

Development, Intent, and Use of the North Carolina Family Assessment Scales, and Their Relation to Reliability and Validity of the Scales

Raymond S Kirk, PhD
Principal NCFAS Scale Developer, and
Director of Research, Independent Living Resources, Inc.
Durham, North Carolina
ray.kirk@ilrinc.com

In cooperation with the
National Family Preservation Network

Introduction

This paper presents historical and philosophical information about the development of the North Carolina Family Assessment Scales (NCFASes) and discusses research conducted by the scale developers to establish the psychometric properties of the scales, specifically the scales' reliability and validity in child welfare practice settings. Research conducted by others that informs the discussion of NCFAS reliability and validity will also be discussed. It is very important to understand the development process and philosophical underpinnings of the NCFAS scales because the intended uses of the NCFAS scales differ from those of many other diagnostic instruments and decision-making/driving models that are also used in child welfare practice setting.

Collectively, the NCFAS scales comprise the original North Carolina Family Assessment Scale (NCFAS), the North Carolina Family Assessment Scale for Reunification (NCFAS-R), the North Carolina Family Assessment Scale for General Services (NCFAS-G), and the NCFAS-G+R, which is a form of the scales that combines the NCFAS-G with the NCFAS-R. Two additional domains have been added that can be used as companions to any of the other NCFASes. These 2 domains are Trauma, and Post-Trauma Well-Being. The key to understanding the intent, use, and validity of the NCFAS instruments lies in understanding their reason for development, the rating strategy, and the definition of "Baseline/Adequate."

Many instruments that claim to have high inter-rater reliability were developed to provide diagnostic capabilities, and for which the developers intend to establish clinical cut scores or universally applicable norms. The NCFAS scales were not intended for that purpose. In fact, during training on the use of the scale, we specifically discourage using the NCFAS scales in this manner. Rather, the NCFAS scales are intended as *case practice tools* to aid in the assessment of family functioning for purposes of service planning and goal setting. They are structured to provide an *organizing framework* for social workers to use in case practice as a vehicle for assuring that a comprehensive assessment is conducted for families. This is necessary because the information in a Child Protective Services report (the original service population for NCFAS development were more serious Abuse/Neglect cases) may be based on the observations of an untrained reporter (e.g., neighbor, teacher, bystander) of the alleged maltreatment. The tendency of reporters to focus on deficits and problems may result in the failure to notice a variety of other issues that must be addressed in order to prevent placement or to effect a successful reunification, or more simply to help a family improve its functioning and harmony as a family system. That is the purpose of the NCFAS scales.

All NCFAS scales use the same rating strategy for all scale items and domains: a 6-point scale ranging from "Clear Strength" to "Serious Problem." In relation to family functioning, strengths lie above "Baseline/Adequate" and problems lie below "Baseline/Adequate." The definition of "Baseline/Adequate" is that "there is no legal, moral, or ethical reason for a public child protection mandate to be exercised." This definition does not imply that everything is fine, that the parents/caregivers are fully competent, or that the family is healthy and well resourced. It means only that conditions in the family do not rise to the level of actionable maltreatment. Being "at or above baseline" in no way precludes or discourages the offer of voluntary services. Practice approaches like differential response, with which the NCFAS-G was originally tested,

accommodate shifting practice ideologies. NCFAS ratings inform these changing practice ideologies.

The reason for this definition of “Baseline/Adequate” is that it is impossible to construct a single definition of “Baseline/Adequate” that is universally applicable. States’ definitions of abuse and/or neglect vary widely. Judges behave differently across jurisdictions. Community standards also vary. Therefore, the application of the definition of “Baseline/Adequate” is dynamic, and should be determined by the agency using the scale in that agency’s legal/judicial/social environment.

Similarly, in actual practice settings, social workers and supervisors come to know what they must take to court, or what they can take to court and “win,” and what they should not take to court, sometimes in spite of the language of the law or even their own policy. Thus, another property of the 6-point rating strategy employed by the NCFASes is that guiding language defining the scale points is provided for the endpoints (“Clear Strength” and “Serious Problem”) and for “Baseline/Adequate,” but the other three scale points are left undefined. The “defining” language, or interpretation of degree of strength or problem is given over to the practitioners using the organizing framework of the NCFASes to assess their families. There is simply too much variability across practice settings, ideologies, laws, and cultures to develop universally applicable definitions; bad or misaligned definitions disable a scale in practice.

An example of this variability can occur where parental drug use is involved. Some communities and judges may view the use of any scheduled substance by parents to be *prima facie* evidence of maltreatment (usually “injurious environment,” even if no maltreatment has actually occurred). Some judges will want to see children in these situations removed from home. Agencies that fail to bring these situations before the judge risk castigation. However, in areas where drug courts have been established, and where drug treatment services are available, the judicial/social culture may permit voluntary services to these families without mandatory court involvement, or at least court approved service plans that permit the family to stay together while the parent(s) receives treatment.

Another reason that the intermediate levels of rating are not defined is that part of the philosophy of the NCFAS assessment strategy is to *encourage interaction* between workers and supervisors. This point cannot be overemphasized as it drives the research designs focusing on determination of psychometric properties, particularly reliability. The NCFAS scales rely on *intra-*, rather than *inter-rater* reliability. The factor structures of the scale have been determined, and the Cronbach’s alphas (reflecting internal consistency of scale content as applied by social workers) have been consistently high in a variety of practice settings. Within the context of assessment, service planning, and goal setting, having the organizational culture move towards agreement on the definitions of, for instance, “mild” or “moderate” is a desirable outcome. Relative movement on the scale is the intent of services, and gradual concordance with ratings should occur over time as supervisors and workers discuss the substance and reasons for particular ratings.

Because the NCFAS ratings involve subjectivity and are not intended to “diagnose” or provide a single index of overall functioning, when used as intended the NCFAS ratings should inform, but not determine, case decisions such as removal or closure. While not intended to determine those

decisions, ratings do, however, provide a structured record of the ratings that influence (or “influenced,” in the case of closure ratings) those decisions. The ratings may also inform decisions about the need for step-down or aftercare services that support reunifications or other forms of permanency.

Rater bias may be a concern when using instruments that claim to have high inter-rater reliability. However, to achieve high inter-rater reliability the workers using such a scale must learn and apply the precise definitions of the scale developer. They are, in essence, taken out of the rating process as contributors, and become mere technicians responsible for adopting the scale developers’ model and applying it proscriptively. Our experience with workers in numerous agencies across numerous jurisdictions and across different service populations (Child Protection Services, Juvenile Justice, Mental Health) is that they do not like being viewed as technicians, they resent being forced to use some instruments that provide to them information that they do not find particularly useful. Similarly, they resent that they cannot influence diagnostic results with what they know about the child or family, but which the scale developer did not anticipate, or did not consider to be important or relevant to the domains being measured. Workers tend to like, rather than dislike, using the NCFAS because it embraces their judgment. It is designed to help them fully engage families, and to respect their training, wisdom, and competence. At the same time, when used as intended, it also requires that they be willing to defend their ratings to supervisors, and to consider the opinions of others in their practice settings. These are good things to strive for in practice environments, particularly collaborative practice environments where the social workers are partners with caregivers, rather than commanders of caregivers.

These points are important when considering rater bias or rater inflation. While rater inflation may be a concern, our experience is that it is not a serious problem. Scales with high inter-rater reliability (i.e., with diagnostic capabilities) tend not to be as capable of detecting change during service delivery. This is true because the rating or scoring strategy is that of the scale developer, not the worker, and because the standardized instruments may not be focusing precisely on the variables of interest. On the contrary, the flexibility of the NCFAS rating strategy permits workers to precisely apply what they observe to the ratings both at intake and closure, and at intermediate times when an updated assessment is desirable. The tendency to inflate those ratings is mitigated in practice by the requirement to defend the ratings. If a worker observes that the family has not made sufficient progress during service (either the family cannot or will not change or improve) and the worker finds it necessary to recommend removal, or to not reunify, or to require additional services, that worker is not likely to inflate the ratings. After all, the ratings reflect the reasons for the decisions or recommendations. However, if progress has been observed, and in the judgment of the worker and supervisor that progress has been sufficient to warrant a change in case status, then the magnitude of the change in the rating is a matter of judgment and agreement by worker and supervisor. Or, if worker and supervisor do not agree, then the worker (or supervisor!) succeeds or fails in relation to the family’s success and future trajectory.

So, having addressed the issues of scale development and intent, use in case practice, and rating strategies, how do these factors relate to reliability and validity? Again, what is the intent? Because the NCFAS scales, since their inception, have relied on *intra*-rater reliability, it is highly

desirable to observe high internal consistency of the use of items within each domain by social workers using the scale in actual case practice. Internal consistency is measured using Cronbach's alpha, and any computation of Cronbach's alpha reflects the internal consistency achieved by a group of social workers using the scale in a particular practice setting with a particular population. Reliability is not an inherent property of the scale itself (although scale construction is very important thereto), which automatically transcends various workers, practice models, and treatment populations. Prudent use of any instrument in any piece of research or program evaluation dictates that Cronbach's alpha be calculated as part of the basic analysis of the research data. Even the most "reliable" instruments can be used "unreliably" by untrained workers or by poor procedures.

Technically, because Cronbach's alpha is based on inter-correlations within the associated domain items, the statistic ranges from "0" (very poor or nonexistent internal consistency) to 1.0 (perfect internal consistency). By convention and psychometric theory (See, e.g., Nunnally, 1978; DeVellis, 1991) alphas of 0.4 are considered acceptable during scale development, alphas of 0.7 or higher are considered to be acceptable for clinical application, alphas of 0.8 or higher are considered to be high, and alphas of 0.9 or above are considered to be very high. In the presentation of research/evaluation findings that follows, Cronbach's alpha is the statistic defining reliability.

Inter-rater reliability (in the case of the NCFAS Scales) refers to the extent to which different workers assign the same ratings (or draw the same conclusions) about the same family. This type of reliability is difficult to determine in practice because it is rare for more than one worker to rate the family simultaneously. However, a 2012 study conducted in Alaska and 3 other national and international sites, suggests that inter-rater reliability of the NCFAS Scales is quite good. In that study, 55 social workers rated the same case scenario at intake and closure, using an Alaska Native family case. Between 70% to 90% of them assigned ratings within one scale increment above and below the overall group average rating. These ratings reflect very good alignment of the social workers ratings with the standard reference ratings for the case study that were established by the authors of the case scenario. Not only were the group mean ratings essentially equivalent to the case standard ratings on almost every domain, but the standard deviations for the group mean ratings was less than 1 in 14 of 20 comparisons, suggesting very little dispersion of ratings about the mean.

Pursuit of validity is less "straight forward" than reliability because there are several types of validity that can be pursued, some of which are subjective (face validity), and some of which require specific statistical computation (e.g., predictive validity, concurrent validity). Others (e.g., construct validity) require the simultaneous administration of other, validated (and reliable!) instruments and computation of correlations with the constructs (domains) of the comparison scale. Establishing construct validity requires expensive and time consuming research, not normally achievable during applied program evaluations. At various times during the development of the NCFAS scales, each of these types of validity has been determined, and results of the validation efforts are also presented in the presentation of research/evaluation studies of the NCFAS scales.

Reliability and Validity of the NCFAS Scales

The NCFAS scales share many common elements (indeed, “whole cloth” combinations in some cases), and the research, evaluation, and discussion of the scales makes it difficult in some cases to isolate one scale from another. In the following presentation of summaries of studies, the specific version of the scale used in the study is identified, as are the types of reliability and validity pursued. The NCFAS is the original scale, comprising the domains of Environment, Parental Capabilities, Family Interactions, Family Safety, and Child Well-Being. The NCFAS-R (for Reunification) followed, and is the original NCFAS with two new domains added: Ambivalence and Readiness for Reunification. The NCFAS-G was developed in order to expand the scales’ applicability not only to the high-risk families but also to lower risk families such as those served in differential response practice settings. All of the original NCFAS domains are retained, and a few of the original NCFAS scale items were used to anchor three new domains: Social/Community Life, Self-Sufficiency, and Family Health. The NCFAS-G+R resulted from adding the Ambivalence and Readiness for Reunification domains (from the “R” version of the NCFAS) to the NCFAS-G. Regardless of risk level, the NCFAS-G is applicable to all families, and the “R” domains can be used with those families who have had a child removed due to child maltreatment. Finally, the stand-alone Trauma and Post-Trauma Well-Being domains were tested in several different practice settings using either the NCFAS-G or NCFAS-G+R (the study sample included both placement prevention and reunification cases).

The first article, below, resulted from the seminal reliability and validity (R&V) study of the original NCFAS, and is the foundation of all future studies.

Article/Study Title	The reliability and validity of the North Carolina Family Assessment Scale
Author(s)	Reed-Ashcraft, K.B., Kirk, R.S., & Fraser, M.W. (2001)
Citation	<i>Research on Social Work Practice, 11</i> (4), July, 503–520
Scale involved	NCFAS
Type of R&V Tested	Cronbach’s alpha for reliability; construct validity using Magura & Moses’ Child Well-Being Scale, Hudson’s Index of Family Relations, and McCubbin’s Family Inventory of Resources for Management
Sample Characteristics and Size	419 families served through intensive FPS programs in North Carolina during a 10-month period from September 1996 to June 1997; high-risk families with placement imminent
Results	Factor analysis resulted in some “tweaking” of scale items, and factor naming, resulting in a reconstruction of the scale to fit the best statistical model. Final Cronbach’s alphas on Version 2.0 ranged from .71 to .94 Hypotheses for 5 of the 6 comparisons to other validated instruments were

	supported, robust and statistically significant. The 6th hypothesis was rejected and the factor dropped from V2.0
Limitations	Although there was broad racial/ethnic representation among both families and staff conducting the assessments, and children of all ages were among the study population, all families were high-risk families who had been substantiated for child maltreatment. Therefore, generalization to lower-risk families is not possible from this study.
Other Notes	The general approach of broad-based family functioning assessments that extend well beyond the issues presented during the child protection reports was affirmed.

Following the success of initial scale development, additional research was conducted using a large retrospective sample to address the issue of predictive validity. This study progressed beyond the theoretical issues of family assessment to its application in a state-wide practice model: Intensive Family Preservation Services. The study population was similarly high-risk as compared to the initial scale development study. The dependent variable was placement or non-placement related to assessment ratings from the NCFAS V2.0, and future placement based on closure ratings. This study is summarized below.

Article/Study Title	Advances in the Reliability and Validity of the North Carolina Family Assessment Scale
Authors(s)	Kirk, R.S., Kim, M.M., & Griffith, D.P. (2005)
Citation	<i>Journal of Human Behavior in the Social Environment</i> , 11 (3&4), 157–176
Scale Involved	NCFAS V2.0
Type of R&V Tested	Cronbach’s alpha for reliability; concurrent validity using placement/non-placement at the close of services, and predictive validity using future placement within one year of services
Sample Characteristics and Size	1,279 families throughout North Carolina that had received Intensive Family Preservation Services during the preceding 4 years. All were high-risk families with one or more children on the verge of placement at the time services began. A broad range of racial and ethnic identification was represented.

Results	Cronbach's alphas indicate that NCFAS V2.0 is very reliable both at Intake and at Closure, with alphas at intake ranging from .72 to .90, and from .79 to .91 at closure. These results indicate that the scale items contribute substantially to the measured constructs and that the internal consistency is high. Both concurrent and predictive validity of the NCFAS were established in relation to placement prevention and future placements of children. Ratings in the strength ranges of the NCFAS at closure were significantly associated with non-placement at the end of service and at one year for each domain, while ratings in the problem ranges were significantly associated with placement. Strong, significant associations were found with non-placement at the end of service and within one year for families rated at or above Baseline/Adequate at closure.
Other notes	Reliability findings strongly confirmatory of previous study. Validity findings strongly support definition of Baseline/Adequate.

Following the success of the IFPS model for dealing with high-risk families, and the success of the development efforts relating to the NCFAS, some IFPS providers began to apply IFPS practice models to family reunification cases. The National Family Preservation Network was a principal player in this movement, and secured funding from The Packard Foundation to assist both with model development and to adapt the NCFAS to support family reunification efforts. The following evaluation summary presents the findings of a 3-year effort to promote both model development and NCFAS adaptation, ultimately culminating with release of the NCFAS-R, which included the new reunification-specific domains. The report cited below is the third and final project report to The Packard Foundation on the use of their grant funds to achieve the intended purposes of the grant.

Article/Study Title	Tailoring Intensive Family Preservation Services for Family Reunification Cases: Final Results of Field Testing and Validation of the North Carolina Family Assessment Scale for Reunification
Authors(s)	Kirk, R.S. (2002)
Citation	National Family Preservation Network Project Reports and Monographs; Packard Foundation Final Project Report, http://www.nfpn.org/Portals/0/Documents/ncfas-r_research_report.pdf
Scale Involved	NCFAS-R V2.0

Type of R&V Tested	Cronbach’s alpha for reliability; concurrent validity, using domain ratings and placement outcomes at the end of the service period. Three different methods of demonstrating concurrent validity were explored: (a) establishing the relationship between treatment outcomes and overall aggregate domain ratings at intake and closure; (b) establishing the relationship between treatment outcomes and improvement or deterioration on domain ratings during intervention; and (c) establishing the relationship between treatment outcomes and being at or above the baseline threshold rating at intake and closure.
Sample Characteristics and Size	170 high-risk families from which one or more children had been removed by the courts for child maltreatment. Racial distribution of primary caregivers: 40% White, 49% Black, 11% Other. Families were served by one of three participating agencies from three different States.
Results	Cronbach’s alphas for all domains, including the new domains added to address issues unique to reunification, ranged at intake from 0.82 to 0.92, and at closure ranged from 0.81 to 0.94. For validity, each of the hypothesized relationships was demonstrated to be statistically significant and robust. Thus, the NCFAS-R was found to be both reliable and valid for use in IFPS-based reunification interventions.
Limitations	Although all families in the study were definitely “reunification” families (i.e., one or more children had been removed from home), there were differences among the reunification service models that may have affected outcomes or contributed variance in unknown ways.
Other notes	In this study, the successful reunification rate was 76%. At that time, the national performance standard set by the Children’s Bureau for reunification was 74%, and at the time less than 10% of states were meeting the criterion. The NCFAS-R likely contributed substantially to assisting these programs to meet the federal requirement, even when their host states were not.

During this same time period, several states were experimenting with a new public practice approach that did not require evidentiary investigations to be conducted following allegations of child maltreatment, if the risk to the child(ren) was deemed to be low at the time that the report was received. Rather than investigating these families for child maltreatment, they were offered voluntary services without the stigma of being labeled as abusive or neglectful of their children. These similar models became known as “differential response” models, and they were found to be successful at keeping many children out of the child welfare system and generally were associated with both the absence of child removal and a reduced likelihood of future child maltreatment. All of these are desirable outcomes for public child welfare, and a large California County approached the National Family Preservation Network and requested that it assist them in developing an assessment model to accompany the County’s implementation of a differential response practice model. They were familiar with the NCFAS, used for high-risk families, and wanted to use the same approach for potential differential response families who may have a variety of issues acting against their ability to properly parent their children, including general family or child health issues, economic self—sufficiency, and community supports. Since all of

the domains on the original NCFAS were also relevant, the NCFAS was modified for differential response by adding three domains relating to these “general” concerns that all families have (health, self-sufficiency, community connectedness) the absence of any of which foster circumstances that may lead to child maltreatment. This initiative was the birth of the NCFAS-G, for “general services.” The study below presents the R&V study of the implementation of the NCFAS-G.

Article/Study Title	Development and Testing of a Family Assessment Scale for Use in Child Welfare Practice Settings Using Differential Response
Authors(s)	Kirk, R.S (2008)
Citation	<i>Protecting Children</i> , 23 (1&2), 71-87
Scale Involved	NCFAS-G
Type of R&V Tested	Cronbach’s alpha for reliability; Concurrent validity with respect to family stability, absence of child removal
Sample Characteristics and Size	123 families (and 252 children) representing a wide range of racial and ethnic identities.
Results	Cronbach’s alphas ranged from 0.83 to 0.95 at intake, and from 0.88 to 0.95 at closure, suggesting very good internal consistency with respect to the DR workers assessing low-moderate risk families in the DR practice environment. For validity, ratings assigned by workers were consistent with the low-moderate risk population targeted by Differential Response program (96% of families were low to moderate risk); there were modest population shifts towards the strengths range of ratings associated with receipt of services; there were no known reports of child maltreatment among the families served; there was a statistically significant increase in the number of families rated at or above Baseline/Adequate at closure, compared to intake.
Limitations	The county in which the NCFAS-G was tested for R&V was implementing DR at the same time. The NCFAS-G performed as expected, but the DR model experienced implementation issues, and exposed several service gaps and family engagement issues that likely attenuated treatment outcomes.

Other Notes	<p>Because a few scale items from the original NCFAS domains were used as anchors for the new domains comprising the “G” version of the NCFAS, it was important to assure that the internal consistency, as reflected in high alphas, was retained, e.g., item on “sufficient household income” lifted from NCFAS Environment to anchor NCFAS-G Self Sufficiency domain, with new scale items added. High alphas confirmed retention of good internal consistency.</p> <p>The NCFAS-G, like the original NCFAS and NCFAS-R, was created to meet a need of the practice community that had been identified in the literature and among early purveyors of Differential Response: tools and methods to assess families broadly and collaboratively rather than forensically.</p>
-------------	--

Up to this point, the studies summarized were conducted by the scale developers with a combination of state and foundation funding, and related to establishing the R&V of the scale forms across different service populations. The National Family Preservation Network also engages in service model development and efficacy testing. Once the R&V of the NCFAS scales had been established, the scales were used as part of the design of those model-testing studies. One such study compared the efficacy of Intensive Family Preservation Services for both placement prevention and reunification but focusing on the ages of youths/children at the time of entry into the child welfare system. This endeavor was a collaborative effort between the National Family Preservation Network and the National Alliance to End Homelessness. As suggested earlier in this paper, R&V should always be part of any research endeavor using standardized instruments, and those analyses were a part of this study, which is summarized below.

Article/Study Title	A Comparative Study of Older-Youth Families and Younger-Youth Families in Placement Prevention and Reunification Cases, Using Traditional Child Protection Services Measures and NCFAS and NCFAS-R Assessment Data
Authors(s)	Kirk, R.S. (2010)
Citation	National Family Preservation Network Project Reports and Monographs; Older Youth Study; http://www.nfnp.org/preservation/research-on-older-youth
Scale Involved	NCFAS and NCFAS-R
Type of R&V Tested	Cronbach’s alpha for reliability (multiple states, multiple programs, numerous workers, varying treatment populations). Validity: worker assigned ratings as a function of placement/non-placement and reunification; worker assigned ratings as a function of type of child maltreatment.

<p>Sample Characteristics and Size</p>	<p>Placement prevention youths sample: N = 648, with 468 youths 0 – 11 years of age (72%) and 180 youths 12 – 17 years of age (28%). Across both groups, about 60% to 66% were White, 10% to 18% Black, and the remainder distributed across Asian, American Indian, and Other categories. Overall, the sample was 50/50 male/female. There are interesting differences between the older and younger sample members with respect to poverty, gender, child’s relationship to caregiver(s), and living arrangements at intake and closure, which are available in the full report, but which are not central to the issues of R&V of the NCFAS scales as used in the study.</p> <p>Reunification youths sample: N = 185, with 147 youths 0 – 11 years of age (79%) and 38 youths 12 – 17 years of age (21%). Sample is predominantly White (about 55% overall, about 10% Black, and about 15% multi-racial, the balance (20%) being Asian, American Indian, or Other. Like the younger youth sample there are interesting differences between the older and younger reunification samples on race, gender, and living arrangement (details available in full report).</p>
<p>Results</p>	<p>Reliability: For placement prevention families (NCFAS) the Cronbach’s alphas ranged from .61 to .89 (all but one was above 0.76; the outlier rose to 0.81 at closure) and all closure ranged from 0.76 to 0.90) suggesting good internal consistency across the various states, programs, workers and treatment populations. For the reunification families (NCFAS-R) Cronbach’s alphas ranged from 0.74 to 0.91 at intake, and from 0.84 to 0.92 at closure, suggesting high internal consistency across the various states, programs, workers and treatment populations.</p> <p>Validity: NCFAS and NCFAS-R ratings closely tracked maltreatment types and other agency concerns recorded independently on other forms. Families with older youths were significantly more seriously rated on Family Interactions at Intake, but not at Closure. The same was true to an even greater degree for Child Well-Being, with the difference maintaining at Closure. Among the reunification sample, families with older youths were more seriously rated on Child Well-Being, and families with younger youths made more progress on Well-Being during the service period than did families with older youths. These NCFAS and NCFAS-R ratings were consistent with treatment outcomes and achievement of service plans.</p>
<p>Limitations</p>	<p>The study relied on a retrospective sample of convenience which is positive in that it reduces the likelihood of reactivity associated with prospective research, but the combined samples from the various states and programs may have unknown selection bias.</p>
<p>Other Notes</p>	<p>The states, programs, service types, and treatment populations were diverse, supporting broad generalization of R&V of NCFAS scales for both placement prevention and reunification.</p>

Researchers and program evaluators outside of NFPN have also used the NCFAS scales in program development and model testing. One such endeavor that contributed R&V information involved a State-funded project to test the use of school-based family team meetings (using the Australia/New Zealand family team decision-making model as its basis). The focus of the initiative was early identification of low/moderate risk children in the school setting and engagement of their families in voluntary services. The NCFAS-G was selected as the family assessment instrument. The study is summarized below.

Article/Study Title	School-Based Child and Family Team Project Report to the North Carolina Department of Public Instruction
Authors(s)	Pennel, J. (2008)
Citation	Center for Family and Community Engagement, North Carolina State University, Campus Box 8622, Raleigh, NC 27695-8622 United States of America
Scale Involved	NCFAS-G
Type of R&V Tested	Cronbach's alpha for reliability, concurrent validity in relation to disciplinary referrals, absences, suspensions, school performance.
Sample Characteristics and Size	N = 16 families, focus child age range 4-17 (of which 56% were 12 or older), 64% were multi-child families, 81% were single-parent families. Child races: 75% White, 19% Black, 6% American Indian. Gender: 56% male
Results	<p>Reliability: Alphas at intake ranged from 0.58 to 0.93, and at closure they ranged from 0.89 to 0.98. Somewhat lower than expected intake reliability attributed to workers not embracing the idea (conveyed in training) of conducting sufficiently lengthy home visits to gather information needed to assign ratings. After many weeks of getting to know family through CFT process, reliability is very high.</p> <p>Validity: Worker assigned ratings of family functioning in keeping with low/moderate risk population, and also school-focused (specific scale items among the domains), perhaps to the exclusion of non-school related items, also associated with lower intake reliability. Pre/post NCFAS-G ratings track data from other sources relating to alleviating family stress to some degree, assisting caregivers to improve their living situations and their abilities to effectively parent their children (or care for the children for whom they are responsible), and helping children adjust to become more successful and less disruptive in school.</p>
Limitations	The small sample size of families serviced during this pilot program implementation period precludes statistical testing of difference scores for statistical significance.

Other Notes	<p>Reliability data for intake underscores the necessity of training on the use of the NCFAS-G, and embracing the need for adequate home visits and gathering of information across the various domains to competently and confidently assign ratings.</p> <p>See also: Parcel, T.L. & Pennell, J. (2012), “Child and Family Teams Building Social Capital for At-Risk Students,” <i>Journal of Sociology and Social Welfare</i>.</p>
-------------	---

Although reliability is not always tested in research endeavors, some studies can contribute substantially to discussions of validity. One such very recent study is summarized below. The NCFAS was used in conjunction with several other standardized instruments, and the between-instrument assessment comparisons are very informative with respect to construct validity.

Article/Study Title	Supportive Housing for Families in Child Welfare: Client Characteristics and Risk Factors at Intake
Author(s)	Randall, K. (2012)
Citation	University of Connecticut, DigitalCommons@UConn.edu
Scale Involved	NCFAS
Type of R&V Tested	Reliability not tested. Validity: Construct by multiple instrument comparisons on the study population, including the Simple Screening Instrument for Alcohol and Other Drugs (SS-AOD); Parenting Stress Index-Short Form (PSI-SF); Brief Symptom Inventory (BSI)
Sample Characteristics and Size	80 clients referred to Supportive Housing for Families, a publicly funded service working in collaboration with Connecticut’s Department of Children and Families; 87% female, 50% White, 26% Black, 21% Latino/Hispanic; mean age = 31 years; 54% high school or higher grade equivalent; 61% unemployed. As a group there were multiple symptomologies: mental health, substance abuse, stress related to parenting/child care
Results	High scores on SSI-AOD tracked increasing problem ratings on NCFAS/Parental Capabilities (i.e., drug use interferes with parenting); NCFAS/Child Well-Being tracked PSI-SF subscales of parent/child dysfunction, difficult child, and total stress. NCFAS/Environment and NCFAS Family Interactions tracked Parental Stress on PSI-SF. NCFAS/Parental Capabilities tracked all subscales of BSI.
Limitations	Unknown
Other Notes	This study offers independent support for the constructs comprising the NCFAS. The academic paper provides detailed discussions of between-scale relationships. Study author also noted the uniqueness of the NCFAS among the instruments employed with regard to its ability to assess strengths.

The preceding study was selected for inclusion due to its direct bearing on the issue of construct validity. Another study from the same laboratory involved the model testing of a child welfare intervention relating to housing and homelessness and its contribution to problems relating to family functioning and child-well-being. Again, the study did not focus on the reliability and validity of the NCFAS, per se, but used the NCFAS and NCFAS-R to explore the impacts of a housing subsidy program for poor families in the child welfare system that were experiencing homelessness. As such, the study population is somewhat narrower than the general child welfare population, regardless of initial risk, but the authors cite numerous concomitant issues that they (and the research literature cited in the article) believed to be impacting these families. They point out that homelessness resulting from structural issues such as few job opportunities, and individual factors, such as limited financial resources or disabilities affecting working capabilities, can impact the family ecosystem and can exacerbate health, education, and social problems. All of these stressors place children in these families at heightened risk. Many of these issues are addressed in the NCFAS and NCFAS-R domains, and the instruments were chosen to be used as part of the assessment arsenal for the study. The study is summarized below, as it relates to informing this paper’s focus on validity (to our knowledge, reliability as reflected in Cronbach’s alpha was not examined, although the workforce is known to be very experienced in using the NCFAS scales).

Article/Study Title	Supportive Housing for Families in Child Welfare: Client Characteristics and Their Outcomes at Discharge
Author(s)	Farrell, A.F., Britner, P.A., Guzzardo, M., & Goodrich, S. (2010)
Citation	<i>Child and Youth Services Review</i> , 32, 145-154
Scale Involved	NCFAS and NCFAS-R
Type of R&V Tested	Reliability not tested, to our knowledge. Validity: Concurrent (ratings at intake tracking the issues hypothesized to be prevalent among the target population; ratings at closure tracking and informing successful housing outcomes); Predictive (ratings at intake predicting successful housing outcomes). Target outcomes included: housing (improved or not, temporary or permanent), access to health care (improved or not), employment (unemployed, part-time, full-time), type of exit (successful/unsuccessful).
Sample Characteristics and Size	From a client population of 1,327 families, NCFAS (or NCFAS-R) intake ratings available for 986 families. Among families, heads of household were: 93% female, 75% single, 39% White, 28% Black, 30% Latino, 3% American Indian; age range 16–64, mean age = 32; mean education = 10.9 years. The 3,797 children in the families were 52% male; mean age = 10.1 years; race/ethnicity of closely tracked heads of household.

Results	At intake, 49% of families were rated in the problem range for Environment (which subsumes housing and employment issues), and 37% were rated at baseline/adequate. The majority of families were baseline/adequate on other NCFAS domains. The majority of exits (73%) were labeled as successful, with the majority being in permanent housing, and with significant increases in numbers employed. These improvements were tracked by significant improvements in NCFAS domain ratings, particularly so for Environment, the key domain associated with housing and employment. Paired t-tests for 541 families for whom both intake and exit NCFAS ratings were available showed significant improvements on all domains. Also, intake domain ratings (except child well-being) predicted successful from unsuccessful exits as a function of being at or above Baseline/Adequate at intake.
Limitations	The researchers started using the NCFAS and NCFAS-R on all families (especially those within the child welfare system) after some cases were already open, so that among the cases included in the analyses, the time interval between initial intake and subsequent ratings was not consistent. Because some families may have already begun to receive services prior to the initial NCFAS rating process, any improvements in those families would likely have attenuated the magnitude of difference scores.
Other Notes	In keeping with the philosophy of the developers and the training provided as part of the acquisition rights for use of the scales, the researchers did not use the predictive validity of the NCFAS intake scores to impute “cut score” capability or use of the findings to select cases more likely to be successful. Rather, they suggest that the findings relating to prediction of success at intake be used to identify those families likely needing extra or more intensive services and engagement efforts.

Another study conducted in the fall of 2012 focused on both the cultural competence of the NCFAS Scales (i.e., their applicability to non-dominant races, cultures and ethnicities), and the psychometric property of inter-rater reliability. The National Family Preservation Network, sponsor of the study, teamed with child welfare workers in Alaska to examine the applicability (validity) of the Scales for working with Alaska Native families. NFPN and a group of Alaskan social workers and administrators (about evenly divided by Native and non-Native participants; the Native participants representing 4 Tribes) provided input for a 2-part case scenario based on a typical Alaska Native family involved with child protection services. The case scenario, ultimately authored by NFPN, was reviewed by members of the Alaska group, who affirmed that it met the goals of representation of Alaska Native families.

A group of 55 volunteer social workers was recruited to review the case scenario and to assess the family at intake and closure using either the NCFAS-R or the NCFAS-G+R. The volunteer workers were from child welfare agencies in three different states and one international agency based outside the US, but familiar with serving indigenous populations. The groups were about equal in size, and represented a broad range of experience, age, education, gender and race/ethnicity.

Although the primary focus of the study was the cultural competence of the scale, NFPN took advantage of the study’s design to examine inter-rater reliability within this group of volunteer workers, using the standardized, 2-part case scenario. Results affirmed both the applicability of the Scales for Alaska Native families, and provided very good inter-rater reliability statistics. The findings support a growing body of knowledge on the reliability and validity of the NCFAS Scales.

Article/Study Title	An examination of the cultural competence and inter-rater reliability of the North Carolina Family Assessment Scales
Author(s)	Kirk, R.S and Martens, P. (2012)
Citation	An examination of the cultural competence and inter-rater reliability of the North Carolina Family Assessment Scales. Available at http://www.nfpn.org/articles/cultural-competence-report
Scale Involved	NCFAS-R and NCFAS-G+R
Type of R&V Tested	Cronbach’s Alpha for internal consistency (reliability); Summed proportions of workers rating within one scale point of scale standard for standardized case, for inter-rater reliability. Validity as reflected in the scales applicability to racial minority family assessment, as measured by differences between raters’ assigned ratings and case standard for a specific racial minority case scenario.
Sample Characteristics and Size	55 social workers who varied by gender, race, age, experience, education, and location of case practice (3 states and 1 international country)
Results	Computed reliability statistics (Cronbach’s Alphas) were adequate to high at intake and high to very high at closure indicating good reliability as measured by internal consistency. Inter-rater reliability appears to be very good: in all but 4 of 20 domain measures 70% to 90% of 55 social workers rating the same case scenario at intake and closure (using an Alaska Native family case) assigned ratings within one scale increment above and below the overall group average rating. In 16 of 20 domain measures the group mean assigned rating was within ½ scale increment of the case standard rating, and with standard deviations of less than 1.0, suggesting little dispersion of ratings about the mean. There were no meaningful differences in the way workers assigned ratings as a function of demographics (age, race, gender, experience or location of practice).
Limitations	This study specifically examined the cultural competence of the NCFAS Scales by having multiple and varied social workers assess an Alaska Native family case scenario, and although the findings appear to be compelling with respect to two types of reliability and cross-cultural validity, only one case was reviewed by the participating workers.

Other Notes	These findings contribute to the growing body of evidence that the NCFAS Scales are applicable across various racial, cultural and ethnically identified groups, and the Scales exhibit good inter-rater reliability.
-------------	---

In 2014, the National Family Preservation Network (NFPN) published findings from a study designed to examine service outcomes for both placement prevention and reunification families as a function of a variety of demographic and service variables. (Kirk & Martens, 2014). The exploratory study compared intake and closure family functioning assessment ratings on the NCFASes and a variety of service variables and service outcomes, principally successful prevention of placement or family reunification. Additional variables were examined to gather descriptive information on their possible relevance to influencing service outcomes. These included: family (primarily caregiver) engagement, family meetings, concrete services, step down services, father involvement, domestic violence, and various mental health issues affecting caregivers (substance abuse, general mental health, depression).

This study was not designed specifically to examine reliability or validity of the NCFASes, per se. However, the reliability of the scale in these practice settings was tested, is it always is when NFPN uses the scales for research or program development purposes. The following description of the study presents the findings of the reliability analysis, as well as some of the findings relating to the family preservation services model with both placement prevention and reunification cases. The report is lengthy, and the reader is encouraged to access the source document for detailed explanation of the multiple research questions explored therein.

Author(s)	Kirk, R.S and Martens, P. (2014)
Citation	Family assessment, family functioning and caregiver engagement in family preservation and reunification programs, and the relation of these and other factors to reunification service outcomes. Available at: http://www.nfpn.org/reunification/reunification-research
Scale Involved	NCFAS-G, NCFAS-R and NCFAS-G+R
Type of R&V Tested	Cronbach's Alpha for internal consistency (reliability);

<p>Sample Characteristics and Size</p>	<p>Four family preservation services programs offering placement prevention and/or reunification services in 4 different states participated in the study. Three additional programs, including 2 programs operating in foreign countries, contributed data relating only to placement prevention programs, and contributed survey data exploring client engagement of caregivers in their placement prevention and differential response programs. Data represented families across the racial/ethnic/international continuum, and the distribution of demographics varied across analyses due to multiple and varied research questions. Sample sizes ranged from 64 families to 184 families depending on the data being analyzed. More complete descriptions of sample characteristics are presented on an analysis by analysis basis in the full body of the report.</p>
<p>Results</p>	<p>Reliability: for placement prevention families where workers were completing the NCFAS-G, Cronbach’s alphas were consistently above .800 and frequently above .900 at both intake and closure. For reunification families, with 2 exceptions at intake only and one exception at closure, Cronbach alphas were consistently above .830 and most frequently above .900.</p> <p>Findings relating to the studies’ research questions include the finding that there was virtually no difference in placement outcomes for either placement prevention or reunification families purely as a function of demographic variables. This is true even for families in which domestic violence, alcohol or other drug use, mental illness (including depression) was evident. Step down services for families who needed them significantly improved the durability of reunifications, and successful engagement with the caregiver and alignment of views of the caregiver and family worker (as determined by independent assessments thereof) were associated with higher reunification rates and a greater likelihood of service completion than when trust was not established early in the case. There is some evidence, requiring further exploration, that biological fathers’ involvement in the lives of their children during reunifications can have a significant impact on the durability of those reunifications although it did not appear to have an impact on initial reunification rates.</p>
<p>Limitations</p>	<p>This study was exploratory and descriptive. The findings are qualitative and each finding will require additional research to verify or continue the line of investigation and inquiry.</p>
<p>Other Notes</p>	<p>These findings contribute to the growing body of evidence that the NCFAS Scales are applicable across various racial, cultural and ethnically identified groups, and the Scales continue to exhibit good to very good reliability statistics in child welfare practice settings both in the United States and abroad. The continued robust reliability of the NCFASes lends credibility to the findings in studies where it is used to provide assessments of family functioning which are often used as the dependent variables in the studies.</p>

In 2015 the National Family Preservation Network released a report on the development and field testing of 2 new domains which can be used in conjunction with various versions of the

NCFAS scales. (agencies participating in the field test were using the NCFAS-G or the NCFAS-G+R). The 2 new domains are Trauma and Post-Trauma Well-Being. Interest in assessment of trauma and posttraumatic well-being in child welfare has grown out of recent research findings from the Centers for Disease Control, specifically from the Adverse Childhood Experiences Survey (ACES), and other sources.

Trauma is now known to affect psycho-social, emotional, cognitive, and even physical development of children. Trauma is also now known to affect parents’ or caregivers’ ability to fulfill their role as parent, protector, and nurturer of their children. It is likely that many families in the child welfare system have histories of trauma, either for the parent/caregiver, the children, or both. Presence of trauma symptomology can adversely impact a family’s ability to engage in services and improve family functioning. However social workers and contracted providers serving families in the child welfare system have rarely been trained in the assessment of trauma, nor have there been practice-friendly tools available to assist them. The development of the new trauma domains, to be used as companions to the broader family assessments conducted using various versions of the NCFAS scales, is intended to encourage social workers to embrace the process of assessing for trauma, to assist them in planning services to ameliorate trauma symptomology, and to reassess families at the end of services with respect to the amelioration of trauma symptomology and improvements in well-being in areas of development and social/psychological functioning following services for trauma.

The 2015 study described below presents the results of the field test of the new domains, including their initial reliability testing when used by social workers in actual case practice settings, with families on their active caseloads. Three sites participated in the field study, and the participating workers provided data on 170 families and more than 350 children.

Results of the study affirmed very good reliability statistics (as measured by Cronbach’s Alpha) for the new domains, reconfirmed excellent reliability statistics for the NCFAS-G and NCFAS-G+R in the same case practice settings, supported hypothesized correlational relationships between scale items on the 2 new domains (trauma and post-trauma well-being) and various domain ratings on the pre-existing domains of the NCFAS scales, suggesting good convergent validity between the new domains and the NCFAS scales..

Article/Study Title	Field Test and Reliability Analyses of Trauma and Post-Trauma Well-Being Domains of the <i>North Carolina Family Assessment Scale for General and Reunification Services</i>
Author(s)	Kirk, R.S and Martens, P. (2015)
Citation	Field Test and Reliability Analyses of Trauma and Post-Trauma Well-Being Domains of the <i>North Carolina Family Assessment Scale for General and Reunification Services</i> Available at: http://www.nfpa.org/articles/trauma-report
Scale Involved	Trauma and Post-Trauma Well-Being Domains for use with the NCFAS-G and NCFAS-G+R

Type of R&V Tested	Cronbach's Alpha for internal consistency (reliability) of both the NCFASes used in the study and the new domains being tested (Trauma and Post-Trauma Well-Being). Convergent validity is exhibited in the form of correlations between the NCFAS-G and NCFAS-G+R domain ratings and ratings on Trauma and Post-Trauma Well-Being scale item ratings.
Sample Characteristics and Size	Over a six-month period in 2014 social workers from 3 sites with family preservation services programs contributed data on 170 families and 352 children. Families were assessed at intake and closure with the NCFAS-G or NCFAS-G+R as well as the new Trauma Domain and Post-Trauma Well-Being Domain, when appropriate. The sample was 66% white, 29% African-American, and 5% other racial identities. Children were approximately equally distributed across the age ranges of 0-2, 3-6, 7-12, and 13-18. About one third of the cases involved reunification, and two thirds were placement prevention cases.

Results	<p>The NCFAS scales continued to show high to very high reliability in the child welfare environment with Cronbach's Alpha for the NCFAS domains ranging from .861 to .964 at intake and from .875 to .957 at closure.</p> <p>Reliability for the new Trauma and Post-Trauma Well-Being domains was also high to very high, with Cronbach's Alpha for Trauma = .811, and for Post-Trauma Well-Being = .905.</p> <p>With respect to validity (all noted correlations are statistically significant) at intake: parental capability was correlated with parental trauma; family interactions were correlated with parental trauma; family safety was correlated with traumatic sexual abuse, physical abuse, and neglect; child well-being was correlated with traumatic sexual abuse, physical abuse, neglect, and emotional/psychological abuse; social and community health was correlated with parental trauma; and family health was correlated with traumatic sexual abuse, physical abuse, neglect, emotional/psychological abuse, and parental trauma. The strength of the correlations suggests that while the trauma information gathered at intake is consistent with assessment of overall family functioning, both positive and negative, the new information provided by the Trauma Domain is uniquely useful, and not duplicative of existing NCFAS domains. At closure: overall parental capabilities was correlated with posttraumatic parent support for child, and parent well-being; family interactions was correlated with posttraumatic parent support for child, and parent well-being; family safety was correlated with posttraumatic parent support for child, and parent caregiver well-being; child well-being was correlated with posttraumatic cognitive/physical well-being of child, emotional/psychological well-being of child, and social functioning of child; social and community life was correlated with posttraumatic parent support for child, and parent well-being; and family health was correlated with posttraumatic cognitive/physical well-being of child, emotional /psychological well-being of child, and parent well-being. Overall the results suggest that the new Trauma and Post-Trauma Well-Being domains are reliable in case practice; traumatic histories/symptomology can affect overall family functioning at intake; and that post-trauma well-being of children and parents is associated with improved family functioning at closure; all of which supports the practice of assessing for trauma as part of a comprehensive overall family assessment at intake and assessing for post-trauma well-being at case closure.</p>
Limitations	<p>These initial findings on validity are based on correlations between the trauma and post-trauma well-being scale items and associated NCFAS-G+R domain ratings. While they were positive and occurred as hypothesized, correlations do not imply causality, so additional examination focusing on validity is recommended. Furthermore it is not known to what degree the available services across the 3 sites may have varied, nor the degree to which social workers in the agencies may have had training on the assessment of trauma or post trauma well-being, other than the training they received as part of their introduction to using the 2 new domains being field tested.</p>

The studies reviewed above illustrate the history of the evolution of the NCFAS scales as case practice tools, the utility of which lies in helping social workers conduct broad-based evaluations for child welfare and, more recently, non-child welfare families. The results of intake assessments are used for comprehensive case planning and as a vehicle for staffing and review of cases among social workers, social work supervisors and service providers. Closure or exit ratings are used to inform case decisions (but not to drive those decisions) and to serve as outcome indicators for families following receipt of services. These purposes are different from (although not unrelated to) those of highly structured and “manualized” diagnostic instruments.

The preceding studies are known specifically to have contributed to knowledge about the reliability and validity of the NCFAS scales when used in case practice. Many other studies (not reviewed in detail, herein) have used the scales as part of their research or program evaluation studies. (See, e.g., Coll, et.al., 2006; de la Rosa, et al., 2009.) These studies and many more like them, while not focusing on R&V as part of their applied studies, employ the information from the various NCFAS scales for efficacy testing and outcome evaluation of programs.

A Google Scholar search using “ncfas g + r + kirk” results in a list of more than 40 citations of such studies. A more general Google search of NCFAS-G+R produces hundreds of citations, including numerous state reports from states that use the NCFAS scales as part of their child welfare systems, and several child welfare instrument comparisons (e.g., Johnson, et. al., 2008; Johnson, et. al., 2006), in which the NCFAS is highly rated on the basis of its focus on child welfare, per se, and its established reliability and validity. Additional research is ongoing, including research being conducted by NFPN, and by others both independently and in collaboration with NFPN. Results will be forthcoming as studies are completed.

References

- Coll, K.L.; Stewart, R.A.; Morse, R.; & Moe, A (2006). “The value of coordinated services with court-referred clients and their families: An outcome study,” *Child Welfare*, 89(1), 61–80.
- De la Rosa, I.; Perry, J., Johnson, V. (2009). “Benefits of increased home-visitation services: Exploring a case management model,” *Family and Community Health*, 32 (1) 58–75.
- DeVellis, R.F. (1991). *Scale development: Theory and applications*, Newbury Park, CA: Sage.
- Farrell, A.F.; Britner, P.A.; Guzzardo, M.; & Goodrich, S. (2010). “Supportive housing for families in child welfare: Client characteristics and their outcomes at discharge,” *Child and Youth Services Review*, 32, 145–154.
- Johnson, M.; Stone, S.; Lou, C.; Vu, C.; Ling, J.; Mizrahi, P.; & Austin, M.J. (2008). “Family assessment in child welfare services,” *Journal of Evidence-Based Social Work*, 5 (1 & 2).

- Johnson, M.; Stone, S.; Lou, C.; Vu, C.; Ling, J.; Mizrahi, P.; & Austin, M. (2006). "Family Assessment in Child Welfare Services: Instrument Comparisons." (This report compiled by the authors at the Center for Social Services Research, Berkeley, CA. The NCFAS and NCFAS-R are rated as the top assessment tools for child welfare.) Report downloadable as a PDF file at: <http://www.nfnp.org/articles/top-assessment-tools-for-cws>
- Kirk, R.S. (2002). Tailoring Intensive Family Preservation Services for Family Reunification Cases: Final Results of Field Testing and Validation of the North Carolina Family Assessment Scale for Reunification," National Family Preservation Network Project Reports and Monographs, Packard Foundation Final Project Report, http://www.nfnp.org/Portals/0/Documents/ncfas-r_research_report.pdf
- Kirk, R.S (2008). "Development and testing of a family assessment scale for use in child welfare practice settings using differential response," *Protecting Children*, 23 (1&2), 71–87.
- Kirk, R.S. (2010). A Comparative Study of Older-Youth Families and Younger-Youth Families in Placement Prevention and Reunification Cases, Using Traditional Child Protection Services Measures and NCFAS and NCFAS-R Assessment Data, National Family Preservation Network Project Reports and Monographs; Older Youth Study; <http://www.nfnp.org/preservation/research-on-older-youth>
- Kirk, R.S.; Kim, M.M.; & Griffith. D.P. (2005). "Advances in the reliability and validity of the North Carolina Family Assessment Scale," *Journal of Human Behavior in the Social Environment*, 11(3/4), 157–176.
- Kirk, R.S and Martens P. (2012). An examination of the cultural competence and inter-rater reliability of the North Carolina Family Assessment Scales. Project Reports: NFPN Press <http://www.nfnp.org/articles/cultural-competence-report>
- Kirk, R.S and Martens, P. (2014) Family assessment, family functioning and caregiver engagement in family preservation and reunification programs, and the relation of these and other factors to reunification service outcomes, <http://www.nfnp.org/reunification/reunification-research>
- Kirk, R.S and Martens, P. (2015) Field Test and Reliability Analyses of Trauma and Post-Trauma Well-Being Domains of the *North Carolina Family Assessment Scale for General and Reunification Services*, <http://www.nfnp.org/articles/trauma-report>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill
- Pennel, J. (2008). *School-based child and family team project report to the North Carolina Department of Public Instruction*. Available from Center for Family and Community Engagement, North Carolina State University, Campus Box 8622, Raleigh, NC 27695-8622, United States of America.

Randall, K. (2012). *Supportive Housing for Families in Child Welfare: Client Characteristics and Risk Factors at Intake*. University of Connecticut, [DigitalCommons@UConn.edu](#).

Reed-Ashcraft, K.B.; 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), July, 503–520.