
*Measurement as Communication in
Outcomes Management: The Child and
Adolescent Needs and Strengths (CANS)*

John S. Lyons and Dana Aron Weiner
Northwestern University

Melanie Buddin Lyons
Buddin Praed Foundation

INTRODUCTION

The children's mental health services system has fallen short of the ideal. This is especially true in the public sector, where children often receive inadequate or appropriate treatment. Several large-scale research projects have found that improving access and coordination within these service systems does not necessarily result in improved outcomes for children (Bickman, Lambert Andrade, & Penalzoa, 2000). These findings call into question the effectiveness of existing services (Bickman, Noser, & Summerfelt, 1999; Bickman, Lambert, Andrade, & Penalzoa, 2000; Weiss, Catron, Harris, & Phung, 1999).

To address this shortcoming, there is growing interest in evidence-based treatment planning (Hoagwood, Burns, Kiser, Ringeisen, & Schoenwald, 2001). However, such treatment planning is predicated on reliable, valid, and feasible outcome measurement methods. Flawed measures implemented within complex systems along with inconsistent involvement of stakeholders have impeded the outcomes measurement process (Lambert, Ogles, & Masters, 2000; Lyons, Howard, O'Mahoney, & Lish, 1997).

Outcome measures should be able to assess clinical status, well-being, level of functioning and quality of life (Lambert, Hansen, & Finch, 2001; Sederer, Dickey, & Hermann, 1996; Steinwachs, Flynn, Norquist, & Skinner, 1996). To be applicable in real-world service settings, measures must balance the need for comprehensiveness with the need for efficiency and brevity. Although the standard goals for children's services can be quite simple (e.g., at home, in school, out of trouble; Rosenblatt, 1993), the full range of outcomes requires more sophisticated measurement to support decision making and quality management. At the same time, measures must be brief and easy to use to facilitate their use by busy providers.

Stakeholders in the children's mental health services system (e.g., child and family, providers, administrators, evaluators) have different priorities for target outcomes and for methods of measurement (Fischer, Shumway, & Owen, 2002). Ideally, the selection of target outcomes for measurement should be a collaborative process involving representatives of various perspectives. Evaluators are likely to advocate for

use of the most reliable and empirically valid measures, but translating these from research into clinical practice can be problematic. Administrators might emphasize the cost and utility of measures for quality improvement and accountability. Service providers might stress meaningfulness and ease of use. Service recipients might advocate for measures with relevance and respect for consumers. Selected outcome measures should honor all of these perspectives.

In response to these complexities, we have evolved an approach to the design of outcome measures that is distinct from psychometric approaches. This paper will briefly review these traditional approaches to outcomes measurement and present the communication model of measurement. The Child and Adolescent Need and Strengths (CANS) measure will be used to illustrate the development and applications of a measure developed from a communication perspective.

APPROACHES TO OUTCOMES MEASUREMENT

Although psychometric theory has resulted in significant contributions in measurement for research in behavioral health, the application of these approaches to the development of measures for outcomes management has some notable drawbacks. They may lack one or more of the prerequisites of comprehensiveness, brevity, relevance, and real-world validity needed to apply findings in meaningful ways in a service delivery environment. In this section, we review psychometric and clinimetric approaches to measurement development and present an alternative approach: a communication model.

Psychometric Approaches

Most of the currently available outcome measures have been developed from either classical test theory (Anastasi, 1968; Nunally 1976) or item response theory (Rost & Langeheine, 1997).. Although the theoretical assumptions and statistical approaches of item response theory are dramatically different from those of classical test theory (Drasgow & Schmitt, 2002), both require multiple items to measure a single construct. Also, both implicitly value precision in measurement over brevity, although item response theory does offer strategies for identifying the most efficient number of items to accurately assess a construct. Multi-item measures can be an implementation challenge in the mental health services context, where multiple dimensions must be assessed in a brief time.

The majority of mental health services aim to impact multiple dimensions of an individual's functioning. For example, crisis intervention may reduce suicide risk and the likelihood of violence, improve self-care, mobilize resources, and stabilize symptoms. Outpatient therapy is seen as having a potential impact on symptoms, subjective well-being, and functioning. Thus, regardless of the specific intervention, comprehensive mental health outcomes measurement requires a multidimensional approach to capture the complexity of the impact of services (Lambert, *Ogles & Masters*, 2000).

Outcome measures often must be completed within a brief time frame, as neither clinicians nor service recipients are inclined or able to spend a great deal of time completing measures. Long measures with overlapping or redundant items, such as those valued by classical test theory (i.e, high internal consistency reliability), can be too time-consuming for respondents to complete. This can result in missing data that

threaten the validity and utility of outcomes assessment for services, programs, systems

The absence of clear relevance to treatment planning can be problematic for the use of a measure. A clinician who does not see the relevance of a measure for his or her work is unlikely to devote time and attention to completing it. Service recipients have the same perspective. Both perspectives demand that a measure have face validity which is less valued by both classical test theory and item response theory.

Clinimetrics

In response to the problems with psychometric approaches previously identified, measurement developers in medicine have utilized a theoretical approach referred to as clinimetrics (Feinstein, 1986). The stated goal of clinimetrics is to convert "intangible clinical phenomenon into formal specified measurement" (Apgar, 1966, p. 125). Virginia Apgar is generally credited with developing the first measure from this perspective (Apgar, 1966). First introduced in 1953, the "Apgar" is routinely utilized as a health status measure at birth. Clinimetric tools are now quite common in medicine (e.g., Bloem, Beckley, van Hilten, & Roos, 1998; Stone, Salonen, Lax, Payne, Lapp, & Inman, 2001; Gates, 2000; Hoff, van Hilten, & Roos, 1999).

Feinstein (1999) has enumerated six principles that speak to the differences between clinimetrics and psychometrics:

1. Selection of items is based on clinical rather than statistical criteria.
2. Factors need not be weighted.
3. Scoring is simple and readily interpretable.
4. Variables are selected to be heterogeneous rather than homogeneous.
5. Measures must be easy for clinicians to use.
6. Face validity is required and subjective states are not measured because of limited sources of observation.

Current applications of clinimetrics have some notable limitations (Marx, Bombardier, Hogg-Johnson, & Wright, 2000; Zyzanski & Perloff, 1999). Many clinimetric scales consist of a single item. When complex phenomena are described, a single item fails to communicate complexity. For example, a Childhood Global Assessment Scale (Endicott, Spitzer, Fleiss, & Cohen, 1976) that ranges from 0 to 100 can provide a general sense of how a child is doing, but it cannot capture individual dimensions of functioning that are useful to clinicians. In addition, single-item measures are not particularly sensitive to change. For these reasons, Zyzanski and Perloff (1999) and others have called for an integration of clinimetric and psychometric approaches to measurement.

Service delivery settings have very different priorities than do research settings. Accommodating these technical and contextual requirements requires a broad scope models of measurement. This model must include guidelines for utility in clinical operations as well as reliability and validity. Measures intended for the evaluation of treatment outcomes should be easy to use and brief. Their output should be clear, unambiguous, relevant, easy to translate into service planning recommendations, and accessible to providers, consumers, and policy makers. Neither classical test theory nor item response theory fully informs the development of measures meeting these requirements in outcome management applications.

-

A Communication Model

Measures used within service delivery operations must be able to easily and accurately communicate relevant results. Feedback about performance is central to quality improvement and outcomes management (Clark, Schyve, Lepoff, & Ruess, 1994; Koike, Unutzer, & Wells, 2002; Krulish, 2002; Schiff & Goldfield, 1994). This feedback requires the integration of measurement into the information feedback loop (Lichtman & Appleman, 1995). Thus, it can be argued that communication is a primary goal of measurement in clinical settings (Howard, Moras, Brill, Martinovich, & Lutz, 1996; Lueger, Howard, Martinovich, Lutz, Anderson, & Grissom, 2001). This includes communication between recipients and providers about perceptions of clients' needs; communication between providers, program administrators, and evaluators about clinical status; communication between providers and payers about medical necessity for benefits from services; and communication among providers and other partners about the goals and outcomes of an integrated children's system of care.

Communication theory is a broad and diverse field that informs improvements in outcomes management strategies. Although this chapter does not present a comprehensive review of the communication theory literature, there are three areas within communication theory that have particularly influenced the development of our approach to measurement development. The first construct is the theory of communicative action. Simply stated, communicative action is a consensus-based approach that relies on mutual definitions of how to reach a goal (Habermas & Seidman, 1989). Kihlstrom and Israel (2002) found that group leadership actions based on communicative action theory lead to greater openness to diversity and individual experiences. Friedland (2001) posited that communication forms the primary ecology of postindustrial communities; this logic is relevant for the children's system of care. According to this theory, the foundation for a system of care would be effective action-oriented communication based on consensus among the partners in that system.

Second, White (2001) highlighted communication as the basis for innovation in science. Within this context, the dissemination of evidence-based practices into the field requires consensus across the field that new practices are better than current practices. This kind of consensus cannot be reached by the publication of randomized clinical trials alone, but by ongoing interaction among service delivery, evaluation, and research (Drake, 2002).

Finally, Harris's work in organizational communication has laid out the importance of communication within business environments (Harris, 2002). He conceptualized communication as a nonlinear process that plays a central role in effective leadership, organizational development, and establishing an organization's culture. This is especially important within the children's mental health service system. Given the organizational complexity of most children's systems of care the system is in need of communication tools to serve these functions.

Application of the communication model within the children's services system will require the establishment of a common ground, and common language, among mental health, child welfare, juvenile justice, and school systems. With these goals in mind we have worked to develop a communication model of measurement. The model builds on some of the tenets of psychometric theory (high face and content validity, high interrater reliability, high concurrent and predictive validity) and the six principles of clinimetrics, adding three additional requirements:

1. All partners in the service delivery system of care should participate in the development and uses of the measure.

2. The goal of item selection is to include single items that represent each of the key constructs, identified by consensus, that inform good decision making and service planning in the service delivery operation.
3. The levels of each item should be directly translatable into action steps for treatment planning.

From the communication perspective, a good measure should be clear, concise, relevant, and comprehensive without being redundant. It should use common, understandable language and be easy to use. Most important, a measure should be useful for the three primary purposes for which one requires these tools in clinical practice: decision support, quality improvement, and outcomes monitoring.

Decision support strategies help ensure that clinically appropriate decisions are made consistently at key points in the service delivery process. Quality improvement activities ensure that potentially effective interventions are provided when needed and that needs are assessed accurately and in a timely fashion. Outcome monitoring efforts inform clinicians, administrators, and evaluators about the impact of an intervention or program. The measurement approach should allow for all three of these tasks to be accomplished for every case, program, and system. For the past decade, we have been working to develop outcomes management tools using a communication model that blends the strengths of psychometric and clinimetric strategies to measurement development. Perhaps the most widely used tool of this type is the Child and Adolescent Needs and Strengths (Lyons, 1999). Versions of the CANS have been developed to guide service delivery for children with mental health needs, developmental disabilities, child welfare, and juvenile justice involvement. In addition, a specific version for children 3 years old and younger has been developed. Development of the CANS is grounded, in part, in our prior work modeling decision making for hospital and residential services for children and adolescents (Lyons, Mintzer, Kisiel, & Shallcross, 1998).

CANS: DEVELOPMENT, STRUCTURE, AND FUNCTION

Development

As part of a major reform of the child welfare service system in Illinois, we assessed the extent to which psychiatric hospitalization and residential treatment services were used appropriately. For this purpose, we developed the Childhood Severity of Psychiatric Illness (CSPI). This measure was designed to assess the dimensions important to good clinical decision-making for intensive and expensive mental health service interventions. We have demonstrated the utility of the CSPI for reforming decision making for residential treatment (Lyons et al., 1998) and for quality improvement in crisis assessment services (Leon, Uziel-Miller, Lyons, & Tracy, 1999; Lyons, Kisiel, Mcan, Cohen, & Chesler, 1997). The strength of the measure has been that it has face validity and is easy to use, yet reliably provides sufficiently comprehensive information regarding the clinical status of the child or youth that can be translated into policy and treatment recommendations.

The Child and Adolescent Needs and Strengths builds on the conceptual approach of the CSPI but expands the assessment to include additional areas of needs and the assessment of strengths. The CANS was developed using focus groups with a variety of participants, including families, representatives of the provider community,

case managers, and staff. Each focus group was charged with answering the question: "What do we need to know about children and their families to effectively plan and monitor services?" Items were then developed and sent back to focus group participants, who gave feedback on the wording and the levels of each item. The item selection process used a clinimetric approach to begin building a common assessment language. Thus, consensus of those using the tool on the inclusion of an item was more important than intercorrelations among items. This process places measure construction decisions in the hands of the people who will be utilizing the tool.

Structure

The CANS consists of dimensions that are either areas of need or areas of strength. Anchors, standard across these dimensions, are used to rate the level of each need or strength. The anchors are worded in terms of the level of intervention needed, which enables the CANS to produce information that is instantly relevant for service planning. A rating of 2 or 3 on a need indicates that this need should be addressed in service planning.

Table 17.1 provides a summary of the dimensions of the CANS-MH. Unless otherwise specified, each rating is based on the child's functioning in the last 30 days. Each of the dimensions is rated on a 4-point scale after routine service contact, a semistructured interview, or a review of notes from case files. Raters are encouraged to use corresponding "action levels" to help make the determination between two adjacent ratings. For example, a rater torn between a rating of 1 or 2 should consider whether the need requires action or continued observation.

Table 17.2 illustrates the CANS-MH using two items, Psychosis and Danger to Self, from the CANS-MH. For the Psychosis item, "No evidence" indicates the absence of any signs of thought disorder (e.g., hallucinations, delusions, bizarre behavior, or

TABLE 17.1
Domains and Individual Items of the CANS-MH

Problem Presentation: Psychosis	Treatment
Attention deficit/impulse control	Transportation
Depression/anxiety	Service permanence
Oppositional behavior	Caregiver Needs & Strengths
Antisocial behavior	Physical/behavioral
Substance abuse	Supervision
Adjustment to trauma	Involvement with care
Situational consistency of problems	Knowledge
Temporal consistency of problems	Organization
Risk Behaviors: Danger to Self	Residential stability
Danger to other	Resources
Elopement	Safety
Sexually abusive behavior	Strengths: Family
Social behavior	Interpersonal
Crime/delinquency	Relationship permanence
Functioning: Intellectual Developmental	Education
Physical/ medical	Vocational
Family	Well-being
School/day care	Spiritual/religious
Care Intensity & Organization	Talents/Interests
Monitoring	Inclusion

TABLE 17.2
 Two Example Items From the Child and Adolescent Needs and Strengths

Psychotic Symptoms

This rating is used to describe symptoms of psychiatric disorders with a known neurological base. DSM-N disorders included on this dimension are schizophrenia and psychotic disorders (unipolar, bipolar, NOS). The common symptoms of these disorders include hallucinations, delusions, unusual thought processes, strange speech, and bizarre or idiosyncratic behavior.

- 0 This rating indicates a child with no evidence of thought disturbances. Both thought processes and content are within normal range.
- 1 This rating indicates a child with evidence of mild disruption in thought processes or content. The child may be somewhat tangential in speech or evidence somewhat illogical thinking (e.g., age inappropriate). This also includes children with a history of hallucinations but none currently. The category would be used for children who are below the threshold for one of the DSM-IV diagnoses previously listed.
- 2 This rating indicates a child with evidence of moderate disturbance in thought process or content. The child may be somewhat delusional or have brief intermittent hallucinations. The child's speech may be at times quite tangential or illogical. This level would be used for children who meet the diagnostic criteria for one of the disorders previously listed.
- 3 This rating indicates a child with a severe psychotic disorder. Symptoms are dangerous to the child or others.

Danger to Self

This rating describes both suicidal and significant self-injurious behavior. A rating of 2 or 3 would indicate the need for a safety plan.

- 0 Child has no evidence or history of suicidal or self-injurious behaviors.
- 1 History of suicidal or self-injurious behaviors but no self-injurious behavior during the past 30 days.
- 2 Recent (last 30 days) but not acute (today) suicidal ideation or gesture. Self-injurious in the past 30 days (including today) without suicidal ideation or intent.
3. Current suicidal ideation and intent in the past 24 hours.

bizarre thinking). A 1 could be used either where there is suspicion of these symptoms or after a youth who was actively psychotic has been stabilized for some time through medications. A 2 indicates that a youth has clear evidence of these problems and requires active treatment. A 3 indicates a level of these symptoms that requires immediate intervention to prevent harm. The CANS-MH is not a diagnostic tool, but because diagnoses also play a role in communicating a child's needs, the CANS is designed to be consistent with the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV, 1994). Thus, for Psychosis, a 2 or 3 is consistent with the presence of one of the disorders whose symptoms are covered by this item.

In the second item in Table 17.2, Danger to Self, no evidence or history of suicidal behavior means that no action is required in this regard. Because we know that children who have attempted suicide in the past are more likely to repeat this behavior, a history of suicide requires monitoring and prevention. A lifetime history of significant suicidality is included within the watchful waiting/prevention level (1). This helps clinicians remain aware of the potential risk. Recent suicidal ideation or behavior requires intervention. Acute suicidal ideation and intent requires immediate action. Psychiatric hospitalization is likely a consideration for children scoring a 3 on this item. As we have shown elsewhere, this level of this item is strongly related to inpatient admissions (Leon et al., 1999).

One of the implications of this item structure is that it is unnecessary to score the CANS in order to understand and apply its results for an individual child family. Whereas scoring is recommended for purposes of outcome monitoring, service planning comes directly from level of rated needs and strengths. For needs, 2 and 3 indicate areas that require inclusion in the service plan, whereas 1 suggests a need for monitoring or preventive activities.

For strengths, 0 indicates a strength that could be the centerpiece of a strength-based plan; a 1 indicates a strength that can be utilized in strength-based planning a 2 indicates an area where a strength has been identified but must be developed. and a 3 indicates areas where no strength is currently identified that require strength identification and building efforts. The individual strengths included in the CANS-MH were derived from focus groups with system partners that were used to develop the Child and Adolescent Strengths Assessment (CASA; Lyons, Kisiel, Sokol, & Reyes, 2000). In this study, which used the CSPI and the CASA, we found that strengths and problems were significantly but independently correlated with level of functioning and the probability of high-risk behavior, suggesting that optimal child outcomes would result from both addressing needs and identifying and building strengths. This finding stimulated the integration of the two tools into the CANS.

Function

The CANS functions as an information integration tool. Whoever completes the CANS must take all the information available (e.g., observation, documentation, or both) and integrate it into his or her best estimate of the level of need or strength.

Family-friendly interview prompts have been developed for the CANS in collaboration with parents of children with serious emotional and behavioral items. However, parent interviews are not always feasible (e.g., some child-welfare applications). The sources of information may vary from child to child; therefore, the CANS method allows the rater to take the information available from all sources and integrate it into the rater's best estimate of the level of needs and strengths.

The CANS is designed to be used either as a prospective assessment tool for decision support during the process of planning services, or as a retrospective assessment tool based on the review of existing information for use in the design of service systems. This flexibility allows for a variety of innovative applications. For example, the CANS can be used to conduct retrospective file reviews for planning purposes. Then, based on the result of the planning study, CANS data can be used prospectively to direct efforts at system evolution (Lyons et al., 1997).

The flexibility of this measurement approach in either clinical operations or records review facilitates measurement audits (Anderson, Lyons, Giles, Jensen, & Estles, 2003; Lyons, Rawal, Yeh, Leon, & Tracy, 2001). By reviewing a randomly selected set of records, one can easily monitor the reliability with which the CANS is completed prospectively. This is an important method for use in situations where there is concern about the over- or underreporting of clinical needs or when there is interest in monitoring the success of ongoing training efforts.

The modular design of the CANSMH allows the tool to be adapted for local applications without jeopardizing its measurement properties. By modular, we mean that because each item can stand alone, consensus process allow partners to decide which items should or should not be included or whether a new item should be created for a local application. For example, in the Alaskan Youth Initiative application, partners in this program expressed a need to include an item called Cycling of Symptoms.

This item was designed for children with episodic symptom presentation to reflect whether the child's current symptom level reflected the best or worst of the cycle. Children whose symptoms do not cycle are rated as 0. For the New Jersey version, Fire Setting was separated out of Danger to Others; however, the logic of Danger to Others was maintained, leaving it comparable to the same item used in other venues. The CANS also can be used to monitor outcomes. This can be accomplished in one of two ways. First, the proportion of cases that move from ratings of 2 or 3 to 1 or 0 can be studied to identify the proportion of cases with resolved needs or bolstered strengths. However, to enhance the measure's sensitivity to change, it is also possible to combine items within dimensions (e.g., Problems) to obtain a dimension score. Changes in dimension scores can then be studied over time. In studies across a range of program and service settings, we have found that the dimension scores of the CANS are sensitive to change after a minimum of 3 months of service delivery. Thus, the CANS is not ideal for measuring the effect of short-term service interventions (e.g., crisis stabilization). It has worked well for monitoring outcomes of intensive community services, intensive outpatient services, treatment foster care, and residential treatment services.

Training

The CANS can only be utilized by individuals who have participated in our sequenced training model. The first phase of the training involves "remoralization." This phase emphasizes the potential importance of outcomes management to the trainee and the system. The goal of this phase of training is to help trainees realize that the CANS is a different type of assessment. The second phase of training is an explanation of the communication model. The third phase involves a detailed review of the manual with case examples for each dimension. In addition, a supplementary glossary is provided with additional helpful information, such as symptoms of major DSM-IV (1994) disorders. In the fourth stage, a practice vignette is distributed, and small groups complete the CANS for the sample case. Working through a case example with others crystallizes the use of CANS for trainees. After feedback and discussion of the group exercise, a practice vignette is completed by each individual trainee. Following feedback and discussion of this vignette, each trainee completes a test vignette to establish reliability.

Only trainees who are reliable at .70 (intraclass correlation) or more are considered trained. Remedial training is provided to individuals who are not reliable. In our experience of training more than 5,000 individual on this model, we have found that about 90% of trainees are reliable at the end of the standard training.

PROPERTIES OF THE CANS MEASURE: RELIABILITY, VALIDITY, AND SENSITIVITY TO CHANGE

Reliability

The following describes a series of reliability studies for the CANS-MH. Two primary methods are used. In one method, a standard vignette is given to a group of individuals who had completed the training protocol previously described. They then completed the CANS-MH based on the case description in a test vignette. In the second method, case records are reviewed independently by two or more reviewers, each of whom

complete the CANS-MH based on the information available in the record. In each case, reliability more than .70 using item-level intraclass correlation coefficients (mixed model) is considered adequate.

Case Vignettes. As a part of standard training in the use of the CANS-MH (as previously described) in five different states, all training participants completed the CANS-MH for a reliability test case vignette. Across these trainings, a total sample of 188 CANS-MH trainees turned in a CANS-MH. Comparisons were made between the CANS-MH ratings by the trainees and those of the first author. Average reliability using intraclass correlation, was .76. One hundred fifty-two (86.7%) of these trainees had reliabilities above .70; 67 (35.6%) had reliabilities above .80. Nineteen (10%) had reliabilities between .63 and .69. Only five trainees (2.7%) had reliabilities below .60.

Case Records. In two record review studies, 19 reviewers were trained to complete the CANS-MH on medical records. Following training, each of these reviewers completed the CANS-MH on a randomly selected record that was independently reviewed by the first author. Average reliability for all 19 reviewers was .83. Intra-class correlations were calculated for each reviewer. Seventeen of the 19 reviewers (89%) had reliabilities more than .70, with two reviewers (11%) having reliabilities of .69. Four reviewers (21%) had reliabilities more than .90. Eight reviewers (42%) had reliabilities between .80 and .89.

Anderson et al. (2003) have recently published a reliability study of the CANS-MH in which two case-record reviews of 60 cases were compared and each was compared to a prospectively completed CANS-MH used in ongoing service delivery, completed by the treating clinician. The results of this study suggest that the CANS is reliable at the individual-item level even in ongoing use in clinical service delivery.

Validity

Concurrent and Discriminate Validity: Comparison With the CAFAS. One of the most commonly used measures in the children's mental health service system is the Child and Adolescent Functional Assessment Survey (CAFAS; Hodges, McKnew, Cytryn, Stern, & Klein, 1982; Hodges & Wotring, 2000). The CAFAS provides ratings on five or eight dimensions that are combined into a single score that represents the child's overall level of functioning across major life domains.

A validity study was conducted using the CANS-MH and the CAFAS, administered with 249 youth served through the Mental Health Juvenile Justice (MH-JJ); Initiative. One of the goals of the study was to demonstrate discriminant validity, that the CANS and the CAFAS measured different facets of the same construct (child functioning) in different ways. Ideally, we hoped to find moderate correlations between CANS-MH and CAFAS scores. Both tools were completed by the MH-JJ liaison at the initiation of service planning.

The correlation between the CAFAS and CANS-MH total score was .63 ($df = 247, p < .001$). Next, we correlated the CAFAS subscales that had a parallel single CANS-MH item. The correlation between the CAFAS School/Work subscale and the CANS-MH school functioning item was .59 ($df = 247, p < .001$); between CAFAS Self-harm and CANS-MH Danger to Self item was .61 ($df = 247, p < .001$); between the CAFAS Substance Abuse and the CANS-MH Substance Use item was .73 ($df = 247, p < .001$); and between the CAFAS Thinking subscale and the CANS-MH Psychosis item the correlation was .54 ($df = 247, p < .001$). Thus, these individual

TABLE 17.3
Correlations Between the CANS and CAFAS at Enrollment Into the Illinois
Mental Health Juvenile Justice Initiative for 249 Youths

CAFAS	<i>Problems</i>	<i>Risks</i>	<i>Functioning</i>	<i>Caregiver</i>	<i>Strengths</i>	<i>Total</i>
School role	.08	.09	-.37**	.10	.28**	.31**
Home role	.18**	.24**	.21**	.04	.27**	.37**
Community role	.27**	.31**	.15**	.08	.15*	.32**
Moods/behavior	.21**	.24**	.25**	.14*	.15*	.31**
Moods/emotions	.05	.19**	.13	.16**	.26**	.24**
Self-harmful behavior	.07	.26**	.09	.11	.23**	.25**
Substance use	.17**	.10	.03	.03	.22**	.21**
Thinking	.17**	.06	.06	.06	.00	.10
Total	.29*	.33**	.23**	.23**	.42**	.63**

items of the CANS-MH which could be tested with reasonable comparisons to the CAFAS appear to be valid indicators.

Table 17.3 presents the correlations between the CAFAS subscales and the CANS dimension scales. Review of this table and the correlations previously presented suggest that there is a significant measurement overlap between the two tools but that they also appear to assess somewhat different aspects of child and family functioning.

Predictive Validity: Predication of Level of Care. In a planning study undertaken for a large northeastern state, 1,592 records were reviewed. The sampling was a stratified random sample in which a set of 20 randomly selected cases from the prior year of service were selected from randomly selected programs within each of the five regions of the state. Reviews were completed by independent reviewers with an average reliability of .86, weighted by the numbers of cases reviewed. Of these cases, 772 were selected, representing three distinct levels of care: residential treatment, intensive community-based treatment (i.e., intensive case management and a Medicaid waiver program that followed a wraparound model), and outpatient treatment.

A discriminant function analysis was used to predict group membership in the three levels of care using the five domain scores of the CANS-MH. With three groups, two discriminate functions were created, and both were statistically significant ($\chi^2 = 352.9$, $df = 10$, $p < .001$ and $\chi^2 = 43.0$, $df = 4$, $p < .001$, respectively). The CANS-MH accurately classified 63% of all cases into their actual level of care. This prediction model represents a statistically significant improvement above chance ($\chi^2 = 11.6$, $df = 12$, $p < .001$). Based on the structure matrix, the first discriminant function was a linear combination of Care Intensity (.847), Risk Behaviors (.750), Problems (.70), and Functioning (.493). The second discriminant function was a linear combination of Caregiver Capacity (.529) and Strengths (-.471).

Table 17.4 presents the predicted and actual level of care for the full sample. All five domains of the CANS provided independent and statistically significant contributions to this prediction model. The significant relationship of strengths to level of care replicates findings from a study in another state using the CSPI and Behavioral and Emotional Rating Scale (Epstein, 1998; Oswald, Cohen, Best, Jensen, & Lyons, 2001).

The data presented in Table 17.4 demonstrate that the CANS is most accurate in identifying children at the highest level of care. They also suggest that decision making on admission in these services is consistent with the original program designs, in that more challenging youth are served in more intensive and expensive programs.

TABLE 17.4
 Classification of Cases Into Three Levels of Care Using
 the Domain Scores of the CANS-MH

<i>Actual</i>	<i>Predicted</i>		
	<i>Low</i>	<i>Intermediate</i>	<i>High</i>
Low	206 (65%)	72 (23%)	37 (12%)
Intermediate	86 (30%)	138 (48%)	66 (22%)
High	14 (7%)	28 (13%)	173 (80%)

--
 Note. Row percents are included to represent the accuracy with which actual LOC is predicted.

Not surprisingly the accuracy is lowest for the intermediate level of care; it appears that a significant number of children are served at this level that are more similar to children served at either a lower or higher level of care. Thus, opportunities for step-down (e.g., the 20% of cases in residential treatment who look more like outpatient or intensive outpatient) can be identified. Likewise, the need for more intensive community treatment slots is represented by the fact that 23% of children served in the lowest level of care had needs consistent with children served in higher levels of care. Consistent with this finding (and others in the overall study), a primary result of this planning study was to increase investment in developing more intensive community services (Lyons, Carpinella, Rosenberg, Zuber, Fazio, & MacIntyre, in press). This illustrates the immediate applicability of CANS data for policy change.

Sensitivity to Change: The CANS as an Outcome-Monitoring Tool

Because of its design (a recommended 30-day window for most ratings), the CANS has limited applicability for the monitoring of change over a short period of time. It could not be expected to be useful for assessing change in an acute care setting, for example. However, it is designed to allow for the monitoring of change over a longer period of time. To allow the items of the CANS to detect reliable change for interventions to extend 3 months or longer, it is possible to move to a psychometric strategy. Because the Cronbach alphas within dimensions (e.g., Problems) are generally above .70, one can add the items within dimension to calculate a score. Calculating dimension scores enhances the statistical power by increasing the variability of the measure.

Table 17.5 presents data studying change over time with two different types of treatment. One hundred randomly selected recipients of Intensive Treatment Services (community-based wraparound model) were assessed at admission and discharge. One hundred twenty-six youth who were arrested and detained and found to have either a psychotic or affective disorder were assessed at enrollment and after 6 months of referral to community treatment and case monitoring activity. With the exception of caregiver capacity and strengths in the MH-JJ cohort, all other CANS dimensions demonstrated reliable improvement over the full samples. The effect sizes ranged from small to moderate to large depending on the dimension and the treatment group. Thus, it appears that the CANS dimension scores are sensitive to change and can be used for monitoring outcomes (i.e., change in status) in a variety of service settings.

LIMITATIONS

Although the CANS offers a number of unique advantages relative to other available measures, it does have some limitations. First, training and ongoing monitoring of

TABLE 17.5
 Comparison of Means and Standard Deviations of Comparison of Admission and Discharge
 CANS-MH Domain Scores for a Random Sample of Children Served in Intensive
 Community Treatment Programs or Residential Treatment Centers

<i>CANS Domain</i>	<i>Intensive Community</i>			<i>Mental Health Juvenile Justice</i>		
	<i>Admit</i>	<i>Discharge</i>	<i>Effect Size</i>	<i>Admission</i>	<i>Discharge</i>	<i>Effect Size</i>
Problems	8.0 (2.5)	5.8 (3.5)	0.79	8.6 (2.5)	6.9 (3.1)	0.59
Risk behaviors	3.4 (2.5)	2.5 (2.6)	0.33	7.0 (2.9)	5.8 (3.0)	0.39
Functioning	5.3 (2.5)	4.8 (1.9)	0.26	2.9 (2.4)	0.9 (1.7)	0.82
Caregiver	5.1 (4.6)	3.8 (4.7)	0.35	5.9 (3.9)	6.5 (4.4)	0.15
Strengths	18.6 (4.3)	17.4 (4.4)	0.42	16.5 (4.3)	14.4 (4.9)	0.49

*p<,01. **p<,001.

reliability is important. It is likely not possible for most individuals to become reliable in the use of the CANS by simply reading and studying the manual. With training, it is clearly possible to use the CANS reliably in the field. Second, because of the item construction, the CANS is likely less sensitive to change, particularly over short periods of time, than other measures. However, particularly when dimension scores are used, the CANS appears to capture meaningful change across a number of program types. Finally, given its modular design, different locales can use different versions with different items. The granting of local control over measurement design can complicate cross-system comparisons. Although it is possible to manage these comparisons given the use of identical items, care must be taken to ensure appropriate cross-site comparisons are made (Rawal, Lyons, MacIntyre, & Hunter, in press).

CONCLUSION

By focusing on the communication role of measurement in outcomes management, the CANS-MH illustrates a novel approach that integrates psychometric and clinimetric approaches to measurement within the children's service delivery system. Although classical test theory and item response theory continue to play an important role in measurement development, additional considerations beyond the psychometric properties of a tool should influence the development, design, and use of outcomes measures.

Evidence from reliability studies indicates that the CANS can be completed reliably by individuals working with children and families. The CANS demonstrates both concurrent validity and predictive validity in initial studies. In addition, effect sizes from two outcomes studies support the use of the CANS for monitoring outcomes. Thus, it appears that the CANS is a reliable and valid tool for use within children's services.

Various versions of the CANS are used currently in at least twenty-one different states. The applications include decision support, quality improvement, and outcomes monitoring. Decision support applications include both establishing appropriate level of care and assisting service planning. Quality improvement activities involve monitoring the match between identified needs and services provided, and the extent to which strengths are identified and included in service plans. Outcomes monitoring applications involve monitoring the effectiveness of services over 3- to 6-month intervals.

It has become increasingly apparent that in order to effectively manage and evolve systems of care for children and families, it is necessary to ensure that service recipients are represented in all aspects of the process. One means of ensuring representation is to consistently collect information on the needs and strengths of children and families who are served using a measure that takes all perspectives into account. In this way, services can be understood from the perspective of their overarching purpose—to help children and families. The CANS is one tool that can facilitate this process. The CANS allows for the reliable and valid communication of needs and strengths to inform decision making at the individual child and family level while enabling administrators to monitor the quality and effectiveness of services at the program and system level.

ACKNOWLEDGEMENTS

A large number of individuals have participated in the development of the CANS tools over the past 5 years, and it is impossible to name everyone. However, the following individuals have made particularly notable contributions to the development of the CANS: M. Connie Almeida, Rachel L. Anderson, Mina Dulcan, M.D., Ardas Khalsa, Cassandra Kisiel, John Lavigne, Michael Lee, Mary Beth Rautkis, Harry Shallcross, and Patricia Sokol. The CANS tools are open-domain tools that are free for anyone to use. They can be obtained from the authors or by visiting www.buddinpraed.org.

REFERENCES

- American Psychological Association. (1994). *Diagnostic and statistical manual of mental disorders*. Washington DC: Author.
- Anastasi, A. (1968). *Psychological Testing*, 3rd ed. Toronto, Ontario, Canada: Macmillan.
- Anderson, R L., Lyons, J. S., Giles, D. M., Price, J. A, Estes, G. (2003). Examining the reliability of the Child and Adolescent Needs and Strengths-Mental Health (CANSMH) Scale from two perspectives: A comparison of clinician and researcher ratings. *Journal of Child and Family Studies*, 12, 279-289.
- Apgar, V. (1966). The newborn (Apgar) scoring system. Reflections and advice. *Pediatric Clinics of North America*, 13, 645-650.
- Beardslee, W. R, Wright, E., Rothberg, P. C., Salt, P., BVersage, E. (1996). Response of Families to two preventive intervention strategies: Long-term differences in behavior and attitude change *JAACAP*, 35, 774-782
- Bickman, L. (1996). Implications of a children's mental health managed care demonstration evaluation. *Journal of Mental Health Administration*, 23, 107-117.
- Bickman, L., Lambert, E. W., Andrade, A R, & Penaloza, R V. (2000). The Fort Bragg continuum of care for children and adolescents: Mental health outcomes over 5 years. *Journal of Consulting and Clinical Psychology*, 68, 710-716.
- Bickman, L., Noser, K, & Summerfelt, W. T. (1999). Long-term effects of a system of care on children and adolescents. *Journal of Behavioral Health Services & Research*, 26, 185-202.
- Bloem, R B., Beckley, D. J., van Hilten, B. J., & Roos, R A. (1998). Clinimetrics of postural instability in Parkinson's disease. *Journal of Neurology*, 245, 669-673.
- Clark, G. B., Schyve, P. M., Lepoff, R B., & Ruess D. T. (1994). Will quality management paradigms of the 1990's survive into the next century? *Clinical Laboratory Management Review*, 8, 426-428, 430-434.
- Drake, R (2002, November). *Psychiatric rehabilitation and evidence-based medicine*. Carl Taube Award Lecture to the American Public Health Association, Philadelphia.
- Drasgow, E, & Schmitt, N. (Eds.). (2002). *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis*. San Francisco: Jossey Bass.
- Endicott, J., Spitzer, R L., Fleiss, J. L., & Cohen, J., (1976). The global assessment scale. A procedure measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry*, 33, 766-771
- Epstein, M. H. (1998). Strength-based assessment: the Behavioral and Emotional Rating Scale. *Behavioral Healthcare Tomorrow*, 7, 46-48.

- Epstein, M. H. & McGee, J. L., (1996). Roles of federal and state government in outcomes assessment. *American Journal of Medical Quality*, 11, 19-21.
- Feinstein, A. R. (1999). Multi-item "instruments" vs. Virginia Apgar's principles of clinimetrics. *Archives of Internal Medicine*, 159, 125-128.
- Fischer, E. P., Shumway, M., & Owen, R. R. (2002). Priorities of consumers, providers, and family members in the treatment of schizophrenia. *Psychiatric Services*, 53, 724-729.
- Friedland, L. A. (2001) Communication, community, and democracy: Toward a theory of communicatively integrated community. *Communication Research*, 28, 358-391.
- Gager, P. J., & Elilas, M. J. (1977). Implementing prevention programs in high-risk environments: Application of the resiliency paradigm. *American Journal of Orthopsychiatry*, 67, 363-373, 1977.
- Gates, G. A. (2000). Clinimetrics of Meniere's disease. *Laryngoscope*, 110, 8-11.
- Habermas, J., & Seidman, S. (1989). Juergen Habermas on society and politics: A reader. Frankfurt, Germany: University of Frankfurt.
- Hammer, K. M., Lambert, E. W., & Bickman, L. (1977). Children's mental health in a continuum of care: Clinical outcomes at 18 months for the Fort Bragg demonstration. *Journal of Mental Health Administration*, 24, 465-471.
- Harris, T. E. (200). *Applied organizational communication: Principles and pragmatics for future practice*. Mahwah, N J: Lawrence Erlbaum Associates.
- Hoagwood, K., Burns, B. J., Kiser, L., Ringeisen, H., & Schoenwalk, S. K. (2001). Evidence-based practice in child and adolescent mental health services. *Psychiatric Services*, 52, 1179-1189.
- Hodges, K., McKnew, D., Cytryn, L., Stern, L., & Klein, J. (1982). The Child Assessment Schedule (CAS) diagnostic interview: A report on reliability and validity. *Journal of the American Academy of Child and Adolescent Psychiatry*, 21, 468-473.
- Hodges, K., & Wotring, J. (2000). Client typology based on functioning across domains using the CAFAS: Implications for service planning. *Journal of Behavioral Health Services and Research*, 27, 257-270.
- Hoff, J. I., van Hilten, B. J., & Roos, R. A. (1999). A review of the assessment of dyskinesias. *Movement Disorders*, 14, 737-743.
- Howard, K. I., Moras, K., Brill, P. L. Martinovich, A., & Lutz, W. (1996). Evaluation of psychotherapy. Efficacy, effectiveness, and patient progress. *American Psychologist*, 51, 1059-1064.
- Kihlstrom, A., & Israel, J. (2002). Communicative or strategic action –An examination of fundamental issues in the theory of communicative action. *International Journal of Social Welfare*, 11. 210-218.
- Koike, A. K., Unutzer, J., & Well, K. B. (2002). Improving the care for depression in patients with comorbid medical illness. *American Journal of Psychiatry*, 158, 1738-1745.
- Krulich, L. H. (2002). A basic and practical overview of the six steps of outcome-based quality improvement: Part 2. *Home Healthcare Nurse*, 20, 585-586.
- Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001). Patient-focused research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology*, 69, 159-172.
- Lambert, M. J., Ogles, B. M., & Masters, K. S. (2000). Choosing outcome assessment devices: An organizational and conceptual scheme. *Journal of Counseling and Development*, 70, 527-532.
- Leon, S. C., Uziel-Miller, N. D., Lyons, J. S., & Tracy, P. (1999). Psychiatric hospital utilization of children and adolescents in state custody. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38, 305-310.
- Lichtman, D. M., & Appleman, K. A. (1995). Measures of effectiveness: A methodology of integrating planning, measurement, and continuous improvement. *Military Medicine*, 160, 189-193.
- Lueger, R. J., Howard, K. I., Martinovich, Z., Lutz, W., Anderson, E. E., & Grissom, G. (2001). Assessing Treatment progress of individual patients using expected treatment response models. *Journal of Consulting & Clinical Psychology*, 69, 150-158.
- Lyons, J. S. (1999). *The Child and Adolescent Needs and Strengths for children with mental health challengers*. Winnetka, IL: Buddin Praed Foundation. Also retrieved from www. Buddinpraed.org
- Lyons, J. S., Howard, K. I., O'Mahoney, M. T., & Lish, J. (1997). *The measurement and management of clinical outcomes in mental health*. New York: Wiley.
- Lyons, J. S., Kisiel, C. L., Dulcan, M., Cohen, R., & Chesler, P. (1997). Crisis assessment and psychiatric hospitalization of children and adolescents in state custody. *Journal of Child and Family Studies*, 6, 311-320.

- Lyons, J. S., Mintzer, L. L., Kisiel, C. L., & Shallcross, H. (1998). Understanding the mental health needs of children and adolescents in residential treatment. *Professional Psychology: Research and Practice*, 29, 582-587.
- Lyons, J. S., Rawal, P., Yeh, I., Leon, S., & Tracy, P. (2001). Use of measurement audit in outcomes management. *Journal of Behavioral Health Services & Research*, 29, 75-80.
- Marx, R. G., Bombardier, C., Hogg-Johneon, S., & Wright, J. G. (2000). Clinimetric and psychometric strategies for development of a health measurement scale. *Journal of Clinical Epidemiology*, 52,105-111.
- Nunnally, J. (1976). *Psychometric theory*. New York: Wiley.
- Ogles, B. M., Lambert, M. J., & Masters, K. S. (1996). *Asscssing outcome in clinical practice*. Needham Heights MA: Allyn L Bacon.
- Orlikoff, J. E. (1994). CQI trends and transitions. *Healthcare Erecutme*, 9, 12-15.
- Oswald, D. P., Cohen, R, Best, A. M., Jensen, C. E., & Lyons, J. S. (2001). Child strengths and level of care for children with emotional and behavioral disorders. *Jounrnal of Emotional and Behanioral Disorders*, 192-199.
- Rawal, P., Lyons, J. S., MacIntyre, J., & Hunter, J. C. (in press). Regional variations and clinical indicators of antipsychotic use in residential treatment. A four-state comparison . *Journal of Behavioral Health Services and Research*.
- Rosenblatt, A. (1993). In home, in school, and out of trouble. *Journal of Child and Family Studies*, 2, 275-282.
- Rost, J., & Langeheine, R (1997). *Applications of latent train and latent class models in the social sciences*. New York: Waxmann.
- Schiff, G. D., & Goldfield, N. I. (1994). Deming meets Braverman: toward a progressive analysis of the continuous quality improvement paradigm. *International Journal of Health Services*, 24, 655-673.
- Sederer, L I, Dickey, B., & Hennann, R C. (1996). The imperative of outcomes assessments in psychiatry. In L. I. Sederer & B. Dickey (Eds.), *Outcomes Assessment in Clinical Practice*. Baltimore: Williams & Wilkins.
- Sheasby, J. E., Barlow, J. K, Cullen, L. A, & Wright, C. C. (2000) Psychometric properties of the Rosenberg self-esteem scale among people with arthritis. *Psychological Reports*, 86,2139-1146.
- Steinwachs, D. M., Flynn, L. M., Norquist, G., & Skinner, E. A. (Eds). (1996). *Using client outcomes information to mprove mental health and substance abuse treatment*. San Fransisco: Jossey-Bass.
- Stone, M., Wonen, D., Lax, M., Payne, U., Lapp, V., & Inman, R (2001). Clinical and imaging correlates of response to treatment with infliximab in patients with ankylosing spondylitis. *Journal of Rheumtology*, 28,1605-1614.
- Stroul, B. A. (1993). *Systems of care for children and adolescents with severe emotional disturbances: What are the results?* Washington, DC: CASSP Technical Assistance Center, Center for Child Health and Mental Health Policy Center, Georgetown University Child Development Center.
- Weiss, B., Cahon, T., Harris, V., Phung, T. M. (1999). The effectiveness of traditional child psychotherapy. *Journal of Consulting & Clinical Psychology*, 67, 82-94.
- White, W. J. (2001). A communication model of conceptual innovation in science. *Communication Theory*, 11, 290-314.
- Zyzanski, S. J., & Perloff, E. (1999). Clinimetrics and psychometrics work hand in hand. *Archives of Internal Medicine*, 159, 1816-1817.