

Family Assessment in Child Welfare Services: Instrument Comparisons

Michelle A. Johnson, M.S.W., BASSC Research Director
Susan Stone, Ph.D., Assistant Professor, School of Social Welfare, UC Berkeley
Christine Lou, M.S.W., BASSC Doctoral Research Assistant
Cathy Vu, M.P.A., BASSC Doctoral Research Assistant
Jennifer Ling, BASSC Master's Research Assistant
Paola Mizrahi, B.A., Research Assistant
Michael J. Austin, BASSC Staff Director

June 2006

Supported by the Bay Area Social Services Consortium and
the Zellerbach Family Foundation

The Center for Social Services Research (CSSR) in the School of Social Welfare at the University of California at Berkeley conducts research, policy analysis, program planning, and evaluation toward the improvement of the publicly supported social services. The focus of the Center is on populations who are considered needy or disadvantaged, including victims of child abuse and neglect, the chronically mentally ill, the aged, the medically indigent, and the poor.

Housed at CSSR, the Research Response Team of the Bay Area Social Services Consortium (BASSC) was organized in 1995 to respond rapidly to the emerging needs of county social service agencies for information for their changing environments. Structured reviews of the research literature are undertaken in close collaboration with agency administrators and program staff.

BASSC was founded in 1987 and is composed of the Directors of Bay Area county social service and human service agencies, deans of the Bay Area graduate social work departments, and foundation representatives. The BASSC Research Response Team would like to thank the members of the BASSC Research Advisory Committee: Margaret Ahern, Sonoma County Human Services Department; Maureen Borland, Former Director, San Mateo County Human Services Agency ; Kilolo Brodie, California State University, East Bay; Brenda Carrillo, San Jose State University; Christopher Cassels, Solano County Health & Welfare; Danna Fabella, Contra Costa County Employment and Human Services; Shaaron Gilson, University of California, Berkeley; Alexis Halley, San Mateo County Human Services Agency; Will Johnson, Alameda County Social Services; Dan Kelly, City & County of San Francisco; Liz Knox, San Francisco State University; Felicia Law-Murray, San Francisco State University; Rodger Lum, San Francisco State University; Heather Ravani, Marin County Health & Welfare; Cynthia Stoops, Santa Clara County Social Services Agency; Robert Taniguchi, Monterey County Department of Social Services; Jennifer Yasumoto, Napa County Health and Human Services; Judy Yokel, Santa Cruz County Human Resources Agency. The involvement of these Research Advisory Committee members in shaping the study and providing feedback on study results and final reports was invaluable. A special thanks to Katherine Strahorn, Center for Social Services Research, for her assistance.

Introduction

For child welfare services to be relevant and effective, workers must systematically gather information and continuously evaluate the needs of children and their caregivers as well as the ability of family members to use their strengths to address their problems. Several kinds of assessments are conducted with children and families that come to the attention of child welfare services, such as risk and safety assessments that are used to guide and structure initial decision-making and predict future harm. However, the states' performance on the federal Child and Family Services Reviews, in both outcomes and systemic factors, suggests that it is not often clear how caseworkers gain a full understanding of family strengths, needs, and resources or how this information is incorporated into ongoing service planning and decision-making (HHS, 2006). Family assessment instruments hold promise for enhancing clinical judgment by structuring the decision making process and demonstrating the linkages between assessment, service provision, and child and family outcomes.

A previous structured literature review, *Risk and Safety Assessment in Child Welfare: Instrument Comparisons* (2005),¹ described approaches to assessing risk and summarized research findings regarding the validity and reliability of existing instruments. The primary focus of this review is to evaluate the family assessment literature and provide recommendations for promising instruments that may be useful in structuring the family assessment process. The report is divided into six sections. We first describe the concept of family assessment in the child welfare context, followed by an overview of the theoretical and disciplinary influences in the family assessment field and key measurement criteria. Next, we present practical considerations in the selection of a family assessment instrument for use in child welfare. The framework and methods of the review are then presented, followed by major findings and implications for practice.

Family Assessment in Child Welfare

Comprehensive family assessment has been defined as the process of identifying, gathering and weighing information to understand the significant factors affecting a child's safety, permanency, and well-being, parental protective capacities, and the family's ability to assure the safety of their children. The Children's Bureau of the U.S. Department of Health and Human Services recently released guidelines for comprehensive family assessment to provide an initial framework to facilitate the development of best practices (HHS, 2006). The guidelines

identify key points in the life of a case for comprehensive family assessment, beginning with the initial contact with the family and continuing through several decision making stages, including placement, reunification, termination of parental rights, and case closure. Other assessment points include decisions related to changes in the service plan or case goal, independent living decisions, formal progress reviews, and anytime new information triggers the need for additional assessment. However, existing guidelines for family assessment in child welfare services typically do not recommend particular tools or instruments for monitoring the complex and often challenging circumstances that bring families to the attention of child welfare services (HHS, 2006; DePanfilis & Salus, 2003).

Previous literature on family assessment instruments for use in child welfare includes descriptions of instruments (Pecora, Fraser, Nelson, McCroskey, & Meezan, 1995; Berry, Cash, & Mathiesen, 2003) and guides for developing comprehensive assessment strategies as part of community-based child welfare services reform (Day, Robison, & Sheikh, 1998). This structured literature review builds on these efforts by identifying the most valid and reliable instruments that address the following four federally-defined domains of family assessment: (1) patterns of social interaction, including the nature of contact and involvement with others, and the presence or absence of social support networks and relationships; (2) parenting practices, including methods of discipline, patterns of supervision, understanding of child development and/or of the emotional needs of children; (3) background and history of the parents or caregivers, including the history of abuse and neglect; and (4) problems in access to basic necessities such as income, employment, adequate housing, child care, transportation, and other needed services and supports (HHS, 2006). Several additional behaviors and conditions have been associated with child maltreatment, such as domestic violence, mental illness, poor physical health, disabilities, and alcohol and drug use. Ideally, a comprehensive family assessment instrument will address these conditions and indicate whether a need for more specialized assessment exists. An objective of this review was to identify measures that addressed these behaviors and conditions as part of a comprehensive family assessment strategy. However, the review of specialized instruments for these conditions and various disabilities was outside the scope of this review. A structured review on the assessment of children and youth in the child welfare system is the focus of a separate review.

Family Assessment and Measurement Criteria

Interest in family relationships began expanding in research and clinical practice with the advent of systems of child protection in the 1970s; however, only in recent years have significant efforts been made to develop family assessment instruments specifically for the child welfare practice setting. Three related sets of literatures, stemming from academic psychology during the 1970s and 1980s and medicine during the 1980s and 1990s, inform the general topic of family assessment (Boss, Doherty, LaRossa, Schumm, & Steinmetz, 1993). Rooted in family systems theory and family therapy research, a first literature seeks to capture overall family functioning, focusing on the family as a primary unit of analysis. Typically, three general within-family dimensions are assessed including overall structural and organizational patterns, communication processes, and affective qualities and cohesiveness. For example, the McMaster Model (Epstein, Bishop, & Levin, 1978), the Circumplex Model (Olson, 2000), and the Beavers Systems Model (Beavers & Hampson, 2000) represent assessment models in this tradition.

Informed by developmental psychology, a second literature includes research on the assessment of parenting. This literature identifies relevant components of parenting and typically relates them to child developmental and functional outcomes. In short, it focuses on the caregiver-child dyad as the key unit of attention. Conceptual and empirical work in this area highlights the following five parenting factors that are particularly salient for assessment: (1) parent beliefs about the child, (2) perceived efficacy in the parenting role, (3) parenting style, (4) parent-child relational qualities, and (5) parenting skills and behaviors. Finally, the stress and coping literature, as well as related literatures on risk and resilience, informs family assessment (see Hill, 1949). For example, McCubbin and McCubbin (1987) provide a model of family stressors (normative or unexpected; acute or chronic) and the extent to which families manage the stressor without negative effects on the family system. Research identifies two protective factors, including the internal and external social support resources of families as well as how the family perceives the stressor (i.e., the extent to which the family views the stressor as manageable). In short, this work places attention on social supports and family appraisal processes as a way to understand and assess family functioning.

These major theoretical and disciplinary influences have given rise to several practical issues when considering the appropriateness of a family assessment measure and method. While there are many approaches, family assessment methods typically fall into three categories: client

self-report, observation, and interviews. Each of these methods has its advantages and disadvantages. A key distinction is the degree to which the method is formalized. Formal methods, such as self-report questionnaires, tend to have procedures that are clearly outlined to facilitate consistently repeated administrations. By contrast, informal methods such as interviews may be less clear in their specification and more variable in terms of administration.

Family assessment measures also vary in terms of the perspective obtained. Typically, child welfare practitioners will consider the perspectives of multiple individuals during the family assessment process, including “insider” reports from family members and children as well as “outsider” reports from school personnel, extended family members, and others that may be involved with the case. Integration of the assessment of multiple reporters with insider and outsider perspectives is reflected in the “multisystem-multimethod” (MS-MM) approach (Cromwell & Peterson, 1983).

Self-report questionnaires provide a unique insider view of family life as well as reliable methods, simplified administration and scoring, and a measurable link between an individual’s perceptions or attitudes and behaviors. Given these advantages, they are by far the most commonly used method in research as well as in practice. Observation rating scales provide another cost-effective method of generating outsider information regarding family interaction patterns that can also be evaluated for reliability and validity. However, rating scales can also be limited in their usefulness by the competence of the rater and the psychometric quality of the scale. Raters must have a clear understanding of the concepts that are measured and the behaviors that represent the concepts in practice. They must also possess adequate knowledge of different populations in order to place observed behavior on a continuum, a concern that adequate training and clinical supervision can begin to address. However, as with self-report measures, evidence of the validity and reliability of an observational rating scale is critical in the instrument selection process, particularly with regard to specific stages of assessment.

Family assessment includes several sequential functions, including (1) screening and general disposition, which typically occur at intake; (2) definition of the problem, which may include diagnostic assessments (or quantification of problem severity) that occur during intake and investigation procedures; (3) planning, selecting, and matching services with identified problems; and (4) monitoring progress and evaluating service outcomes (Hawkins, 1979). Validity and reliability are the primary psychometric issues when selecting family assessment

measures. Briefly, validity is the degree to which the instrument measures what it intends to measure (e.g., family functioning or perceptions of family life) whereas reliability is a measure of consistency. In other words, a high level of reliability indicates confidence in the fact that similar results will be obtained if similar procedures are used and if the results are assessed in the same manner time after time. As Figure 1 suggests, there are many types of validity and reliability to consider for each stage of assessment when selecting a family assessment instrument. Appendix A provides more details about these measurement criteria.

Figure 1. Stages of Assessment, Criteria for Evaluation, and Child Welfare Services (CWS) Decision Making (adapted from Carlson, 1989)

CWS Decision-Making Stage	Assessment Stage	Clinical Criteria	Measurement Criteria
Intake	Screening	Detects the nature of a problem; provides guidance as to further assessment; cost-effective	Adequacy determined by predictive validity
Investigation	Diagnosis	Confirms hypotheses regarding family functioning; quantifies or measures the severity of dysfunction; determines the primary locus of the problem; provides standardized measures and validated clinical cutoff scores	Adequacy determined by discriminative and differential predictive validity
Case Planning	Service Planning	Specifies objectives for change; analyzes factors that produce and maintain problematic behavior; identifies family strengths and resources; determines both intervention sequence and level of change adequate for treatment termination; may require multimethod assessment approach if multiple goals cannot be systematically measured using a single method	Adequacy determined by content validity and inter-rater reliability regarding specific behavioral patterns relevant to the problem
Continuing Services/Placement and Reunification Decisions	Monitoring Progress/Evaluation	Focuses on the behavior to be changed; amenable to repeated-measures; generalizable beyond the treatment setting; sensitive to change; easily administered	Unresponsive to spurious influences such as retesting effects and instrument decay

Considerations for Selecting Family Assessment Instruments for Use in Child Welfare

There are many clinical and measurement criteria for evaluating the adequacy of a family assessment method and they vary depending upon the function for which they are developed and

used. In the child welfare setting, the choice of method will also be governed by the following practical considerations (adapted from Johnson & Wells, 2000):

1. Will the instrument be used for initial assessment only or for the monitoring of progress?

If it is the latter, is the instrument sensitive to clinical change? Many instruments are designed to detect the existence of a given condition, not to measure improvement in a child or family's functioning over time. Only instruments sensitive enough to detect client change can reliably measure it, a distinction that may not be apparent to many users. Since child welfare decisions are often made when there appears to be a "lack of progress" on the part of a client, assessment instruments need to be very sensitive to measuring change.

2. What domains of family assessment are assessed? Family assessment instruments cover a wide array of factors, from tangible outcomes such as the cleanliness of the home environment, to less tangible factors such as self-esteem. Before selecting measures, such as parental functioning, parental behavioral health, or quality of the home environment, it is important for agencies and programs to clearly identify the goals and desired outcomes of services for children and families.

3. How long does it take to administer the instrument? Child welfare workers generally have limited time to spend with clients. Therefore, the time needed to administer an assessment instrument needs to be brief. Managers will also want to consider the time it takes to train workers to use the instrument and the length of time required to interpret the results.

4. What is the developmental stage or age focus with respect to the instrument? The broad range of ages of parents and children served by the child welfare system will require agencies to select multiple instruments in most cases.

5. Is it useful with the intended target group of clients? For example, if an agency works primarily with Latino clients, knowing that a particular instrument has been tested with Latino individuals will be a defining factor in selection. As most instruments have been normed with white English speaking individuals in research settings, serious consideration needs to be given to the appropriateness of using instruments in practice that are not culturally validated. Managers will also need to consider how the instrument is administered. If a client completes the form, it is important to consider the reading level of the instrument and the languages available.

6. What are the advantages and disadvantages of using this instrument? Certain clinical instruments have the advantage of assessing a range of child or family functioning. Other instruments are useful in that they can be used along with other tools as part of a "package." Any time an instrument can provide information on multiple outcomes, managers are able to conserve resources. Several instruments may only tap one aspect of family functioning, or are useful only with a particular population. For example, some instruments may be written for a higher reading level than would be sensible for use with an agency's client population. Managers and administrators also need to consider the costs of purchasing copyrighted materials or reproducing other instruments.

7. What does the instrument tell a practitioner, administrator, or policy maker? Decisions about instruments should be guided by a clear idea of what information is needed, how it will be used, and who will be using it.

8. Is psychometric data available? Again, reliability and validity indices establish the credibility of instruments. Without this information, various alternative explanations for the findings (e.g., examiner bias, chance, and effects of maturation) cannot be ruled out, which seriously restricts the usefulness of findings.

Given measurement criteria and practical considerations, the goal of this review is to identify instruments that (a) comprehensively address the major domains of family assessment, (b) are valid and reliable for the appropriate stage of assessment, and (c) are practical for use in child welfare settings.

Methods

Search Strategy

This review used pre-determined search terms and search sources to identify research literature within a given topic. This method of searching can reduce the potential for bias in the selection of materials. Using specified search terms, we searched numerous social science and academic databases available through the University of California library. In addition, we conducted overall internet searches and also searched the websites of research institutes and organizations specializing in systematic reviews, conference proceedings databases, dissertation databases, internet databases. In order to gather information on research that has not been published, inquiries were sent to national child welfare resource centers, federal agencies such as the Children's Bureau, and child welfare researchers (see Appendix B for a description of the search strategy). The references in reviews and primary studies were scanned to identify additional articles. The references reviewed were limited to those printed in the English language.

Evaluation Methods

The instruments that were obtained through the structured search strategy were evaluated with regard to their appropriateness for child welfare settings based on seven criteria: (1) their relationship to the family assessment domains identified and their comprehensiveness in relation to these domains; (2) the appropriateness of the assessment methods employed; (3) the number of stages of assessment addressed, with emphasis on the appropriateness for use at multiple points in the life of the case; (4) the populations with which instruments were normed; (5) ease

of administration, in terms of time, instructions, scoring, and clarity of interpretation; (6) other advantages and disadvantages related to use in the child welfare setting, such as the reading level required of clients or prior use by caseworkers; and (7) psychometric properties. The psychometric properties of the instruments were rated on a four point scale, from those having the least psychometric information available to those having psychometric information available for all of the stages of assessment that the instrument addressed.

A ten percent sample of the instrument evaluations was reviewed by an independent reviewer to establish the inter-rater reliability of the evaluation process. Two discrepancies were found with regard to the comprehensiveness of the family assessment domains that an instrument addressed and in one case, with regard to the stage of assessment that the instrument addressed. These differences were reconciled with the introduction of additional sub-criteria for evaluation.

Major Findings

Overview

Eighty five (n=85) instruments pertaining to family assessment were evaluated (see Appendix C). Of these, the majority typically addressed one to two domains of family assessment, such as patterns of social interaction and parenting practices. The majority of the instruments relied on self-report methods and/or observational rating scales (80%). A smaller number of instruments included structured interviews (15%) and methods relying on structured tasks such as games (4%).

Figure 2. Instruments/Models Addressing Family Assessment Domains (n=85)²

Family Assessment Domains	Number of Instruments Addressing Domain
Patterns of Social Interaction (the nature of contact and involvement with others, the presence or absence of social support networks and relationships)	58
Parenting Practices (methods of discipline, patterns of supervision, understanding of child development and/or of emotional needs of children)	43
Background and History of Caregivers (including the history of abuse and neglect)	20
Problems in Access to Basic Necessities (such as income, employment, adequate housing, child care, transportation and needed services and supports)	23
Other Behaviors and Conditions (domestic violence, mental illness, physical health, physical, intellectual, and cognitive disabilities, alcohol and drug use)	18

In terms of measurement criteria, half of the instruments (50%) had some type of information available about their reliability and/or validity. In twenty-two cases (26%), psychometric information was available for (1) some stages of assessment but not all, or (2) for specific stages but overarching psychometric properties of the instrument had yet to be established (such as content validity or test-retest reliability). Ten instruments had information available for all stages of the assessment addressed (12%), while another ten provided little to no psychometric information (12%).

As mentioned, seven criteria were used to evaluate the 85 instruments with regard to their appropriateness for child welfare settings. Seven instruments appeared to be the most comprehensive and appropriate for use in the child welfare setting. These are presented first, followed by instruments that appear to be promising for specialized purposes within specific domains. For example, the specialized assessments of patterns of social interaction presented (n=4) might be made to better target referrals for mental health services or family therapy. Similarly, the assessments of parenting practices identified (n=5) might be made to refer clients to the most appropriate parenting program. Community-based providers of mental health services and parenting programs might also use these specialized instruments to assess family strengths and needs, develop service plans, and monitor and report on progress. Promising instruments for the specialized assessment of background characteristics (n=3) and basic needs (n=2) are also discussed.

Comprehensive Measures of Family Assessment

As noted in Figure 3, the seven family assessment instruments that are the most comprehensive and appear the most promising for child welfare practice include three instruments that have been developed specifically for use in child welfare settings: (1) the North Carolina Family Assessment Scale (NCFAS) and two modified versions of the NCFAS, (2) the NCFAS for Reunification (NCFAS-R) and (3) the Strengths and Stressors Tracking Device (SSTD). Four additional instruments include (4) the Family Assessment Form (FAF), (5) the Family Assessment Checklist (FAC), (6) the Ackerman-Schoendorf Scales for Parent Evaluation of Custody (ASPECT), and (7) the Darlington Family Assessment System (DFAS). Each instrument is discussed briefly.

North Carolina Family Assessment Scale (NCFAS) and Related Instruments. The NCFAS (Reed-Ashcraft, Kirk, & Fraser, 2001) was developed in the mid-1990s to allow caseworkers working in intensive family preservation services (IFPS) to assess family functioning at the time of intake and again at case closure. The 39-item instrument was designed to assist caseworkers in case planning, monitoring of progress, and measuring outcomes. The NCFAS provides ratings of family functioning on a six-point scale ranging from “clear strengths” to “serious problems” in the following five domains: (1) environment, (2) parental capabilities, (3) family interactions, (4) family safety, and (5) child well-being. Internal consistency and construct validity have been established for early versions as well as the most recent version of the NCFAS (Version 2.0; Reed-Ashcraft et al., 2001, Kirk et al., in press) and the instrument is able to detect changes in functioning over time. The instrument also appears to have some degree of predictive validity in relation to placement prevention; however, the authors caution that the relatively weak capability of the intake ratings to predict placement at closure or thereafter suggest that the NCFAS should not be used as a device to screen out families from service at the time of intake (Kirk et al., in press). Additional research with sufficiently large samples is necessary to establish predictive validity for outcomes of interest.

Figure 3. Promising Measures of Comprehensive Family Assessment for Child Welfare by Assessment Domain

Instruments	Family Assessment Domains				
	Patterns of Social Interaction	Parenting Practices	Background of Caregivers	Basic Needs	Other Behaviors and Conditions
North Carolina Family Assessment Scale (NCFAS)	X	X	X	X	X
NCFAS-Reunification (NCFAS-R)	X	X		X	X
Strengths and Stressors Tracking Device (SSTD)	X	X	X	X	X
Family Assessment Form (FAF)	X	X		X	
Family Assessment Checklist (FAC)		X	X	X	X
Ackerman-Schoendorf Scales for Parent Evaluation of Custody (ASPECT)	X	X	X	X	X
Darlington Family Assessment System (DFAS)	X	X	X		X

X = assesses family functioning in this domain

The NCFAS for Reunification (NCFAS-R), a collaborative effort between the National Family Preservation Network and the University of North Carolina at Chapel Hill, is an assessment instrument used to assist caseworkers using intensive family preservation service strategies to successfully reunify families where children have been removed from the home due to substantiated abuse and or neglect, juvenile delinquency, or the receipt of mental health services in a “closed” treatment setting (Reed-Ashcraft et al., 2001). The scale provides family functioning assessment ratings on seven domains relevant to reunification: (1) environment, (2) parental capabilities, (3) family interactions, (4) family safety, (5) child well-being, (6) caregiver/child ambivalence, and (7) readiness for reunification. Like the NCFAS, change scores for the NCFAS-R illustrate the amount of measurable change that is achieved during the service period from intake ratings through closure ratings. Internal consistency and concurrent validity in relation to the success or failure of reunification cases have been established for this measure.

The Strengths and Stressors Tracking Device (SSTD) is another modification of the NCFAS that assesses the strengths and needs of families at intake to help guide case planning and evaluate the effectiveness of treatment. The SSTD is shorter than the NCFAS in its two-page form but includes an additional 16 items (for a total of 55 items). The SSTD can be completed by caseworkers in less than 30 minutes. Like the NCFAS, family strengths and stressors are rated on a Likert type scale, with -2 indicating a serious stressor to +2 indicating a clear strength. Unlike the NCFAS, psychometric information for the SSTD is somewhat limited. In a small validation study in a single agency, the SSTD demonstrated high internal consistency in all domains, distinguished between physical abuse and neglect cases at intake, and appeared to be sensitive to specific changes made by families during the treatment period (Berry, Cash, & Mathiesen, 2003). However, further use and validation are needed to establish content and criterion related validity, including predictive validity, as well as test-retest reliability.

Family Assessment Form (FAF). The FAF is a practice-based instrument that was developed by workers at the Children’s Bureau of Los Angeles, a nonprofit child welfare agency, to help practitioners improve the assessment of families receiving home-based services. It includes 102 items that relate to the following five factors: (1) living conditions, (2) financial conditions, (3) interactions between adult caregivers and between caregivers and children, (4) support available to the family, and (5) developmental stimulation available to children. The

FAF is completed at assessment and termination along with a two-page termination review. A comparison of initial and termination scores provides data on changes during the service period so workers and families can evaluate progress and plan for the future. Content validity for the FAF was developed through a committee and reliability testing has yielded positive results for its internal consistency and inter-rater reliability (McCroskey & Nelson, 1989; McCroskey, Nishimoto, & Subramanian, 1991; Children's Bureau of Southern California, 1997). However, its consistency in repeated administrations and its ability to distinguish between groups and predict outcomes of interest is unclear.

Family Assessment Checklist (FAC). The FAC is a comprehensive assessment of family problems and strengths that was developed for use in an urban, home-based child welfare program to assist workers in establishing goals, planning services, and monitoring changes. The FAC addresses seven major areas: (1) financial status, (2) condition of the home environment, (3) developmental level of the client, (4) the developmental level of the child(ren), (5) parenting skills, (6) nutrition knowledge and practice, and (7) physical and mental health of family members. The FAC is sensitive to changes in family functioning over the course of home-based services. It appears to be economic in terms of personnel demands and time expenditure given that it can be completed by caseworkers based upon observations made in the routine course of service. In a single study, the FAC appeared to have high inter-rater reliability and convergent validity (Cabral & Marie, 1984). However, like the FAF, its consistency in repeated administrations and its ability to distinguish between groups and predict outcomes of interest is unclear.

Ackerman-Schoendorf Scales for Parent Evaluation of Custody (ASPECT). The ASPECT was designed to assist mental health professionals in making child custody recommendations by assessing characteristics of parents and parent-child interactions that are related to effective parenting. The scales include 56 items and represent a system that combines the results of psychological testing, interviews, and observations of each parent and child to provide data regarding the suitability of the parent for custody. While the scales are comprehensive in relation to the family assessment domains, obtaining the data needed for the ASPECT involves considerable time and entails several assessment steps. Nonetheless, the scale has adequate internal consistency and inter-rater reliability and correctly predicted the final disposition of court orders regarding custody in approximately 75% of cases. However, it is

important to note that these scales were developed and tested primarily with predominantly white, married and well-educated parents; therefore, the generalizability of the scale to child welfare populations is unknown (Heinze & Grisso, 1996; Touliatos, Perlmutter, & Holden, 2001).

Darlington Family Assessment System (DFAS). The DFAS is a multisystem-multimethod assessment that consists of three components: (1) the Darlington Family Interview Schedule (DFIS), a structured family interview with an integrated rating scale called the Darlington Family Rating Scale, (DFRS), (2) a battery of self-report questionnaires, including the Social Support Index, Goldberg's General Health Questionnaire, the Eyberg Child Behavior Inventory, the Marital Satisfaction Index, and the McMaster Family Assessment Device, and (3) a task with an associated behavior coding system. DFAS measures twelve problem dimensions using four major perspectives: (1) child-centered (including physical health, development, emotional behavior, relationships, and conduct), (2) parent-centered (including physical health, psychological health, marital partnership, parenting history, and social supports), (3) parent-child interactions, including care, and control, and (4) the whole family/total system perspective (closeness and distance, power hierarchies, emotional atmosphere and rules, and family development). The DFIS requires approximately 1 ½ hours to complete the interview, twenty minutes for clients to complete the self-report questionnaire battery, and fifteen minutes for completion of the task activity. The DFIS has been developed and tested with psychiatric and healthy populations and may be helpful to novice and non-specialty practitioners as a training device. Experienced practitioners may use DFIS to organize clinical observations and inferences and the DFRS can assist practitioners with summarizing clinical observations and treatment planning. The DFIS enhances understanding of both objective and subjective views of family problems, is useful as an integrated package of tools, and appears promising in guiding therapeutic strategies. While it has relatively good inter-rater reliability, concurrent and content validity, and is sensitive to clinical change, the DFIS has not been used with child welfare populations (Wilkinson, 2000; Wilkinson, & Stratton, 1991).

Figure 4. Promising Measures of Comprehensive Family Assessment for Child Welfare

Instruments	Child Welfare Decision/Stage of Assessment			
	Intake/ Screening	Investigation/ Diagnosis	Case Planning	Continuing Services, Placement & Reunification/ Monitoring & Evaluation
North Carolina Family Assessment Scale (NCFAS)		X	X	X
NCFAS for Reunification (NCFAS-R)			X	X
Strengths and Stressors Tracking Device (SSTD)		X	X	X
Family Assessment Form (FAF)		X	X	X
Family Assessment Checklist (FAC)		X	X	X
Ackerman-Schoendorf Scales for Parent Evaluation of Custody (ASPECT)	X		X	
Darlington Family Assessment System (DFAS)	X	X	X	X

X = child welfare decision/stage of assessment for which instrument is used

In summary, of the seven most promising assessment instruments, the NCFAS and the NCFAS-R appear to be the most relevant for use in child welfare settings due to its strengths-based orientation and extensive testing with child welfare populations, despite some of its psychometric limitations. The Darlington Family Assessment System (DFAS) also appears promising given its multi-system, multi-method approach, which mirrors the family assessment process in child welfare by using multiple methods to gain multiple perspectives in a case. It has excellent psychometric properties and is comprehensive nature. However, more research is needed to establish its validity with child welfare populations and to evaluate its feasibility due to lengthy administration time.

Patterns of Social Interaction and Support

We identified four measures for specialized assessment for use at multiple points in the life of the case that focus on patterns of social interaction (including the nature of contact and involvement with others, and the presence or absence of social support networks and relationships at multiple points in the life of the case). As noted in Figure 5, these instruments include the McMaster Model, the Assessment of Strategies in Families-Effectiveness (ASF-E), the Circumplex Model, and the Family Assessment Measure III.

Figure 5. Promising Measures of Social Interaction

Instrument	Child Welfare Decision/Stage of Assessment			
	Intake/ Screening	Investigation/ Diagnosis	Case Planning	Continuing Services, Placement & Reunification/ Monitoring & Evaluation
McMaster Model	X	X	X	X
Assessment of Strategies in Families- Effectiveness (ASF-E)	X	X		X
Circumplex Model		X	X	X
Family Assessment Measure III		X		X

X = child welfare decision/stage of assessment for which instrument is used

McMaster Model. The McMaster Model relies on multiple instruments to assess six dimensions of functioning: (1) problem solving, (2) roles, (3) communication, (4) affective responsiveness, (5) affective involvement, and (6) behavior control. The three complementary instruments include: the Family Assessment Device (FAD), a 60-item self-report questionnaire; the McMaster Clinical Rating Scale (MCRS), an observational rating used by clinician or other observer; and the McMaster Structured Interview of Family Functioning (McSiff), which provides a series of structured questions on each of the six domains. The MCRS and the FAD provide a single score for each of the six dimensions, and the McSiff is used to obtain a reliable clinical rating on the MCRS. The clinical utility and psychometric validity and reliability of the McMaster instruments have been documented in several studies (Epstein et al., 2003; Miller et al., 2000). The FAD is easy to administer and cost effective, has predictive validity for several clinically relevant outcomes, can differentiate between clinical and non-clinical families and is available in at least sixteen languages (Epstein et al., 2003; Miller et al., 2000). The Chinese and Spanish versions of the FAD appear to possess good psychometric properties (Shek, 2001; Shek, 2002; Walrath et al., 2004). While the instruments presently lack normative data on child welfare populations, they may provide early identification of families who may benefit from therapy despite reluctance to seek services (Akister & Stevenson-Hinde, 1991; Miller et al., 2000).

Assessment of Strategies in Families-Effectiveness (ASF-E). The ASF-E is a brief, 20-item screening instrument to determine the perceived need for therapy and to determine progress as a result of family therapy in clinical settings. The ASF-E measures congruence and family health on four dimensions of family behavior patterns and strategies; namely, stability, growth, control, and connectedness/spirituality. High internal consistency and validity have been

established for the ASF-E in the U.S. and the measure has been tested with populations internationally (Friedemann, Astedt-Kurki, & Paavilainen, 2003).

Circumplex Model. The Circumplex battery of instruments integrates three dimensions of family functioning (communication, cohesion, and flexibility) and is designed for use in clinical assessment, treatment planning, and family intervention research. The Circumplex Model includes the Family Adaptability and Cohesion Scale (FACES), a self-report questionnaire that has gone through multiple revisions over the past 20 years to improve the reliability and validity of the instrument. The latest version, the FACES IV, has been found to be reliable and valid for clinical use (Olson & Gorall, 2003). Additional Circumplex measures include the Clinical Rating Scale (CRS) for rating couples and family systems based on clinical interviews or observations; the Family Communication Scale, which focuses on the exchange of factual and emotional information; the Family Satisfaction Scale to determine the family's satisfaction with their functioning; the Family Strengths Scale, which focuses on family characteristics and dynamics that enable families to demonstrate resilience and deal with family problems; and the Family Stress Scale, which taps into levels of stress currently being experienced by family members within their family system (Olson, 2000; Olson & Gorall, 2003). While the CRS has been validated, it is unclear whether self-report questionnaires other than the FACES IV have established validity and reliability.

Family Assessment Measure III. The FAM III is a set of self-report questionnaires that measure family strengths and weaknesses in the seven constructs related to: (1) task accomplishment, (2) role performance, (3) communication, (4) affective expression, (5) affective involvement, (6) control, and (7) values and norms. While the concepts are similar to those measured in the McMaster Model, the FAM III is unique in assessing family strengths and weaknesses from perspectives on three scales: the family as a system (general scale), various dyadic relationships (dyadics scale), and individual family members (self-rating scale). The collection of data from all three perspectives facilitates the analysis of family processes from multiple system levels. The FAM III consists of 94 items and can be completed by family members at least 10-12 years of age. Numerous studies attest to the clinical utility of the FAM III, including its ability to differentiate between clinical and non-clinical families and its predictive validity in relation to children's problems. The FAM III has demonstrated sensitivity

to change in treatment, has been developed and tested with clinical and non-clinical families, and has twenty years of research to support its efficacy (Skinner et al., 2000).

In summary, research has found the FACES, the FAM III, and the FAD to be highly correlated, to suggest that these three instruments may be interchangeable (Olson, 2000; Beavers & Hampson, 2000). Although the Circumplex instruments appears best at providing a multisystem-multimethod assessment of the family, the McMaster instruments provide the clearest link with a therapeutic model of intervention (Carlson, 2003). McMaster instruments also have demonstrated superior sensitivity in identifying families with clinical needs and greater correspondence between clinical rating scales and family member self-report inventories when compared to the Circumplex instruments (Drumm & Fitzgerald, 2002). More studies comparing the treatment utility of the various instruments are needed, especially with respect to child welfare populations.

Parenting Practices

In addition to the seven comprehensive measures of family assessment and the four specialized measures of patterns of family social interaction, five measures were identified as promising for the specialized assessment of parenting practices among families that have come to the attention of the child welfare system: (1) the Adult-Adolescent Parenting Inventory (AAPI); (2) the Child Abuse Potential Inventory (CAPI); (3) the Parental Empathy Measure (PEM); (4) the Parenting Stress Index (PSI); (5) and the Beavers Model of Family Assessment (see Figure 6).

Figure 6. Promising Measures of Parenting Practices

Instrument	Child Welfare Decision/Stage of Assessment			
	Intake/ Screening	Investigation/ Diagnosis	Case Planning	Continuing Services, Placement & Reunification/ Monitoring & Evaluation
Adult-Adolescent Parenting Inventory	X	X		X
Child Abuse Potential Inventory (CAPI)	X			X
Parenting Stress Inventory (PSI)	X			
Parental Empathy Measure (PEM)	X			
Beavers Model of Family Assessment	X	X	X	X

X = child welfare decision/stage of assessment for which instrument is used

Adult-Adolescent Parenting Inventory (AAPI). The AAPI is designed to identify high-risk parenting activities and behaviors that are known to be attributable to child abuse and

neglect, and may also be used to assess patterns of family social interaction. It is a self-report inventory consisting of forty five-point Likert scale items, and can be administered at multiple points over the course of a child welfare case for the purposes of screening, diagnosis, and monitoring progress and clinical change over time. Advantages of this instrument include a brief administration time of approximately twenty minutes and suitability for parents with a fifth-grade reading level or above. Additionally, the AAPI can be read orally to non-readers and a Spanish version is available for Spanish-reading parents. Over twenty years of research have provided considerable evidence of the psychometric strength of this instrument, including high internal consistency and significant diagnostic and discriminatory validity in discerning non-abusive parents and known abusive parents (Bavolek & Keene, 2001).

Child Abuse Potential Inventory (CAPI). The CAPI is also designed to identify parents who are most likely to be at risk for child abuse by assessing problematic parenting practices and social interaction, and was developed as a tool specifically for child protective services workers in their investigations of reported child abuse cases. While the CAPI was originally designed as a preliminary screening tool to discriminate between abusive and non-abusive parents, treatment/intervention programs have successfully used the CAPI at pre- and post-treatment to assess progress and clinical change (Milner, 1994). It is a self-administered 160-item questionnaire that assesses six primary clinical factors: (1) distress, (2) rigidity, (3) unhappiness, (4) problems with child and self, (5) problems with family, and (6) problems with others. Additionally, the instrument includes three validity subscales that help mitigate potential self-report bias. Other advantages of the CAPI include a brief administration time of approximately twenty minutes and a third-grade reading level requirement that permits its suitability for use with parents with limited literacy proficiency. Similar to the AAPI, the CAPI has undergone substantial psychometric evaluations and has demonstrated significant discriminative validity, high internal consistency and test-retest reliabilities (Heinze & Grisso, 1996).

Parental Empathy Measure (PEM). The PEM is a promising instrument for screening for abusive or neglecting parenting behaviors/practices. It is a semi-structured interview with open-ended questions assessing parental attention to signals, attributes, emotional/behavioral responses to, and perceptions of their children. In addition to these parenting practices and behaviors, the PEM includes items addressing past involvement with child protective services. One of the strongest features of this instrument is the comprehensiveness of its psychometric

evaluation; reliability and validity tests indicate that the PEM has good sensitivity for identifying abusive parents, good inter-rater reliability, high internal consistency, and high construct reliability when measured against the CAPI. Furthermore, the PEM also includes a measure of social desirability that was found to be effective in detecting biased responses. However, the PEM lacks the advantage of administrative brevity exhibited by the previous measures; for example, the PEM contains open-ended items and general administration time cannot be estimated because it depends on specific case characteristics (Kilpatrick, 2005).

Parenting Stress Index (PSI). The PSI also screens for abusive or neglecting parenting behaviors/practices, and assesses social interaction characteristics that may affect the quality of family functioning. The current version of the PSI contains 101 self-report items assessing the parenting domain (competence, social isolation, attachment to child, health, role restriction, depression, spouse) and the child domain (distractibility, adaptability, parent reinforcement, demandingness, mood, acceptability). An optional nineteen-item life stress scale is also provided (Terry, 1991a). The advantages include a brief administration time of 20-25 minutes for the full instrument (recommended for a more comprehensive assessment), and a 36-item short form is also available for situations requiring more rapid assessment. Additionally, the PSI is available in eight languages permitting its use with non-English reading populations. Psychometric evaluations have demonstrated high internal consistency, high correlations with instruments measuring the same construct, and relatively good test-retest reliabilities (Terry, 1991b). However, evaluators caution that low ratings on the PSI do not necessarily indicate the absence of problems, in part due to the lack of validity measures that address potential social desirability bias (Touliatos et al., 2001).

Beavers Model of Family Assessment. The Beavers Model of Family Assessment consists of three instruments that assess parenting practices using a combination of self-report and observational methods: (1) the Beavers Self-Report Family Inventory (SRFI) which measures self-reported parenting practices and competence; and (2) the Beavers Interactional Style Scale (BISS) and the Beavers Interactional Competence Scale (BICS), which are both scored using observer ratings of parenting style and competence based on a ten minute observation of a semi-structured episode of family interaction (Beavers & Hampson, 2000). The Beavers instruments may be administered throughout the course of a child welfare case, and consequently, assist with multiple stages of assessment, including screening, diagnosis, treatment

planning, and monitoring progress/follow-up. The SFRI is a 36-item Likert-format questionnaire that may be completed by family members eleven years of age or older, and is brief and easy to score (McCubbin, McCubbin, Thompson, & Huang, 1989). Psychometric evidence of its reliability and validity is substantial; studies demonstrate a 91% correct classification of clinical versus non-clinical cases, high test-retest reliability, high internal consistency, and concurrent validity (Halvorsen, 1991). The BICS also has demonstrated strong reliability and validity; studies indicate that this instrument has a 65% sensitivity rate for clinical families, a 90% specificity rate for non-clinical families, high inter-rater reliability, high overall test-retest reliability, and high construct validity (Carlson, 2003). Psychometric evidence of the reliability/validity of the BISS is still in progress; however, one study suggests that it has limited descriptive and discriminative power in comparison to the other two Beavers measures (Drumm, Carr, & Fitzgerald, 2000). Although some studies have administered and evaluated these instruments separately, the developers of the Beavers model indicate that a more comprehensive family assessment would be facilitated by the conjunctive use of all three instruments (Beavers & Hampson, 2000).

Background Characteristics

Three measures were identified as possible candidates for the assessment of family background characteristics related to a history of child abuse and neglect; namely, the Family Systems Stressors Strength Inventory (FSSSI), the Hispanic Stress Inventory (HIS), and the Ontario Child Neglect Index (CNI). While these measures have been designed for clinical use, more psychometric evaluation is needed to determine their validity and reliability.

Family Systems Stressors Strength Inventory. The FSSSI is a 53-item self administered questionnaire that is designed to identify the perceptions of family members regarding general and specific family stressors and strengths. When used as a clinical tool, the instrument can provide direction for intervention planning and has the advantage of assessing family strengths as well as difficulties. Content validity was assessed through inter-rater agreement for conceptual fit and for clarity of items. However, as previously mentioned, very little psychometric data are available for this instrument and reliability of this instrument is unknown (Touliatos et al., 2001).

Hispanic Stress Inventory. The HIS is designed as a culturally appropriate tool for assessing stressors within Hispanic families, including marital stress, family stress, occupational stress, economic stress, discrimination stress, and acculturation stress. Two versions of the

instrument are available, a 73-item self-report questionnaire designed for use with immigrant families and a 59-item self-report questionnaire adapted for US-born family members. A key advantage of the HIS is its culture-specific application for diagnosing and planning interventions for Hispanic families, and its subscales have been found to have high internal consistency and high test-retest reliabilities (Cervantes, Padilla, & De Snyder, 1991). Additional psychometric tests should be conducted in order to further substantiate its reliability and validity (Touliatos et al., 2001).

Ontario Child Neglect Index (CNI). The Ontario CNI is a brief 6-item caseworker-rated instrument, which is designed to identify the type and severity of neglect that children experience from their primary caretakers. In addition to evaluating history of physical abuse, sexual harm and criminal activity, the CNI can also be used to identify problematic areas in basic needs provision, including nutrition, clothing and hygiene, physical care, mental health care, and developmental/educational care. The brevity of the instrument helps facilitate an immediate screening and diagnostic impression of the family, however, may also pose a potential limitation through loss of accuracy, comprehensiveness, and susceptibility to bias (Touliatos et al., 2001). The CNI has demonstrated a high level of consistency in repeated administrations and high inter-rater reliability (Trócme, 1996).

Assessing Basic Needs

Few measures have been developed for the sole purpose of assessing a family's basic needs. As previously mentioned, the Ontario Child Neglect Index includes items that screen for potential deficiencies in basic needs provision; however, it does not provide a thorough assessment of this domain. The Home Observation for the Measurement of Environment (HOME) is perhaps the most comprehensive and widely used measure that assesses the family's capacity to fulfill basic needs, in addition to assessing patterns of social interaction and parenting practices. The HOME may be used clinically for screening and intervention planning purposes. Several versions of the HOME that are tailored to age-specific populations are available, including versions suitable for assessing families with infants/toddlers (age 0-3 years), children in early childhood (age 3-6 years), children in middle childhood (age 6-10 years), and early adolescents (age 10-15 years). Although different versions of the measure vary in number of items, ranging from 45-60 items, all versions employ observation and semi-structured interviewing methods to obtain evaluation scores for the family and can be administered in about

one hour. The HOME has been used in a number of studies with minority and special needs populations, and versions adapted for these populations are also available (Caldwell & Bradley, 2003). Psychometric properties of the HOME include high inter-rater reliability and high internal reliability (Elardo & Bradley, 1981).

The Family Economic Strain Scale (FES) is another measure that is potentially promising for the assessment of basic needs fulfillment for families in the child welfare system. It is a 13-item self-administered questionnaire that is designed to evaluate the financial difficulties of single and two-parent families (Hilton & Devall, 1997). Preliminary reliability tests have demonstrated high internal consistency for the measure, however additional psychometric evaluations should be conducted to ensure its reliability and validity (Touliatos et al., 2001).

Implications for Practice

Rather than replacing clinical judgment, psychometrically validated family assessment instruments can enhance the family assessment process by structuring the collection of information and ensuring that relevant categories of family assessment are evaluated. Practitioners can use the results of these assessments to appropriately refer clients to services and to demonstrate the linkages between assessment, referrals, service provision, and child and family outcomes to supervisors, the courts, and other professionals working on the case, and to monitor client progress over time. At the programmatic level, assessment results can be aggregated and analyzed to assess overall program performance and to identify service areas in need of improvement.

Additional Psychometric Testing

The large number of measures related to patterns of social interaction and parenting practices suggest that the family assessment field has been rapidly expanding based on theoretically diverse but overlapping research traditions (including family systems theory, family therapy research, the literature on risk and resilience and the assessment of parenting). Significant effort has been made to bridge these research traditions to produce comprehensive family assessment instruments that meet the needs of child welfare practitioners. These efforts, which have been made incrementally by a small number of researchers over the past fifteen years, are reflected in the introduction and refinement of measures such as the Family Assessment Form and the North Carolina Family Assessment Scale (NCFAS). In the case of new instruments, it can take several years to establish their structural components and validate them

(Skinner, 1987). However, establishing additional psychometric information for existing measures that appear appropriate for child welfare services represents a task that agencies can manage through pilot testing and smaller scale studies by way of university-agency partnerships, inter-agency research consortiums, or independent contracting.

Key Administrative Supports

In addition to carefully reviewing the measurement criteria and the practical implications for use of a family assessment instrument in child welfare, it is important for managers to assess the agency resources that may be necessary to successfully integrate family assessment. Comprehensive family assessment is a process rather than the simple completion of a tool; therefore, once decisions are made regarding the selection of instruments, consideration will need to be given to how the agency will build or modify the existing infrastructure to support it. The family assessment process includes at least nine components: (1) the evaluation of information; (2) interviewing; (3) obtaining and integrating information from more specialized assessments; (4) identifying family strengths and needs; (5) decision-making; (6) documenting and maintaining records; (7) linking assessments to service plans; (8) evaluating outcomes; and (9) disseminating information to other providers, as needed (HHS, 2006). Figure 7 outlines four areas of administrative support (adapted from HHS, 2006).

Figure 7. Administrative Supports for Family Assessment

Administrative Support	Description
Policies	Policies that require family assessment processes; incorporate workload requirements into staffing needs and time frames; define its characteristics and its overarching framework, distinguish it from risk and safety assessment, and clarify relationships between multiple assessments
Training, Clinical Supervision, and Mentoring	Training, supervision and mentoring on the family assessment process and its relationship to other types of assessment; training on engagement; interviewing skills; interpretation of specialized assessments; information integration and decision-making; documentation; linkage of assessment information to service planning; coordination of information with other service providers; timing of re-assessment; outcomes evaluation
Systems of Accountability and Evaluation	Systems that assure that the family assessment process takes place and guidelines are followed; assessing how results are used in service plans; evaluating whether service needs are addressed and how case progress is tied to the assessment; establishing a system whereby the entire process of assessment, service planning, service delivery, progress reviews, key decisions, and program outcomes are documented and evaluated
Contracting	Requirements such as purchase of service contract provisions for family assessment that are consonant with child welfare family assessment processes; clarification of reporting requirements and policies and processes of information sharing; examination of cross-training opportunities

For example, policy needs to reflect the institutional support for the family assessment process, the parameters and expectations of the family assessment process, and the needed staffing support. A comprehensive family assessment process incorporates information collected through other assessments, such as safety, risk, and child assessments. Policies also need to address how these multiple assessments are conducted in day-to-day practice and how this information will be incorporated into the development of service plans that address the major factors that affect safety, permanency, and child well-being over time.

Given that the engagement and building of worker-client rapport are of central importance in gathering information from families regarding their needs and strengths, organizational and administrative supports are necessary for implementing family assessment techniques. These include allocating staff time for assessment, formal training, clinical supervision, and mentoring in areas such as completing comprehensive assessments in a culturally sensitive manner, engaging families in a change process, and reaching the appropriate conclusions about the meaning of the information gathered.

Systems of accountability, such as quality assurance programs, represent a key support for building the infrastructure that links assessment information to service plans. To illustrate, Figure 8 demonstrates one approach to quality assurance that is currently in place at a local community-based agency that provides differential response services. After the agency receives the child welfare referral, a new worker and a Master's level mentor meet with the family, make their observations, and then jointly complete the NCFAS afterwards during a case conference to establish inter-rater reliability. Results of the NCFAS are then used to develop the service plan, which guides the provision of services. Case notes are used to continuously update the case and to document decisions. The NCFAS is conducted at multiple points during the case to monitor progress and to evaluate the outcomes of service at case closure. The quality assurance component of the process is enhanced through a peer review process using accreditation standards for child welfare developed by the Council on Accreditation. This process is used to monitor and evaluate the linkage of assessment information, service plan specifications, case notes, and service outcomes.

Figure 8. Quality Assurance Steps in Family Assessment

Step 1: Evaluation of Child Welfare Referral	Step 2: Joint NCFAS Assessment (Worker & Master's Level Mentor)	Step 3: Development of Service Plan, Service Provision and Linkage	Step 4: Linkage of Case Notes to Service Plan	Step 5: Repeat NCFAS to Monitor Progress and Evaluate Outcomes of Service at Case Closure	Step 6: Council on Accreditation Peer Review Process
---	--	---	--	--	--

Quality assurance programs represent an important administrative support for monitoring and evaluating the implementation and outcomes of the family assessment process and can also be used to identify needs for changes in policies, training, clinical supervision, and mentoring. While child welfare agencies have the ultimate responsibility for the case plan, increasingly, community-based organizations are often the contracted providers of services. Therefore, systems of accountability naturally extend to services that are provided through other agencies. In relation to family assessment, contract provisions and memoranda of understanding represent the mechanisms through which family assessment processes and information sharing can be coordinated and clarified.

Appendix A
Measurement Criteria (Perlmutter & Czar, 2001; Pedhazur & Schmelkin, 1991)

Measurement Criteria	Definitions
VALIDITY	A valid instrument measures what it claims to measure. Major types of validity include content-related validity, criterion-related validity, and construct-related validity.
Content-Related Validity	A well-constructed instrument can be considered to adequately represent a specified area of knowledge as well as avoid the effects of unrelated variables. Judges that render opinions about the suitability of items in an instrument commonly assess content validity. There is no accepted standard of agreement established for retaining an item, however, better scales tend to elicit greater agreement among judges.
Criterion-Related Validity	Relates to predicting an individual's performance against a score on an existing instrument or a future outcome. Two methods for establishing criterion-related validity include concurrent and predictive strategies. <i>Concurrent validity.</i> Using this method, tests are administered to individuals for whom criterion data are already available for accurate indicators of the construct under study (e.g., an existing measure of family functioning). Correlations between scores or ratings on the new instrument and those obtained on the existing instrument are then established. <i>Predictive validity.</i> Denotes an instrument's ability to predict future outcomes or status from scores on an instrument. <i>Differential predictive validity</i> refers to an instrument's ability to predict these outcomes for different groups (e.g., abusive vs. non-abusive families).
Construct-Related Validity	Relates to the degree to which an instrument successfully measures a theoretical concept. Two methods, convergent and divergent validity, establish the construct-related validity of a test. <i>Convergent-Divergent Validity.</i> When different measures of a concept yield similar results they converge; demonstrating convergent validity typically involves correlating two existing measures or correlating a new measure with an existing measure. When concepts can be empirically differentiated from other concepts they diverge; therefore, measures of different constructs will possess low linear correlations.
RELIABILITY	Reliability is an index of the degree to which individual differences in scoring reflect actual differences in the characteristic under consideration versus chance errors. There are four types of reliability: test-retest, alternate form, inter-rater, and internal consistency.
Test-Retest Reliability	Refers to the degree to which generalizations can be made about test scores from one administration to the next. Established by correlating results of baseline and subsequent administrations. Higher test-retest coefficients generally mean that scores are less susceptible to random changes in the condition of test takers or the testing environment.
Alternate-Form	When equivalent forms of a test are administered to the same person on two occasions, the reliability coefficient is the correlation between scores obtained on the two tests. The higher the scores, the more likely it is that the different test forms are measuring the same characteristics.
Inter-rater Reliability	Relates to the extent to which two or more people arrive at the same result when observing and/or rating the same event. It is most frequently reported for observational techniques when events are recorded or ratings of behaviors are made.
Internal Consistency	Cronbach's alpha is the measure of internal consistency by which we infer that items within a scale or subscale measure the same construct. Alpha rises when the average inter-item correlation between items increases. It also increases with increased numbers of items, so long as the quality of those items remains high. Therefore, when considering the value of alpha for a given instrument or subscale, both the reported alpha and the number of items must be considered.

Appendix B
BASSC Search Protocol

Search Terms

- 1) assessment of families
- 2) family assessment
- 3) family assessment and child welfare
- 4) family assessment and clinical
- 5) family assessment and device
- 6) family assessment and evaluation
- 7) family assessment and guide
- 8) family assessment and intervention
- 9) family assessment and measure
- 10) family assessment and measurement
- 11) family assessment and mental health
- 12) family assessment and model
- 13) family assessment and research
- 14) family assessment and scale
- 15) family assessment and school
- 16) family assessment and service
- 17) family assessment and social services
- 18) family assessment and therapy
- 19) family assessment and treatment
- 20) family functioning assessment
- 21) family strengths assessment

Databases

Academic databases for books and articles

Pathfinder or Melvyl

ArticleFirst

ERIC

Expanded Academic ASAP

Family and Society Studies Worldwide

PAIS International

PsychInfo

Social Science Citation Index

Social Services Abstracts

Social Work Abstracts

Sociological Abstracts

Systematic Reviews

Campbell Collaboration – C2-Spectre & C2-Ripe

Children and Family Research Center

Cochrane Library

ESRC Evidence Network
NHS Centre for Reviews & Dissemination
Social Care Institute for Excellence

Research Institutes

Brookings Institute
Manpower Demonstration Research Corporation
Mathematica Policy Research, Inc.
Urban Institute
RAND
GAO
National Academy of Sciences
Chapin Hall
CASRC (San Diego)

Conference Proceedings

PapersFirst (UCB Database)
Proceedings (UCB Database)

Dissertation Abstracts

DigitalDissertations (UCB database)

Professional Evaluation Listserves

EVALTALK
GOVTEVAL
ChildMaltreatmentListserve

Internet

Google
Dogpile

George Warren Brown School of Social Work at Washington University in St. Louis, Center for Mental Health Services Research : <http://gwweb.wustl.edu/cmhsr/measure/category.html>

Administration for Children and Families, DHHS, Resources for Measuring Services & Outcome in Head Start Programs Serving Infants & Toddlers:
http://www.acf.hhs.gov/programs/opre/ehs/perf_measures/reports/resources_measuring/res_measures_toc.html

Violence Institute of New Jersey at UMDNJ:

http://www.umdj.edu/vinjweb/research_projects/instrument_inventory/instrument_inventory.html

http://www.lib.berkeley.edu/SOCW/socw_instruments.html

Appendix C Instruments and Models Evaluated

Abusive Behavior Inventory	Ackerman-Schoendorf Scales for Parent Evaluation of Custody
Adult-Adolescent Parenting Inventory	Assessment of Strategies in Families-Effectiveness (ASF-E)
Attitudes toward Wife Abuse Scale	Beavers Model
Brother-Sister Questionnaire	Child Abuse Blame Scale- Physical Abuse
Child Abuse Potential Inventory	Childhood Level of Living Scale
Circumplex Model (FACES IV)	Colorado Family Assessment
Culturagrams	Darlington Family Assessment System (DFAS)
Family Adaptation Model	Family APGAR
Family Assessment Checklist	Family Assessment Form
Family Assessment Measure III	Family Behavioral Snapshot
Family Concept Assessment Method (FCAM)	Family Daily Hassles Inventory
Family Distress Index	Family Economic Strain
Family Emotional Involvement and Criticism Scale (FEICS)	Family Environment Scale
Family Evaluation Form (FEF)	Family Functioning Index (FFI)
Family Functioning Questionnaire	Family Functioning Style Scale
Family Hardiness Index	Family Impact Questionnaire
Family Interaction Q-Sort	Family Profile II
Family Strengths Inventory	Family Strengths Scale
Family Systems Stressor Strength Inventory	Global Assessment of Relational Functioning (GARF)
Global Coding Scheme	Global Family Interaction Scales (FIS-II)
Hispanic Stress Inventory	Home Observation for Measurement of the Environment
Interparental Conflict Scale	Interpersonal Support Evaluation List (ISEL)
Inventory of Battering Experiences	Inventory of Socially Supportive Behaviors
Kempe Family Stress Inventory	Life Stressors and Social Resources Inventory – Adult & Youth Forms
Maternal Characteristics Scale	McMaster Model (Family Assessment Device)
MMPI Family Scale	Mother’s Activity Checklist
North Carolina Family Assessment Scale (NCFAS)	North Carolina Family Assessment Scale for Reunification (NCFAS-R)
Ontario Child Neglect Index	Parent Child Conflict Tactics Scale
Parent Child Relationship Inventory	Parental Empathy Measure (PEM)
Parental Stress Scale	Parent-Child Relationship Scale (PCRS, subset of CAPA)
Parenting Stress Index	Parenting Stress Inventory
Physical Abuse and Psychological Abuse	Propensity for Abusiveness Scale
Revised Conflict Tactics Scales (CFTS-2/RCTS)	Scale of Neglectful Parenting
Severity of Violence Scales	Simulated Family Activity Measure (SIMFAM)
Social Environment Inventory	Standardized Clinical Family Interview
Strengths and Stressors Tracking Device (SSTD)	Stress Index for Parents of Adolescents
Structural Family Interaction Scale (SFIS)	Structural Family Systems Rating Scale (SFSR)
Structured Clinical Interview (SCI)	Yale Guide to Family Assessment

References

- Akister, J., & Stevenson-Hinde, J. (1991). Identifying families at risk: Exploring the potential of the McMaster Family Assessment Device. *Journal of Family Therapy, 13*(4), 411-421.
- Bavolek, S. J., & Keene, R. G. (2001). *Adult-Adolescent Parenting Inventory AAPI-2: Administration and Development Handbook*. Park City, UT: Family Development Resources, Inc.
- Beavers, R., & Hampson, R. B. (2000). The Beavers System Model of Family Functioning. *Journal of Family Therapy, 22*(2), 128-143.
- Berry, M., Cash, S. J., & Mathiesen, S. G. (2003). Validation of the strengths and stressors tracking device with a child welfare population. *Child Welfare, 82*(3), 293-318.
- Boss, P., Doherty, W., LaRossa, R., Schumm, W., Steinmetz, S. (Ed.). (1993). *Sourcebook of Family Theories and Methods*. New York: Plenum Press.
- Cabral, R. J., & Strang, M. (1984). Measuring child care: An examination of three assessment measures. *Journal of Social Service Research, 7*(2), 65-77.
- Caldwell, B. M., & Bradley, R. H. (2003). *HOME Inventory Administration Manual*. Little Rock, AR: Print Design, Inc.
- Carlson, C. I. (1989). Criteria for family assessment in research and intervention contexts. *Journal of Family Psychology, 3*(2), 158-176.
- Carlson, C. I. (2003). Assessing the family context. In C. R. Reynolds, & Kamphaus, R. W. (Ed.), *Handbook of Psychological and Educational Assessment of Children: Intelligence, Aptitude, and Achievement* (2nd ed., pp. 473-492). New York: Guilford Press.
- Cervantes, R. C., Padilla, A. M., & de Snyder, N. S. (1991). The Hispanic Stress Inventory: A culturally relevant approach to psychosocial assessment. *Psychological Assessment, 3*(3), 438-447.
- Cromwell, R. E. P., G. W. (1983). Multisystem-Multimethod Family Assessment in Clinical Contexts. *Family Process, 22*(2), 147-163.
- Day, P., Robison, S., & Sheikh, L. (1998). *Ours to Keep: A Guide for Building a Community Assessment Strategy for Child Protection*. Washington, D.C.: CWLA Press.
- DePanfilis, D., & Salus, M. K. (2003). *Child Protective Services: A Guide for Caseworkers*. Rockville, MD: U.S. Department of Health & Human Services.
- Drumm, M., Carr, A., & Fitzgerald, M. (2000). The Beavers, McMaster and Circumplex clinical rating scales: A study of their sensitivity, specificity and discriminant validity. *Journal of Family Therapy, 22*(2), 225-258.
- Elardo, R. B., Robert. (1981). The Home Observation for Measurement of the Environment: A review of research. *Developmental Review, 1*, 113-145.
- Epstein, N. B., Bishop, D. S., & Levin, S. (1978). The McMaster model of family functioning. *Journal of Marriage and Family Counseling, 40*, 585-593.
- Epstein, N. B., Ryan, C. E., Bishop, D. S., Miller, I. W., & Keitner, G. I. (2003). The McMaster Model: A view of healthy family functioning. In F. Walsh (Ed.), *Normal Family Processes: Growing Diversity and Complexity* (3rd ed., pp. 581-607). New York: Guilford Press.
- Friedemann, M., Åstedt-Kurki, P., & Paavilainen, E. (2003). Development of a Family Assessment Instrument for Transcultural Use. *Journal of Transcultural Nursing, 14*(2), 90-99.
- Halvorsen, J. G. (1991). Self-Report Family Assessment Instruments: An Evaluative Review. *Family Practice Research Journal, 11*(1), 21-53.

- Hawkins, R. P. (1979). The functions of assessment: Implications for selection and development of devices for assessing repertoires in clinical, educational, and other settings. *Journal of Applied Behavior Analysis, 12*, 501-516.
- Heinze, M. C., & Grisso, T. (1996). Review of instruments assessing parenting competencies used in child custody evaluations. *Behavioral Sciences and the Law, 14*, 293-313.
- Hill, R. (1949). *Families under Stress*. Westport, CT: Greenwood Press.
- Hilton, J. M., & Devall, E. L. (1997). The Family Economic Strain Scale: Development and evaluation of the instrument with single and two-parent families. *Journal of Family Economics Issues, 18*(3), 247-271.
- Johnson, M. A., & Wells, S. J. (2000). *Measuring Success in Child Welfare, National Study of Outcome Measurement in Public Child Welfare Services: Results and Recommendations*. Urbana, IL: Children and Family Research Center, School of Social Work, University of Illinois at Urbana-Champaign.
- Kilpatrick, K. L. (2005). The Parental Empathy Measure: A New Approach to Assessing Child Maltreatment Risk. *American Journal of Orthopsychiatry, 75*(4), 608-620.
- Kirk, R. S., Kim, M. M., & Griffith, D. P. (in press). Advances in the reliability and validity of the North Carolina Family Assessment Scale. *Journal of Human Behavior in the Social Environment*.
- McCroskey, J., & Nelson, J. (1989). Practice-based research in a family-support program: The Family Connection project example. *Child Welfare, 68*(6), 573-587.
- McCroskey, J., Nishimoto, R., & Subramanian, K. (1991). Assessment in family support programs: Initial reliability and validity testing of the family assessment form. *Child Welfare, 70*(1), 19-33.
- McCubbin, M. A., & McCubbin, H. I. (1987). Family stress theory and assessment: The T-double ABCX model of family adjustment and adaptation. In H. I. McCubbin, & Thompson, A. I. (Ed.), *Family Assessment Inventories for Research and Practice* (pp. 1-32). Madison, WI: University of Wisconsin-Madison.
- McCubbin, H. I., McCubbin, M. A., Thompson, A. I., & Huang, S. (1989). Family assessment and self-report instruments in family medicine research. In C. Ramsey (Ed.), *Family Systems in Medicine* (pp. 181-214). New York: Guilford Press.
- Miller, I. W., Ryan, C. E., Keitner, G. I., Bishop, D. S.; & Epstein, N. B. (2000). The McMaster Approach to Families: Theory, assessment, treatment and research. *Journal of Family Therapy, 22*(2), 168-189.
- Milner, J. (1994). Assessing physical child abuse risk: The Child Abuse Potential Inventory. *Clinical Psychology Review, 14*(6), 547-583.
- Olson, D. H. (2000). Circumplex Model of Marital and Family Systems. *Journal of Family Therapy, 22*(2), 144-167.
- Olson, D. H., & Gorall, D. M. (2003). Circumplex Model of Marital and Family Systems. In F. Walsh (Ed.), *Normal family processes: Growing diversity and complexity* (3rd ed., pp. 514-548). New York: Guilford Press.
- Pecora, P. J., Fraser, M. W., Nelson, K. E., McCroskey, J., & Meezan, W. (1995). *Evaluating Family-Based Services*. New York: Aldine de Gruyter.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, Design, & Analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Perlmutter, B. F., & Czar, G. (2001). Developing, interpreting, and using family assessment techniques. In J. Touliatos, Perlmutter, B. F., & Holden, G. W. (Ed.), *Handbook of Family Measurement Techniques*. Thousand Oaks, CA: Sage Publications, Inc.
- Reed-Ashcraft, K., Kirk, R. S., & Fraser, M. W. (2001). The reliability and validity of the North Carolina Family Assessment Scale. *Research on Social Work Practice, 11*(4), 503-520.
- Shek, D. T. L. (2001). The general functioning scale of the Family Assessment Device: Does it work with Chinese adolescents? *Journal of Clinical Psychology, 57*(2), 1503-1516.
- Shek, D. T. L. (2002). Assessment of family functioning in Chinese adolescents: The Chinese version of the Family Assessment Device. *Research on Social Work Practice, 12*(4), 502-524.
- Skinner, H. A. (1987). Self-report: Instruments for family assessment. In J. Theodore (Ed.), *Family Interaction and Psychopathology: Theories, Methods, and Findings* (pp. 427-452). New York: Plenum Press.
- Skinner, H., Steinhauser, P., & Sitarenios, G. (2000). Family Assessment Measure (FAM) and Process Model of Family Functioning. *Journal of Family Therapy, 22*(2), 190-210.
- Terry, D. J. (1991). Predictors of subjective stress in a sample of new parents. *Australian Journal of Psychology, 43*(1), 29-36.
- Terry, D. J. (1991). Stress, coping and adaptation to new parenthood. *Journal of Social and Personal Relationships, 8*, 527-547.
- Touliatos, J., Perlmutter, B., & Holden, G. (2001). *Handbook of Family Measurement Techniques: Abstracts* (Vol. 2). Thousand Oaks, CA: Sage Publications.
- Trocme, N. (1996). Development and preliminary evaluation of the Ontario Child Neglect Index. *Child Maltreatment, 1*(2), 145-155.
- U.S. Department of Health and Human Services. (2006). *Comprehensive Family Assessment Guidelines for Child Welfare*. Washington, DC: Administration for Children and Families Children's Bureau.
- Walrath, C., Franco, E., Liao, Q., & Holden, W. (2004). Measures of child emotional and behavioral strengths and family functioning: A preliminary report on the reliability and validity of their Spanish translations. *Journal of Psychoeducational Assessment, 22*, 209-219.
- Wilkinson, I. (2000). The Darlington Family Assessment System: Clinical guidelines for practitioners. *Journal of Family Therapy, 22*(2), 211-224.
- Wilkinson, I., & Stratton, P. (1991). The reliability and validity of a system for family assessment. *Journal of Family Therapy, 13*(1), 73-94.

¹ Available at http://cssr.berkeley.edu/bassc/projects_practice.asp

² Numbers do not add to 85 given that instruments address multiple domains