225, with those to CBCL Total T (0.454), Internalizing T (0.385) and externalizing T (0.445) being similarly significant (all $p < 0.001$). The correlations to the YSR and TRF sub-scales, total T scores, and Internalizing and Externalizing T scores were similarly significant, all being significant and ranging from 0.432 to 0.180, except for the YSR Somatic and TRF Somatic and Withdrawn sub-scales (which were non-significant). CASII total scores and recommended levels of care are correlated to actual level of care placement for the children surveyed ($p = 0.011$). They are also highly correlated to CPORT child and family outcome indicators (most at the $p < .001$ level and no less than 0.018) and to many of the system outcome indicators, though strongly correlated to the summary indicator (Overall Adequacy of Services, $p < .001$).

Conclusions

The CASII has demonstrated its capability of functioning within an inter-agency context with equal reliability and validity as in child mental health contexts. It also promises to be a measure that promotes better child welfare outcomes in this era of resource constraints. CASII has potential utility in placement decisions on child welfare as well as in utilization review and as systems of care planning for children in child welfare.

References

Achenbach, T.M. (1991), *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, Vermont: University of Vermont Department of Psychiatry.

American Academy of Child and Adolescent Psychiatry Work Group on Systems of Care, American Association of Community Psychiatrists (2004), *Child and Adolescent Service Intensity Instrument User's Manual* Washington, DC: Author.

Hodges, K. & Wong, 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.

Lyons, J. S., Abraham, M. E: *Designing level of care criteria in the integrated behavioral health continuum*. Kiser, L. J., Lefkowitz, P., Kennedy, L. (2001). Washington, DC, APA Press.

Pumariega, A. J., Nace, D., England, M. J., Diamond, J., Mattson, A., Fallon, T., Hanson, G., Lourie, I., Marx, L., Thurber, D., Winters, N., Graham, M., & Wiegand, D. (1997) Community-based systems of care approach to children's managed mental health services. *Journal of Child and Family Studies 6*(2): 149-164

Pumariega, A. J. & Winters, N. C. (2003). *Handbook of Community systems of care; The new child & adolescent community psychiatry.* San Francisco: Jossey Bass.

CONTRIBUTING AUTHORS

Andres J. Pumariega M.D.
Professor and Director, Child & Adolescent Psychiatry, Department of Psychiatry and Center of Excellence for Children in Custody, East Tennessee State University, 212 McWherter Hall, Box 70567, Johnson City, TN, 37614, pumarieg@mail.etsu.edu and pumariega@aol.com

Pat Wade, M.A.
Director, CPORT Review, Tennessee Commission on Children and Youth, 710 James Robertson Parkway, Nashville, TN 37243-0800, 615-741-2633, fax: 615-741-5956,
email: pat.wade@state.tn.us

Udema Millsaps, M.Ed.
Research Assistant, Department of Psychiatry and Center of Excellence for Children in Custody, East Tennessee State University, 212 McWherter Hall, Box 70567, Johnson City, TN, 37614,
email: childres@etsu.edu

Michele Moser, Ph.D.
Assistant Professor of Psychiatry, Department of Psychiatry and Center of Excellence for Children in Custody, East Tennessee State University, 212 McWherter Hall, Box 70567, Johnson City, TN, 37614, email: moser@etsu.edu

Terrence Clark, M.D.
Assistant Professor of Psychiatry, Department of Psychiatry and Center of Excellence for Children in Custody, East Tennessee State University, 212 McWherter Hall, Box 70567, Johnson City, TN, 37614, email: clarkt@etsu.edu

The Use of Qualitative Methods in Systems of Care Research

Sharon Hodges
Allison Pinto
Mario Hernandez
Caitlin Uzzell

Introduction

Qualitative research methods are valued in the development of knowledge though the exploration of experience and context, in understanding multiple perspectives on an issue or topic, and in understanding the complexity in which phenomena exist. In children's mental health, a field that is largely informed by the results of quantitative research, there has been no analysis of the contribution of qualitative research has made to the knowledge base or the range and depth of qualitative research in this field. This paper presents a review of the proceedings of the annual research conference, *A System of Care for Children's Mental Health: Expanding the Research Base* (RTC Conference), for the purpose of summarizing the qualitative content presented at this conference since its inception in 1988. This review was undertaken as an initial and rudimentary effort to understand the impact of qualitative methods on the field of children's mental health. The goal of the review was to assess both the focus and frequency of qualitative research presented at the RTC conferences from 1988 through 2003. It is hoped that this work will set the stage for developing a more in-depth understanding of how qualitative methods have contributed to this research base and provide a platform for developing recommendations regarding future qualitative research to advance our knowledge of systems of care.

Method

The proceedings of the annual research conference are presented in an edited volume published annually by the Research and Training Center for Children's Mental Health at the University of South Florida. Each year this national conference is host to researchers, policy makers, service professionals, educators, and family members who gather to share their research findings, insights, and experiences in an effort to improve mental health services and outcomes for children and their families. Immediately following the conference, submissions to the proceedings are solicited from all presenters, with the goal of providing brief snapshots of the discussions, papers, and posters presented at this conference. The proceedings are designed to identify those contributing to current work in the field and to promote future research on aspects of design, implementation, and evaluation of systems of care for children and families.

The data presented in this review resulted from an analysis of the presentation summaries included in the 16 volumes of Proceedings from conferences held from 1988 – 2003[1]. It should be noted that RTC conference presenters are not required to prepare and submit summaries to the published proceedings, and therefore the contents do not capture all of the research presented at the conference. Consequently, results of this analysis are based only on presentation summaries formally submitted for publication to the Proceeding editors.

The first step in conducting this review was to establish a shared definition for qualitative research among members of the research team so that initial criteria identification of the published findings as qualitative research could be established. These criteria were applied to three years of proceedings and then refined as summaries were reviewed and the research team discussed studies on a case-by-case basis. The revised criteria that were applied across all 16 years of the published proceedings were as follows:

- Analysis of narrative data that generated themes and patterns was considered qualitative.
- Studies identified as qualitative by the authors were included, unless reported results were based solely on quantitative analysis.

---

[1] *Electronic versions of the Proceedings from 1996 to the present are available at http://rtckids.fmhi.usf.edu/conference_proceedings.htm Information regarding previous volumes is available from the editors, Research and Training Center for Children's Mental Health, liberton@fmhi.usf.edu, 813-974-4661.*

- Interviews were generally considered qualitative, unless results were reported according to a rating system or checklists and there was no evidence of narrative analysis.
- Focus groups of any kind were considered qualitative.
- Documents were considered qualitative if they were treated as narrative data sources and analyzed thematically.
- Literature review was considered qualitative if information pulled was highly thematic.
- Concept mapping was considered qualitative when data analysis focused on the themes identified in the generated statements.
- Backward mapping approaches to policy analysis were considered qualitative if they considered patterns and themes in the data.
- Summaries reporting evaluation results that met the criteria described above were considered qualitative.
- If research was described by authors as using mixed method approaches, this analysis focused on the qualitative aspects, but noted the mixed method approach.
- Coded case record review was considered qualitative when there was thematic emphasis in the review.

Based on the review, a data matrix was developed to record information about presentations that were determined to fit the criteria of qualitative research described above. This matrix included information in the following domains: presentation title, topic, researchers, funders, research design, sampling, data collection, data analysis and results. Information was entered into the data matrix in the form of direct quotes from the RTC conference proceedings. The data listed in each domain were reviewed across studies and across years. Graphs and tables were created to represent and communicate the patterns and themes that emerged in the data, in order to draw conclusions and clarify recommendations.

Results

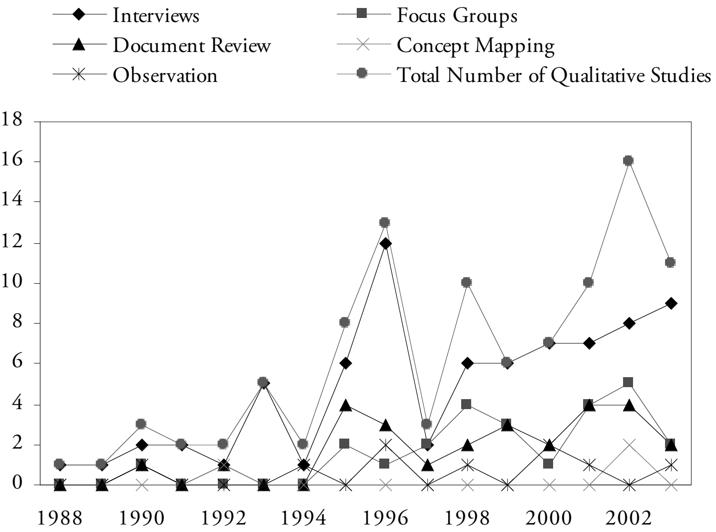
A total of 100 studies were identified as qualitative in this review of conference proceedings. The number of qualitative studies presented at the conference and included in the proceedings has generally been on the increase (1988, *N* = 1; 2003, *N* = 11). Federal agencies (*N* = 15) and private foundations (*N* = 8) were the most commonly identified funding sources for qualitative research; however, the majority of qualitative studies (*N* = 72) did not identify a funding source. The most common topics of study were (a) stakeholder perspectives (*N* = 18); (b) program description or evaluation (*N* = 17); (c) system-of-care description or evaluation (*N* = 15); (d) understanding mental health services funding (*N* = 8), and (e) understanding process of collaboration (*N* = 6). Although many studies (*N* = 28) addressed multiple service delivery subsystems, many more studies (*N* = 47) specifically addressed the mental health subsystem. Table 1 provides descriptions of the most common qualitative study topics.

With regard to research design, data collection, and analysis, most studies (*N* = 85) did not specify research design independently of describing methods used. Half of the studies (*N* = 50) gathered and compared data from multiple informants. Interviews were the most commonly identified method of data collection (*N* = 76), and semi-structured interviews were most commonly reported (*N* = 23) in studies that specified the type of interview conducted (*N* = 42). Document reviews (*N* = 12) and record reviews (*N* = 11) were more often reported than literature reviews (*N* = 5). Observation was rarely reported as a method of data collection (*N* = 9). Most studies (*N* = 71) did not specify the method used to analyze data.

Table 1
Author-Provided Descriptions of Most Common Qualitative Study Topics

Topic of Study	*Descriptions*
Understanding Stakeholder Perspectives	"To provide information about the evaluation interests and technical assistance needs of three key stakeholders" "To learn more about what caregivers included in a very broad definition of 'culture'" "[The study] examines the impact of inclusion on special educators working with nine children with Emotional Behavioral Disturbances" "[A] critical issues analysis of the permanency and support needs of children at risk of parental loss due to HIV/AIDS"
Service Delivery Description & Evaluation	"Discovering the basic social processes and values underlying the agency's service philosophy" "An assessment…to address the following goals: (a) identify key elements of the program, (b) describe relationships between program elements, (c) identify interpretable dimensions of the program, (d) determine the perceived importance and effectiveness of services and (e) identify categories of services" "To gain knowledge about the factors related to positive treatment outcome, barriers to accessing services, and methods for improving service delivery for adolescents receiving school-based mental health services" "Evaluation of the On-Campus Intervention Program" "To evaluate a behavior management system created to improve interpersonal functioning and classroom achievement"
System of Care Description & Evaluation	"This detailed case study represents one portion of an evaluation of the Vermont Community Integration Demonstration Project" "The impact of CASSP and changes in California's service delivery system over the last four years" "summarizes preliminary findings and their relationship to evaluative challenges encountered while measuring child and family outcomes when ___ multiple innovative services at the local level, within a dynamic and complex state administered system" "To determine the relationships among policy development, policy implementation, and the development of collaboration in systems of care" "To elucidate the local theory of change behind implementation and dissemination of the High Point Initiative's application of a system-of-care approach to the area of juvenile justice"
Understanding Mental Health Services Funding Processes	"To report…the effectiveness of the pilot program and the challenges encountered during the transition from a traditional program to a managed care mental health treatment program for youth" "To understand the impact of the new regulations on families and children whose serious emotional disability had previously qualified them to receive SSI disability benefits, and for whom the continuation of this benefit is now in jeopardy or has already been lost" "To identify the various funding sources used by programs which serve and support youth and young adults"
Understanding Collaborative Processes	"A qualitative evaluation of a locally-based effort to increase levels of interaction between schools and parents from East Tampa, Florida" "To better understand the factors that affect the provision of effective service coordination" "[To study] (a) the nature and extent of interagency involvement in systems of care; (b) what approaches are used to enlist interagency involvement in system of care governance, program operations, and direct services; (c) what challenges impede interagency involvement; and (d) the relationship between interagency involvement and referral patterns" "To identify the most effective strategies that promote collaborative relationships on evaluation teams"

Figure 1
Method of Data Collection in Qualitative Studies (*N* = 100)
(Method of data collection was explicitly identified in data extracted from Proceedings)



Discussion

A system of care has been defined as a comprehensive spectrum of mental health and other necessary services which are organized into a coordinated network to meet the multiple and changing needs of children and adolescents with severe emotional disturbances (Stroul & Friedman, 1994). This review of RTC conference proceedings suggests a growing recognition of the value of qualitative research in studying systems of care, particularly for the purpose of studying the complex, dynamic and emergent properties of these systems. While the Proceedings do not capture all research presented at the annual conference, the contents do reflect themes in research and evaluation methodology over time. Based on this review of qualitative research presented at RTC Conferences since 1988, the following recommendations for researchers as well as funding and research institutions are offered below.

Many of the qualitative studies included in the RTC conference proceedings did not describe the specifics of research design, methodology, or data analysis. It is not possible, on the basis of this review, to determine whether this absence of specificity regarding the research process is a result of space constraints imposed by the conference proceeding format or if it reflects the need for more attention to methodological issues among researchers conducting qualitative studies. However, it is recommended that researchers make every effort to present qualitative findings in a manner that conveys the conceptual, methodological and analytical rigor with which they have approached their work, in terms of: (1) being explicit in describing the research design and how the design is expected to contribute to expanding the knowledge base regarding systems of care, (2) clearly distinguishing between the research design and the methods of data collection and analysis employed within that design, and (3) specifically describing efforts to address issues of reliability and validity.

Funding agencies and institutions that support research and evaluation activities are encouraged to continue using qualitative approaches in the study of systems of services and supports in children's mental health. Qualitative studies are especially recommended when the goal is to: (a) clarify local need and the perspectives of multiple stakeholders; (b) better understand systems-of-care efforts within context; (c) gain insight into complex systems-level phenomena (such as collaboration, cultural competence, access, financing and governance); or (d) track the impact of policy changes and system reform.

Because the findings of qualitative studies can be used to inform future qualitative research efforts and might also generate hypotheses that could be tested in future quantitative research, institutions that provide funding or conduct research and evaluation are encouraged to review existing qualitative studies for results and conclusions that elucidate real world phenomena linked to systems-of-care theory and practice. Furthermore, these institutions are encouraged to provide opportunities for training and consultation in order to build understanding of the potential contributions of qualitative research, to build understanding of the nature of qualitative evidence, and to build qualitative research skills. Finally, funding agencies and institutions that support research and evaluation activities are encouraged to provide formatting guidelines and structure for grant applications that support the proposal and conduct of qualitative research in ways that take full advantage of the investigative opportunities offered by these methods and fully support their rigorous design and implementation.

Institutions and organizations that support peer reviewed publications and conference submissions are encouraged to provide opportunities for the dissemination of qualitative research findings in ways that maximize the contributions of these methodological approaches to the knowledge base.

Conclusion

There is a gap between what we know and what we need to know about systems of services and supports in child mental health. In order to bring about positive change in the structure and function of systems that serve children with serious emotional disturbance and their families, we must build a body of knowledge that integrates a rich and in-depth understanding of the how and why of services and supports with our more developed knowledge of the what, when, how many, and how much. It is hoped that this review of qualitative research presented at the RTC conferences from 1988 through 2003 will contribute to our understanding of the contributions of these methods and further suggest ways of closing the gap between what we know and what we hope to learn about serving children with serious emotional disturbance and their families.

Reference

Stroul, B. A. & Friedman, R. M. (1994). *A system of care for children and youth with severe emotional disturbances.* (Revised edition). Washington, DC: Georgetown University Child Development Center, CASSP Technical Assistance Center.

.

CONTRIBUTING AUTHORS

Sharon Hodges, Ph.D.
Department of Child and Family Studies, Louis de la Parte Florida Mental Health Institute, 13301 Bruce B. Downs Blvd., MHC 2437, University of South Florida, Tampa, FL 33612, 813-974-6460, fax: 813-974-7563, e-mail: shodges@fmhi.usf.edu

Allison Pinto, Ph.D.
Department of Child and Family Studies, Louis de la Parte Florida Mental Health Institute, 13301 Bruce B. Downs Blvd., MHC 2222, University of South Florida, Tampa, FL 33612, 813-974-9179, fax: 813-974-7563, e-mail: apinto@fmhi.usf.edu

Mario Hernandez, Ph.D.
Department of Child and Family Studies, Louis de la Parte Florida Mental Health Institute, 13301 Bruce B. Downs Blvd., MHC 2437, University of South Florida, Tampa, FL 33612, 813-974-4637, fax: 813-974-7563, e-mail: mhernandez@fmhi.usf.edu

Caitlin Uzzell, B.A.

Assessing Outcomes Over Time: Questioning Measurement Precision

Ann Doucette

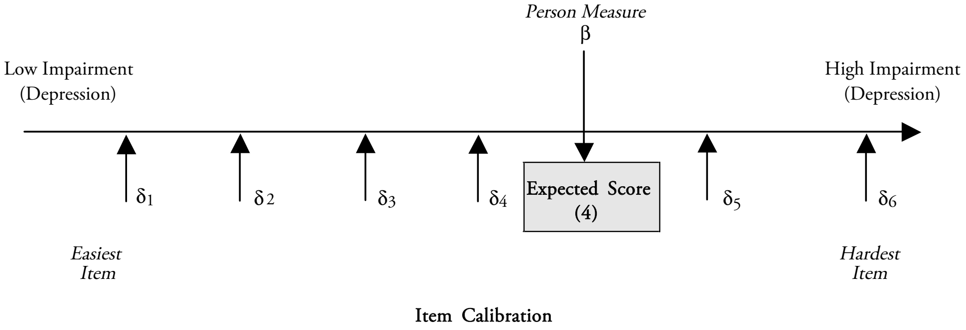
Introduction

Our interest in the psychometric properties of measures is essentially an assertion that measurement is not error-free. We typically rely on reliability estimates as the degree to which measurement is free from error. The more reliable a measure is, the greater the confidence we have in the score obtained on that measure. However, as a group is exposed to treatment and intervention, the group is hypothesized to move toward improvement reflected by a change in scores. While reliability may remain consistent as the relative position of individuals in the groups remains stable as their scores move on the measure as a group, the measure may not adequately assess individuals who have experienced improvement. Most assessment measures used in behavioral healthcare are developed using Classical Test Theory (CTT; Lord & Novick, 1968). From a CTT perspective, it is important to note that the error associated with measurement is equally applicable to all respondents. Intuitively this assumption is not reasonable. Educational assessment provides a clear-cut illustration—more error would be anticipated for low ability students on a difficult test than would be expected for high achieving children on the same test, yet the error of measurement would be equal for both groups of students under the CTT model. In addition, reliability cannot be generalized beyond the sample on which it is estimated—reliability is *sample specific*. This limits the usefulness of measures when they are applied to populations beyond the samples on which they were developed, and may have consequences as samples change over time as a result of exposure to interventions. Many studies assume the reliability reported from normative samples, and fail to report the reliability of measures applied to study samples or consider that reliability may change as the sample is studied over time.

Item Response Theory (IRT) is an alternative to CTT. IRT is a model-based measurement, latent-trait theory approach that addresses many of the shortcomings of CTT (Embretson & Reise, 2000; Embretson, 1996). IRT provides item estimates that are not sample dependent (Wright & Douglas, 1977). For example, items on a depression scale would be identified in terms of levels of impairment and accordingly matched to persons having more or less of the attribute, in this case depression. Any person with severe depression would be expected to respond in a similar fashion. Figure 1 identifies the respondent's position on the impairment continuum: above the easier items indicate lower levels of impairment/milder-moderate depression, and below the two items indicating a higher level of impairment. A person with this impairment level would be expected to endorse the first four depression items, but would not likely endorse items five and six. For example, an adolescent diagnosed with mild to moderate depression would not be expected to endorse depression items indicating suicide and self-harmful behaviors; but a suicidal youth would be expected to endorse items indicating sadness and feeling blue. However, as the adolescent improves he/she would be expected to endorse fewer and fewer items at the difficult end of the continuum, which may result in an assessment that is unable to reflect continuing change because there are few or no items at the mildest end of the impairment continuum. Measures that tap narrow bands of impairment would be insensitive to other impairment levels (high or low depending on the measure) and the results of using such measures may be interpreted as ineffective interventions, as opposed to measurement artifacts.

This paper presents analyses that illustrate that measures may be sensitive only to specific segments of the construct that is being assessed. For example the Child Behavior Checklist (CBCL; Achenbach, 1991) *narrow-band scales* (e.g., anxious-depressed, delinquency, etc) assess moderate to high levels of impairment. As individuals improve, there are few if any items in the scale that allow respondents to demonstrate continued improvement.

Figure 1
Defining Psychological Consturcts using IRT Model-Based Measurement



Method

A secondary analysis of existing data was used to illustrate the precision of measures used in longitudinal studies. The Anxiety-Depressed narrow band scale on the Child Behavior Checklist (CBCL; Achenbach, 1991) was used in the following analyses. The CBCL specifies that the caregiver report on symptoms and behaviors of the youth in his/her care over the last six months. Baseline, six and 12 month caregiver responses to the CBCL Anxiety-Depressed scale were used in the IRT analyses. Data samples included families participating in the Nashville Connection (*N* = 103; no missing CBCL data), a Substance Abuse and Mental Health Services Administration/Center for Mental Health Services-funded system of care initiative, and the Ft. Bragg Evaluation (*N* = 439; no missing CBCL items; Bickman, et al., 1995).

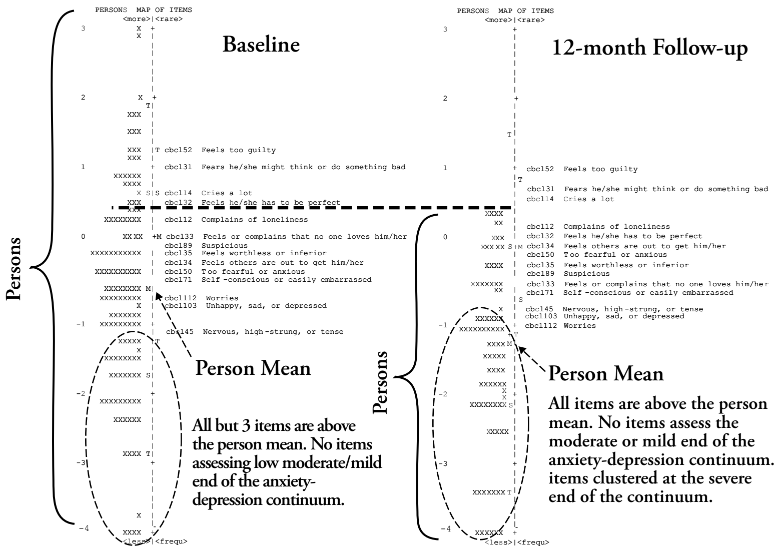
The Rasch measurement model (Rasch, 1966) was used to conduct an examination of the depression/anxiety narrow band scale of the CBCL. Rasch is often referred to as a one-parameter IRT model. While some psychometricians (Goldstein, 1979) argue that the use of two and three-parameter models provide better precision, many contend that the Rasch model more adequately addresses the fundamental issues associated with measurement development (Bond & Fox, 2001; Overton, 1999). Essentially, the Rasch model questions how well empirical data (measure scores/response) fit in terms of the measurement model constraints. Two and three parameter models ask an opposite question: how can additional parameters be manipulated to increase model fit to the available data? The additional parameters are item discrimination (2-parameter) and guessing (3-parameter).

Results

Figure 2 below illustrates that many children have improved to the extent that there are no items on the CBCL Depression-Anxiety scale that are sensitive to assessing continuing change. Data were collected over a 36-month period in both the Nashville Connection and the Ft. Bragg Evaluation. The trend illustrated in Figure 2 continues, as the number of children with impairment at the level tapped by the CBCL Anxiety-Depressed decrease as a result of improvement. While these children may continue to have more moderate and mild problems, the CBCL is not sensitive to assess this level of difficulty.

The Rasch analysis provides an estimate called a separation index, the number of statistically distinct strata of “trait difficulty” (anxious-depressed) that can be represented in the sample using this measure. At baseline the separation index is 2.65, indicating that measure can reliably identify three statistically distinct groups of youth: those youth with mild, moderate and severe impairment. At six- and twelve-month follow-ups the separation index decreases to 1.90 and 1.85 respectively, indicating that there are two statistically distinct groups, those with some clinical concerns and those below the clinical thresholds.

Figure 2
Sensitivity of the CBCL Depression-Anxiety Scale



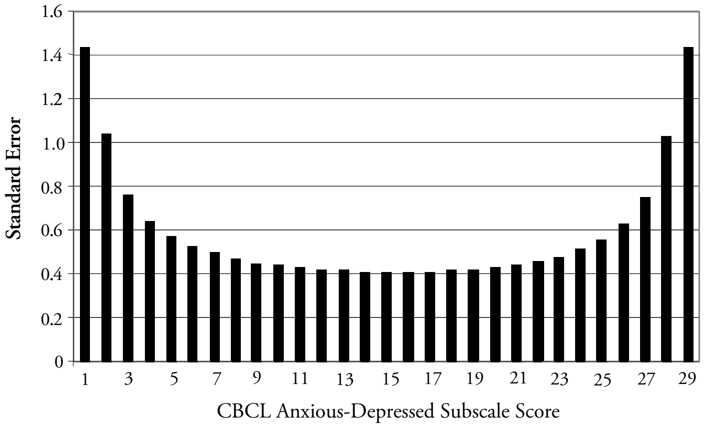
```
    PERSONS  MAP OF ITEMS
          <more>|<rare>
 3            X  +
              X  |
                 |
                 |
                 |
                 |
                 |
 2             X  +
                T|
           XXX   |
                 |
           XXX   |
                 |
           XXX   |T cbcl52  Feels too guilty
           XXX   |
 1               +  cbcl31  Fears he/she might think or do something bad
        XXXXXX   |
          XXXX   |
             X S|S cbcl14  Cries a lot
           XXX   |  cbcl32  Feels he/she has to be perfect
           XXX   |
      XXXXXXXX   |  cbcl12  Complains of loneliness
                 |
 0        XX XX  +M cbcl33  Feels or complains that no one loves him/her
                    cbcl89  Suspicious
   XXXXXXXXXXX   |  cbcl35  Feels worthless or inferior
                 |  cbcl34  Feels others are out to get him/her
    XXXXXXXXXX   |  cbcl50  Too fearful or anxious
                    cbcl71  Self-conscious or easily embarrassed
      XXXXXXXX M|
     XXXXXXXXX   |  cbcl112  Worries
             X   |  cbcl103  Unhappy, sad, or depressed
       XXXXXXX   |
-1   XXXXXXXXX   +
                 |  cbcl45  Nervous, high-strung, or tense
         XXXXX   |T
             X   |
      XXXXXXXX   |
                 |
       XXXXXXX S|
                 |
-2               +
     XXXXXXXXX   |
                 |
        XXXXXX   |
                 |
                 |
                 |
          XXXX T|
-3               +
                 |
                 |
                 |
                 |
                 |
             X   |
                 |-
-4        XXXX   +
         <less>|<frequ>
```

```
    PERSONS  MAP OF ITEMS
          <more>|<rare>
 3               +
                 |
                 |
                 |
                 |
                 |
                 |
 2               +
                 |
                 |
                 |
                T|
                 |
                 |
 1               +  cbcl52  Feels too guilty
                 |T
                 |  cbcl31  Fears he/she might think or do something bad
                 |  cbcl14  Cries a lot
                 |
          XXXX   |
            XX   |  cbcl12  Complains of loneliness
 0         XXX   |  cbcl32  Feels he/she has to be perfect
        XXX XX S+M cbcl34  Feels others are out to get him/her
                 |  cbcl50  Too fearful or anxious
          XXXX   |  cbcl35  Feels worthless or inferior
                 |  cbcl89  Suspicious
       XXXXXXX      cbcl33  Feels or complains that no one loves him/her
            XX   |  cbcl71  Self-conscious or easily embarrassed
                 |S
             X   |  cbcl45  Nervous, high-strung, or tense
        XXXXXX   |  cbcl103 Unhappy, sad, or depressed
-1  XXXXXXXXXX   +  cbcl112 Worries
                 +T
          XXXX M|
         XXXXX   |
          XXXX   |
        XXXXXX   |
             X
-2           X   +
      XXXXXXXX S|
                 |
         XXXXX   |
                 |
-3               +
                 |
                 |
       XXXXXXX T|
                 |
                 |-
-4      XXXXXX   +
         <less>|<frequ>
```

Forty-five percent of youth at baseline assessment had scores below the clinical cutoff indicating mild to moderate impairment; 76% of those youth with mild/low moderate scores maintained that status between baseline and follow-up (6 six and 12 months) assessments. The stable scores may be interpreted to mean these youth made no progress during the 12 month period, as opposed to reporting that the CBCL was no longer sensitive for youth at the milder ends of the impairment continuum.

The Rasch analyses indicate that items differentially contribute to scale scores, and that error contributes to items. Figure 1 illustrates that items such as feeling guilty, cries a lot, loneliness, and fears about having to be perfect tap a higher level of impairment than do items such as unhappiness and nervousness. Additionally, as Figure 3 depicts, error is associated with the extreme (high/low) CBCL Anxiety-Depressed scores. It is important to remember that errors of measurement are assumed to be uncorrelated.

Figure 3
Standard Error: CBCL Anxious-Depressed Subscale Scores



Discussion

We typically assume that measures sensitive to the initial, more severe levels of impairment continue to be sensitive to later moderate and/or mild impairment levels. Often we fail to consider whether there are sufficient items at the milder end of the continuum to provide an opportunity to demonstrate improvement for youth with low baseline scores, or youth making rapid improvement. As children improve the measures we depend upon to reflect the continued effectiveness of interventions may not be sensitive beyond specific points on the impairment continuum of constructs that are of interest in children's behavioral healthcare. The lack of measurement sensitivity may compromise the meaningfulness and the actionability of the information that is yielded by these measures. For example, a child having stable scores may be misinterpreted as not making progress, when it is an artifact of the measures and not a reflection of the intervention or treatment. While the analyses presented in this brief summary focus on anxiety and depression, the same trends were evident for externalizing behaviors (delinquency, conduct and ADHD) narrow-band CBCL scales using the same analytic approach.

In conclusion, the measures used at baseline in longitudinal studies may have compromised utility in providing meaningful and actionable information as the study progresses and youth improve. Measures having sensitivity across a broad impairment continuum would undoubtedly be lengthy and burdensome for the caregiver and/or youth to complete. However, measures developed and/or later calibrated using an IRT approach can be equated, providing researchers with an opportunity to link measures assessing severe and moderate impairment levels with measures addressing milder levels that more sensitively reveal continued improvement. As stated earlier, the lack of demonstrated improvement may well be an artifact of the measures used than a reflection of ineffective intervention and/or treatment.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist. 4-18 and 1991 profile.* . Burlington, VT: University of Vermont Department of Psychiatry.

Bickman, L., Guthrie, P., Foster, E., Lambert, E., Summerfelt, W., Breda, C., & Heflinger, C. (1995). *Evaluating managed mental health services: The Fort Bragg experiment.* New York: Plenum.

Bond, T., & Fox, C. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Embretson, S. (1996). The new rules of measurement. *Psychological Assessment, 8*, 314-349.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc., Publishers.

Goldstein, H. (1979). The mystifications of assessment. *Forum for the Discussion of New Trends in Education, 22*(14-16).

Lord, F., & Novick, M. (1968). *Statistical theories of mental tests.* New York: Addison-Wesley.

Overton, W. F. (1999, August). *Construct validity: A forgotten concept in psychology?* Paper presented at the Annual Meeting of the American Psychological Association, Boston.

Rasch, G. (1966). An item analysis which takes individual differences into account. *British Journal of Mathematical and Statistical Psychology, 19*, 49-57.

Wright, B. D., & Douglas, G. A. (1977). Best procedures for sample-free item analysis. *Applied Psychological Measurement, 1*, 281-294.

CONTRIBUTING AUTHORS

Ann Doucette, Ph.D.
Senior Research Scientist, Center for Health Services Research and Policy, The George Washington University Medical Center, 2021 K St., NW, Suite 800, Washington, DC 20006, 202-530-3921, fax: 202-315-3304, e-mail: Doucette@gwu.edu

Exploring Relationships among Child Outcomes in Kentucky's System of Care

Carla Crane-Mahan
Vestena Robbins

Acknowledgements: Support for a portion of this research related to the Bridges Project was provided by a grant from the Substance Abuse and Mental Health Services Administration Center for Mental Health Services (4 HS5 SM52273-06-2).

Introduction

The assessment of child and family outcomes has become a fundamental component of children's behavioral health service delivery. A growing number of studies have been conducted to examine the impact of service on outcomes, such as adaptive functioning or symptom severity (Epstein, Kutash, & Duchnowski, 1998). Despite an increased focus on outcome, studies of the relationship among outcomes for children with serious emotional and behavioral disabilities are scant (Rosenblatt & Rosenblatt, 2002). The current study examined changes in functioning and hopefulness across time from both the child and parent perspective, and the predictive properties of selected outcome measures on improved functioning after six months in service. Child outcomes were derived from two behavioral health programs in Kentucky: the Bridges Project, a school based mental health program in eastern Kentucky, and IMPACT Plus, a statewide Medicaid funded community-based program. Determining how child outcomes, such as hopefulness and functioning, interact is a critical step in improving services for children and families.

Study 1 – Bridges Project

Study 1 Method

Participants. This study included a subset of 115 children and youth with serious emotional disabilities (72 males and 43 females) served in the Bridges Project for whom six-month follow up data were available. As part of the national evaluation of the Center for Mental Health Services-funded grant communities, a comprehensive evaluation system is in place, including the gathering of longitudinal outcome data. Child and family outcome data were gathered for children between the ages of five and 17.5 years at program entry whose caregiver provided consent for participation. The average age of the sample was 11.4 years. Mirroring the population of eastern Kentucky, the sample was predominantly White/NonHispanic (96%). The most common youth diagnoses were Attention Deficit Hyperactivity Disorder (26%) and Oppositional Defiant Disorder (17%). Most families received Medicaid assistance (90%) and 43% of caregivers had completed high school or higher levels of education.

Measures and procedures

Symptomatology – Caregiver Report. The Child Behavior Checklist (CBCL; Achenbach, 1991) measures the caregiver perspective of a child's competencies and problem behaviors. The problem behavior scale includes 118 items yielding a Total raw score that was used to measure the caregivers' perspective of their child's symptomatology.

Functioning – Caregiver Report. The Child and Adolescent Functional Assessment Scale (CAFAS; Hodges, 1990) yields a measure of functioning across life domains. The total functioning score was used to measure the caregivers' perspective of their child's functioning.

Strengths – Caregiver Report. The Behavioral and Emotional Rating Scale (BERS; Epstein & Sharma, 1997) is a 52-item measure assessing five strength dimensions. The Strengths Quotient was used to measure the caregivers' perspective of their child's overall strengths.

Caregiver Strain – Caregiver Report. The Caregiver Strain Questionnaire (CGSQ; Brannan, Heflinger, & Bickman, 1998) uses a 21-item scale to measure the level of strain experienced by the caregiver while caring for a child with a serious emotional disability. The 21-item total score was used to measure the caregivers' self-report of strain.

A paired-samples *t*-test was used to compare mean functioning scores as measured by the CAFAS at intake and six-month follow up. A multiple regression was used to explore predictive properties of symptomatology, strengths, and caregiver strain on functioning at six-month follow up. The influence of baseline functioning on follow up functioning was controlled by including it as a predictor in the model.

Study 1 Results

Change in functioning. Caregivers (n = 115) reported a significantly lower (better) mean functioning score from baseline (M = 100) to six-month follow up (M = 77.6), $t(114) = 6.18$, $p = .000$.

Predictive properties. Correlations among the variables included in the regression analysis were analyzed to examine the strength of covariance among the constructs (Table 1). As expected, all correlations were statistically significant with child strengths being negatively associated with child functioning ($r = -.47$, $p < .001$) and negatively associated with caregiver strain ($r = -.36$, $p < .001$). A multiple regression was conducted using the CAFAS functioning score at six months as the dependent variable. Baseline functioning and caregiver strain contributed significantly to the prediction of functioning; however, after controlling for baseline functioning, only an additional 6% of the variability in functioning at six months was predicted (see Table 2). Baseline functioning had the greatest influence on follow up functioning ($\beta = .39$), followed by caregiver strain ($\beta = .27$).

Table 1
Intercorrelations Between Variables Included in Study 1 Multiple Regression Model (*n* = 115)

Variable	*1*	*2*	*3*	*4*	*5*
1. Functioning T1	–	.54***	.53***	-.47***	.49***
2. Functioning T2		-	.39***	-.30**	.48***
3. Symptoms			–	-.48***	.52***
4. Strengths					-.36***
5. Strain					

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

Table 2
Summary of Step Wise Regression Analysis for Study 1 Variables Predicting Youth Functioning at 6-Month Follow Up (*n* = 115)

Variable	*Unstandardized Coefficients*		*Standardized Coefficients*
	B	*SE*	β
Step 1			
Functioning T1	.54	.08	.54*
Step 2			
Functioning T1	.39	.10	.39*
Caregiver Strain	.26	.09	.27*
Symptoms	.04	.10	.04
Strengths	-.001	.09	.001

Notes:
Adj. R^2 for Model 1 = .29, $F(1, 113) = 47.4$, $p = .000$
R^2 change for Model 2 = .06, $F(3, 110) = 3.48$, $p = .02$
*$p < .05$.

Study 2 – IMPACT Plus

Study 2 Method

Participants. This study included a subset of 84 youth (45 males and 39 females) served in IMPACT Plus. The sample was restricted by including youth who were old enough, and had parent consent, to complete outcome measures. All participants received Medicaid assistance and all participants had a severe emotional disability contributing to behavioral difficulties across home, school, and community settings. The sample of 84 included 75 Whites, five Blacks, and four unknown. Standardized outcome data are collected by a service coordinator at intake and six month intervals.

Measures and procedures.

Hopefulness and Functioning – Parent Report. The Ohio Scales, Parent Short Form (Ohio Scales-P; Ogles, Melendez, Davis, & Lunnen, 1999) yields a parent hopefulness subscale (4 items) and a child functioning subscale from the caregivers' perspective (20 items).

Hopefulness and Functioning – Youth Report. The Ohio Scales, Youth Short Form (Ohio Scales-Y; Ogles et al., 1999) parallels the parent version of the same scale, yielding a child hopefulness subscale and a child functioning subscale from the child's perspective.

Symptom Severity – Parent Report. As in Study 1, the CBCL total raw score was used to measure symptom severity.

A paired-samples *t*-test was used to compare mean child functioning scores from both the parent perspective and the child perspective as measured by respective short-form versions of the Ohio Scales at intake and six-month follow up. A paired-samples *t*-test was also used to compare the mean child hopefulness scores with the mean parent hopefulness scores as measured by the respective short-form versions of the Ohio Scales. Lastly, a multiple regression was used to explore predictive properties of child's perceived hopefulness, and parent's perceived symptom severity on functioning from the child perspective at follow up. As in Study 1, the influence of baseline functioning on follow up functioning was controlled by including it as a predictor in the model.

Study 2 Results

Changes in functioning and hopefulness. Youth (n = 84) reported a significantly higher mean functioning score from baseline (M = 55.01) to six-month follow up (M = 57.87), $t(83) = -2.08$, p = .041, as measured by Ohio Scales-Y; whereas caregivers (n = 84) reported a nonsignificant higher mean child functioning score from baseline (M = 42.02) to six-month follow up (M = 45.00), as measured by Ohio Scales-P. Youth also self-reported a higher mean hopefulness score at baseline (M = 17.88) when compared to their parent report of mean hopefulness (M = 14.77). Little change occurred with the youth mean hopefulness score at six-month follow up (M = 17.45); however, the parent mean hopefulness score had a statistically significant positive change (M = 16.42), $t(83) = -.288$, p = .005.

Predictive properties. Correlations among the variables included in the regression analysis were analyzed to examine the strength of covariance among the constructs (see Table 3). Most correlations were statistically significant with child hopefulness at baseline being positively associated with child functioning at baseline (r = .71, $p < .01$) and positively associated with child functioning at six months (r = .47, $p < .01$). A multiple regression was conducted using the Ohio Scale-Y functioning score at six months as the dependent variable. As expected, baseline functioning scores were predictive of functioning scores at six months; however, no other variables were significant predictors of functioning at six months (see Table 4).

Table 3
Intercorrelations Between Variables
Included in Study 2 Multiple Regression Model ($n = 84$)

Variable	*1*	*2*	*3*	*4*
1. Symptoms	–	-.19	-.19	-.13
2. Functioning T1		–	.71**	.59**
3. Hopefulness			–	.47**
4. Functioning T2				–

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

Table 4
Summary of Step Wise Regression Analysis for Study 2 Variables
Predicting Youth Functioning at 6-Month Follow Up ($n = 84$)

Variable	*Unstandardized Coefficients*		*Standardized Coefficients*
	B	*SE*	β
Step 1			
Functioning T1	.63	.09	.60*
Step 2			
Functioning T1	.55	.13	.53*
Symptoms	-.02	.09	-.02
Hopefulness	.10	.13	.10

Notes:
Adj. R^2 for Model 1 = .35, $F(1, 82) = .00$, $p < .001$
R^2 change for Model 2 = .36, $F(2, 80) = .73$, $p < .001$
*$p < .05$.

Summary of Study 1 and Study 2

Significant correlations between different constructs coupled with divergent perspectives between caregivers and their children support the complexity of various relationships within the context of behavioral health treatment. At the same time, divergent respondent perspectives should provide greater insight for more comprehensive treatment planning (Rosenblatt & Rosenblatt, 2002). The significant predictive value of caregiver strain validates the importance of family support within the treatment planning process (Heflinger, Northrup, Sonnichsen, & Brannan, 1998).

Although divergent respondent perspectives (i.e. parent versus child) can be useful for individual treatment planning (Rosenblatt & Rosenblatt, 2002), more sophisticated research methodologies are needed to better understand how differing perspectives and interactions thereof should be considered when evaluating service outcomes. As such, analyses of differential patterns of change by subgroups (e.g., level of severity, by diagnoses, by gender) may provide more specific information for interpreting outcome data. Due to the small sample sizes, the number of predictor variables included in each study was limited.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.

Brannan, A. M., Heflinger, C. A., & Bickman, L. (1998). The Caregiver Strain Questionnaire: Measuring the impact on the family living with a child with serious emotional disturbance. *Journal of Emotional and Behavioral Disorders, 5*(4), 212-222.

Epstein, M., Kutash, K., & Duchnowski, A. (Eds.). (1998). *Outcomes for children & youth with behavioral and emotional disorders and their families: Program & evaluation best practices*. Austin, TX: Pro-ed.

Epstein, M., & Sharma, J. (1997). *Behavioral and Emotional Rating Scale: A strength-based approach to assessment*. Austin, TX: PRO-ED.

Heflinger, C. A., Northrup, D. A., Sonnichsen, S. E., & Brannan, A. M. (1998). Including a family focus in research on community-based services for children with serious emotional disturbance: Experiences from the Fort Bragg Evaluation Project. In M. Epstein, K. Kutash, & A. Duchnowski (Eds.), *Outcomes for children & youth with behavioral and emotional disorders and their families: Program & evaluation best practices* (pp. 261-293). Austin, TX: Pro-ed.

Hodges, K. (1990). *The Child and Adolescent Functional Assessment Scale*. Ann Arbor, MI: Author.

Ogles, B. M, Melendez, M. S., Davis, D. C., & Lunnen, K. M. (1999). *The Ohio Youth Problems, Functioning, and Satisfaction Scales* (Short Form): User's Manual. Athens, OH: Ohio University.

Rosenblatt, A., & Rosenblatt, J.A. (2002). Assessing the effectiveness of care for youth with severe emotional disturbances: Is there agreement between popular outcome measures? *Journal of Behavioral Health Services & Research, 29*(3), 259-273.

CONTRIBUTING AUTHORS

Carla Crane-Mahan, Ed.S., A.B.D.
Program Evaluation Specialist, 502-564-4860, fax: 502-564-3844,
e-mail: carla.mahan@ky.gov

Vestena Robbins, Ph.D.
Research Specialist, Research and Practice Improvement Administrator, 502-564-4860,
fax: 502-564-3844, e-mail: vestena.robbins@ky.gov

All Authors: Kentucky Department for Mental Health and Mental Retardation
Services, Division of Administration and Financial Management, 100 Fair Oaks Lane,
4E-A, Frankfort, KY 40621