

Evidence check

Family functioning assessment tools

8 October 2026

Evidence check question

What are the available family functioning assessment tools and what are their intended populations, domains measured, psychometric properties, strengths and limitations, and practical implementation considerations?

- How do these tools compare to the North Carolina Family Assessment Scale (NCFAS) that is currently being used by Whole Family Teams in NSW?
- Are there family assessment tools designed for, adapted to, or validated with Aboriginal families or culturally and linguistically diverse families; and what is the evidence for their cultural validity or appropriateness?

Summary

Sixteen family functioning assessment tools were identified, differing in purpose, target population, domains measured, complexity and psychometric strength. Several tools are available in multiple versions or variants.

- 2 tools (one of which includes 3 variants) are practitioner completed¹⁻⁴
- 14 self-report tools (or clusters of tools) are completed by family members⁵⁻¹⁴
- 2 of the self-report tools have been specifically developed for Aboriginal and Torres Strait Islander populations¹⁰

The NCFAS practitioner-completed tool and its variants have a broader ecological and child protection focus,¹⁻⁴ while the remaining practitioner-completed and self-report tools focus on family functioning, relationships, cohesion, resilience or family health.⁵⁻¹⁴

Practitioner-completed tools

- The North Carolina Family Assessment Scale (NCFAS) and its variants (NCFAS-R, NCFAS-G, NCFAS-G+R) are distinctive because they are practitioner-completed tools designed for child welfare, family preservation and reunification services, rather than self-reported measures. They assess domains such as environment, parenting capabilities, family interactions, family safety and child wellbeing, with some versions adding social or community functioning and reunification readiness.¹⁻³
 - Psychometric evidence for the NCFAS is strong, with good internal consistency, construct validity, concurrent validity, predictive validity and sensitivity to change.¹⁻³
 - NCFAS is resource intensive as it requires practitioner administration.¹⁻³

Evidence check peer-reviewed

Rapid evidence checks are based on a simplified review method and may not be entirely exhaustive but aim to provide a balanced assessment of what is already known about a specific problem or issue. This evidence check should not be a substitute for individual clinical judgement, nor is it an endorsed position of NSW Health.

- NCFAS incorporates assessment of family strengths and risks across parenting, safety, wellbeing, environmental conditions and service-related outcomes, making it particularly relevant for vulnerable families involved with child protection or reunification services.¹⁻³
- Family Star Plus is also a practitioner-completed measure of family functioning and progress in family support services. It has evidence of reliability, construct validity, responsiveness and predictive validity.⁴

Self-report tools

- Comprehensive instruments such as the McMaster Family Assessment Device (FAD),⁶ Family Assessment Measure III (FAM-III),⁶ Family Adaptability and Cohesion Evaluation Scales (FACES-IV),⁶ Systemic Clinical Outcome and Routine Evaluation (SCORE)^{6, 14} and Family Health Scale (FHS)¹¹ assess multiple dimensions of family functioning and demonstrate acceptable to excellent psychometric properties.
- The FAD and SCORE have strong evidence bases, including reliability, construct validity, criterion validity and ability to distinguish between clinical and non-clinical families.^{6, 14}
- SCORE-15 and SCORE-28 additionally demonstrate responsiveness to therapeutic change, making them useful for routine outcome monitoring.^{6, 14}
- Brief measures such as the General Functioning FAD (GF-FAD),⁷ Brief Assessment of Family Functioning Scale (BAFFS),⁸ and Family Functioning and Cohesion Scale (FFCS)⁹ provide shorter assessment options with good reliability and validity but are less comprehensive than the NCFAS.

Language and cultural adaptation

Several measurement tools have been translated into other languages, and some of these translated versions have undergone psychometric validation. However, the translation and validation studies were predominantly conducted in countries where the target language is the primary language spoken, rather than among immigrant or culturally and linguistically diverse populations living in English-speaking countries. Furthermore, while linguistic translation was undertaken, none of the identified tools were reported to have undergone formal cultural adaptation processes, and no studies specifically evaluated their cultural validity.^{2, 15-31}

More specifically, there is limited evidence that mainstream family functioning measures have been adequately validated with Aboriginal and Torres Strait Islander populations. An Indigenous adaptation study of FACES-II found weaker evidence of construct validity, with concerns that conventional concepts of family cohesion and communication may not adequately reflect Indigenous understandings of family, kinship structures and parenting practices.³²

In contrast, the Mayi Kuwayu Family Functioning Measure (FFM) and the Western Australian Aboriginal Child Health Survey (WAACHS) Family Functioning Scale were identified as culturally adapted and validated measures of family functioning.¹⁰

- Both were developed specifically for Aboriginal and Torres Strait Islander populations and were informed by Indigenous concepts of family, resilience and wellbeing.¹⁰
- The FFM underwent extensive face validity testing with Aboriginal and Torres Strait Islander peoples across urban, regional and remote contexts, and demonstrated excellent reliability and good convergent and divergent validity.¹⁰

- The WAACHS scale was developed in partnership with Aboriginal health professionals and uses language designed to be accessible to speakers of Aboriginal English and Aboriginal languages, although psychometric evidence is more limited.¹⁰

Evidence

Sixteen family functioning tools were identified. A description of each tool as well as a summary of the structure, scoring and psychometrics is provided in this section. Table 1 provides more information on each tool's structure and scoring, Table 2 provides more information on the administration and feasibility of each tool, and Table 3 provides the psychometric properties of each tool.

North Carolina Family Assessment Scale (NCFAS)

The NCFAS was developed by family preservation service providers, policymakers and evaluators to support family preservation practice and program evaluation. It was informed by existing family assessment tools, child maltreatment literature and practice-based approaches.² The reading difficulty level of the NCFAS is 10.9, according to the Flesch-Kincaid Grade Level scale.¹

The original NCFAS assessed 5 domains: environment, parental capabilities, family interactions, family safety and child wellbeing.² The NCFAS-R (for reunification) later added 2 new domains: caregiver/child ambivalence and readiness for reunification.² The NCFAS-G expanded the tool for use in both high-risk and lower risk families by adding social/community life, self-sufficiency and family health domains.² The NCFAS-G+R combines these broader family assessment domains with the reunification domains, making it applicable across risk levels and for families involved in child reunification.²

- The NCFAS is a 25 item assessment tool scored on a 6-point scale from clear strength (+2) to serious problem (-3), and is completed by workers at both intake and case closure to measure family functioning and change over time.¹
- The NCFAS, NCFAS-R, and NCFAS G demonstrated strong internal consistency and high reliability at both intake and case closure, indicating the tool reliably measures family functioning over time.^{1, 2}
- Factor analysis supported the instrument's underlying structure, identifying 5 key factors – environment, parenting capabilities, family interactions, child well-being, and family safety.¹
- The NCFAS demonstrated good construct, concurrent and predictive validity, with significant correlations to established family assessment measures and evidence that scores were associated with important outcomes, including child placement prevention and future placement risk.^{1, 2}
- The NCFAS-R demonstrated good construct and concurrent validity with significant associations between treatment outcomes and domain ratings at intake and closure, changes during intervention, and baseline threshold status.²
- The NCFAS-G demonstrated good construct validity through accurate classification of the target service population, responsiveness to intervention-related change and sensitivity to clinically meaningful improvement.² The NCFAS-G demonstrated limited predictive validity with risk classifications consistent with the absence of subsequent maltreatment reports.²
- The NCFAS tools have been translated into Spanish.^{2, 33}

Family Star Plus

The Family Star Plus was developed as a collaborative assessment and outcomes measurement tool for services supporting families, with the aim of improving family functioning, parenting capacity and children's wellbeing. The tool is designed to facilitate structured conversations between practitioners and parents about strengths, needs and progress, while also providing a measure of change over time. The assessment is based on the Outcomes Star methodology and conceptualises change as a 5-stage Journey of Change from being stuck to effective parenting.⁴

The Family Star Plus assesses family functioning across 10 domains: physical health, your wellbeing, meeting emotional needs, keeping your children safe, social networks, education and learning, boundaries and behaviour, family routine, home and money, and progress to work.⁴

- The Family Star Plus is a practitioner-facilitated assessment comprising 10 outcome areas relevant to parenting and family functioning. Ratings are completed collaboratively by the practitioner and parent using a 10-point scale, and are plotted on a star-shaped profile to identify strengths, needs, and areas for intervention.⁴ The Family Star Plus generates domain-specific scores and an overall profile of family functioning. Repeated assessments can be used to monitor progress over time and support goal setting, care planning and evaluation of service outcomes.⁴
- The Family Star Plus demonstrated good inter-rater reliability, with practitioners showing a high level of agreement when applying the scoring framework, indicating that the scales can be used consistently following training.⁴
- The Family Star Plus showed good internal consistency, and a unidimensional factor structure, supporting the interpretation that the 10 domains collectively measure the broader construct of family functioning.⁴
- The Family Star Plus demonstrated responsiveness to change, with significant improvements observed across all 10 domains between assessment points. Effect sizes were particularly strong among families with greater initial levels of need, suggesting the measure is sensitive to meaningful change over time.⁴
- The Family Star Plus showed evidence of predictive validity, with higher scores on the education and learning, and boundaries and behaviour domains associated with lower levels of subsequent unauthorised school absence, indicating that assessment scores are meaningfully related to important child and family outcomes.⁴
- The Family Star Puls tool has been translated into Romanian (short version), Czech (full), French (resources) and Welsh (resources).¹⁵

Emotional Tone Index for Families (ETIF)

The Emotional Tone Index for Families (ETIF) was developed to assess family systems and support interventions aimed at improving family functioning. The measure focuses on the emotional climate within families by assessing emotional closeness across different types of family relationships and from multiple family members perspectives, while also generating an overall family-level score.⁵

The ETIF was informed by Berscheid's work on emotional tone with items measuring positive and negative affect adapted from the Positive and Negative Affect Schedule (PANAS). The ETIF is also designed with a Flesch reading ease score of 84.3 and Flesch-Kinkaid grade level score of 2.7, making it suitable for use with children aged 8 years and older.⁵

- The ETIF is a 14-item web-based self-report measure (7 positive and 7 negative affect items) that assesses emotional closeness within family relationships. Respondents rate the extent to which each affective term describes their relationship with another family member using a 5-point Likert scale.⁵
- The ETIF generates multiple levels of scores, including unidirectional scores (one family member's view of another), bidirectional scores (the average rating within a family dyad), and an overall family score. Scores reflect positive affect, negative affect and a total emotional closeness score, providing both relationship-specific and whole-family assessments of emotional functioning.⁵
- The ETIF demonstrated good internal consistency, as well as adequate test-retest reliability over time.⁵
- The ETIF showed evidence of construct and convergent validity, with total and subscale scores correlating in the expected direction with established measures of family functioning and cohesion.⁵

Circumplex Model Family Adaptability and Cohesion Evaluation Scales (FACES)

The Circumplex Model Family Adaptability and Cohesion Evaluation Scales (FACES) was designed to assess family functioning based on the 3 Circumplex Model dimensions: family cohesion, flexibility and communication.⁶ The FACES tool was developed for clinical assessment, treatment planning and research on outcome effectiveness of marital and family therapies.⁶ The FACES tool has been revised 3 times and is now its fourth edition, FACES-IV.⁶ A 24-item version has also been developed and is comprised of 4 unbalanced scales: enmeshed, disengaged, chaotic and rigid.⁶

- The FACES-IV contains 42 items with 5-point response formats, which assess family functioning on the 2 main dimensions of the Circumplex Model, cohesion and adaptability, across 6 scales.⁶ Two scales assess the 2 balanced constructs of the model, cohesion and flexibility. Higher scores on the balanced scales are associated with better family functioning.⁶ Four scales assess the 4 unbalanced constructs of the model, disengaged and enmeshed (cohesion), and rigid and chaotic (flexibility).⁶ Higher scores on the unbalanced scales are associated with problematic family functioning.⁶
- Along with the balanced and unbalanced total scores, FACES-IV yields a Circumplex total ratio figure, which provides a summary of a family's balanced and unbalanced characteristics in a single score, which is recommended for use in research.⁶
- The 42-item FACES IV demonstrated acceptable to good internal reliability across most scales, although the enmeshed scale showed lower than acceptable internal consistency.⁶ The 24-item version showed acceptable internal reliability for most scales, although the rigid scale demonstrated lower than acceptable internal consistency.⁶
- The 42-item FACES IV demonstrated construct validity, with strong correlations between most FACES IV scales and other family functioning measures (except the enmeshed and rigid scales), as well as criterion validity through its ability to distinguish between clinical and non-clinical families. Factor analyses also supported a stable factor structure.⁶
- The 24-item version demonstrated acceptable structural validity, with factor analyses supporting the proposed model and overall evidence indicating it is a reliable and valid measure of family functioning, including the capacity to differentiate between clinical and non-clinical populations.⁶
- The FACES IV has been translated into Spanish, Portuguese, Italian, Greek, Hungarian and Turkish, and has demonstrated good psychometrics and similar factor structure in these languages.¹⁶⁻²⁰

- The Indigenous-adapted FACES II showed limited validity for family cohesion and caregiver–child communication, suggesting that standard family and parenting measures may not adequately capture Indigenous family structures, relationships and local parenting practices.³²

The McMaster Family Assessment Device (FAD)

The McMaster Family Assessment Device (FAD) was designed for use as a screening instrument to identify problem areas of family functioning, for self-report completion by family members 12 years and older.⁶ The FAD comprises 6 dimensions associated with dysfunctional family patterns: problem solving, communication, role functioning, affective responsiveness, affective involvement and behaviour control.⁶ The main limitation of the FAD relates to the continued controversy over its factorial validity.⁶

The general functioning subscale of the FAD (GF-FAD) consists of 12 items and has been used as a short form of the FAD to identify healthy and unhealthy area of functioning within families.⁷ The Brief Assessment of Family Functioning Scale (BAFFS) is an ultra-brief measure of general family functioning that can easily be used in routine clinical practice. The measure is comprised of the 3 items with the highest loading on the principal component analysis of the GF-FAD.⁸

- The FAD is a 60-item measure of family functioning comprising 7 domains (with 6 to 12 items per domain), scored using a 4-point Likert scale. It provides both domain-specific scores and an overall family functioning score, with scores ranging from healthy to unhealthy functioning.⁶
- The GF-FAD is a 12-item short form of the FAD that uses only positively worded items and the same 4-point Likert response scale. Item scores are averaged to produce a single overall score, with higher scores indicating more positive perceptions of the family environment and functioning.⁷
- The BAFFS is a 3-item ultra brief version of the GF-FAD.⁸
- The FAD, GF-FAD and BAFFS demonstrate good internal consistency, indicating that the instruments reliably measure family functioning.⁶⁻⁸
- The FAD has demonstrated criterion validity through its ability to distinguish between clinical and non-clinical families,⁶ and construct validity through moderate correlations with other family functioning measures and clinician ratings,^{6, 8} while showing low correlations with social desirability measures.⁶
- The GF-FAD has shown good construct and convergent validity, demonstrated by moderate to strong correlations with other established measures of family and relationship functioning and confirmatory factor analyses supporting an adequate and stable one-factor structure across adolescents and caregivers over time.⁸ The GF-FAD has demonstrated the ability to detect changes in family functioning over time in longitudinal studies, although observed effect sizes were small, suggesting modest sensitivity to change.⁸
- The BAFFS showed good construct and concurrent validity, with strong correlations with self-report and clinician rated family and relationship functioning measures.⁸ The BAFFS was able to detect changes in family functioning over time in longitudinal treatment studies, although the magnitude of change was small, indicating modest sensitivity to change.⁸
- The FAD, GF-FAD, and BAFFS have been translated into Spanish and French, with the Spanish version demonstrating reliability.²¹ The GF-FAD has also been translated into Arabic, and demonstrated reliability and validity.²²

Family Assessment Measure III (FAM III)

The Family Assessment Measure III (FAM III) is based on the process model of family functioning, which is comprised of 7 main concepts related to family functioning: task accomplishment, role performance, communication, affective expression, involvement, control, values and norms.⁶ FAM III was created as an assessment measure in clinical and community practice, a measure of therapy outcomes, and also for use in research on family processes.⁶ FAM III has been designed for use with family members 10 to 12 years and older.⁶

- The FAM-III is a comprehensive self-report measure of family functioning comprising 3 scales: a 50-item general scale assessing overall family system functioning, a 42-item dyadic relationships scale examining functioning between 2 family members, and a 42-item self-rating scale assessing an individual's functioning within the family. These scales generate overall family functioning scores as well as scores across multiple domains of family functioning. Shorter 14-item versions of the general, dyadic relationships, and self-rating scales are also available, each providing an overall measure of family functioning.⁶ All FAM-III items are rated on a 4-point Likert scale.⁶
- The FAM-III demonstrated excellent internal consistency for its 3 main scales (general, dyadic relationships, and self-rating), although internal reliability for some general scale subscales ranged from acceptable to poor. Evidence for test-retest reliability was less favourable, with the general scale and its subscales showing limited stability over time.⁶
- The general scale demonstrated construct validity, with moderate to strong correlations with other established family functioning measures, and both the general and self-rating scales demonstrated criterion validity through their ability to distinguish between problem/clinical and non-problem/non-clinical families.⁶
- Evidence for the underlying structure of the FAM-III is mixed. While the measure was originally conceptualised as a 7-factor model, subsequent analyses supported a single-factor model of family functioning, suggesting uncertainty regarding its dimensional structure.⁶
- The FAM-III has been translated into Danish, Italian, French and Spanish.²³

Family Environment Scale (FES)

The Family Environment Scale (FES) was developed using a social systems ecological model and is not based on a particular theory of family functioning.⁶ It was designed for use with children 11 years and older to assess the social and environmental characteristics of the family.⁶

The FES provides 3 perspectives of family members' accounts of their family environment: real (how it is), ideal (how it would be in an ideal situation), and expected (how it is likely to be in future situations).⁶ The FES can be used in a variety of ways, including family assessment and formulation, assessing different perspectives on family functioning, and as a tool to monitor changes in family functioning and program evaluation.⁶

- The FES is comprised of 90 items across 10 subscales and 3 dimensions of family environment: family relationship, personal growth and system maintenance. Items are answered on a dichotomous scale of true or false.⁶ The short-form version of the FES, which is comprised of 40 items, is answered on a 4-point Likert scale.⁶
- The FES demonstrated less than adequate internal consistency reliability, indicating that some scales may not consistently measure the intended family functioning constructs.⁶

- Evidence supports criterion validity for 5 of the 10 FES subscales, which were able to significantly distinguish between clinical and non-clinical families.⁶
- Several studies have challenged the stability of the original 3-dimension, 10-scale structure of the FES. Alternative factor configurations have been identified, suggesting an unstable factor structure and limited support for the measure's original conceptual framework.⁶
- The short version of the FES has less than adequate reliability and the factor structure did not fit the original model.⁶
- The FES has been translated into more than 20 languages.²⁴

Family Functioning and Cohesion Scale (FFCS)

The Family Functioning and Cohesion Scale (FFCS) was developed as a brief practical self-administered instrument for health professionals to assess family functioning and cohesion. Drawing on existing family assessment measures, the developers selected items that were commonly represented across scales and considered important for describing family characteristics. It has a 3-factor structure: communication, anger/resentment/aggression, and values and beliefs.⁹

The FFCS was designed to provide an easy-to-use tool that supports practitioners in evaluating family relationships and tailoring care to families' specific needs.⁹ The FFCS focuses on key aspects of family life, including interpersonal relationships among family members, personal beliefs, and patterns of interactions and behaviours in the family.⁹

- The FFCS is a 14-item self-report measure that respondents rate their agreement with statements about family life on a 4-point Likert scale.⁹ Scores are calculated by assigning higher scores to positive family functioning items (e.g. mutual support) and reverse-scoring negative items (e.g. family violence). Higher total scores indicate stronger family functioning and cohesion.⁹
- The FFCS demonstrates strong internal consistency and excellent test-retest reliability, indicating that the scale produces stable and consistent results over time.⁹
- The FFCS showed high external validity, with a strong correlation with the GF-FAD, suggesting it measures family functioning in a manner consistent with established instruments.⁹
- The FFCS performed similarly across participants from different world regions, with no significant differences in overall scores between country groups, supporting its cross-cultural robustness and generalisability.⁹

Mayi Kuwayu Family Functioning Measure (FFM)

The Mayi Kuwayu Family Functioning Measure (FFM) is a unidimensional measure of family functioning at the national Aboriginal and Torres Strait Islander population level.¹⁰ The FFM was modified from the Western Australian Aboriginal Child Health Survey family functioning scale. Modification and extensive face validity testing occurred with Aboriginal and Torres Strait Islanders peoples representing saltwater, freshwater, desert and island Aboriginal and Torres Strait Islander mobs across urban, regional and remote Australia.¹⁰ The language of the measure was adapted to reduce wordiness, and increase Aboriginal and Torres Strait Islander participant understanding and cultural relevance, while maintaining the underlying family and resilience protective concepts.¹⁰

- The FFM consists of 9 items rated on a 4-point Likert scale, with responses summed to generate an overall family functioning score; “unsure” responses are treated as missing, with a single missing

item replaced with the participants mean score across the remaining items.¹⁰ Total scores are categorised into four levels of family functioning: low (9 to $\leq$ 24); moderate ($>$ 24 to $\leq$ 29); high ($>$ 29 to $\leq$ 33); very high ($>$ 33 to 36).¹⁰

- The FFM demonstrated excellent internal consistency, indicating very strong reliability.¹⁰
- The FFM demonstrated good convergent and divergent validity. Higher family functioning scores were strongly and significantly associated with better family financial security and lower levels of pain, while no association was found with cardiovascular disease.¹⁰

The Family Health Scale (FHS)

The Family Health Scale (FHS) was developed to address the need for a validated measure of family health for use in research, healthcare and screening settings. Its development was guided by the Family Health Maternal and Child Health Measurement Resource Network and informed by Denham's Family Health Framework, which conceptualises family health across relational, contextual and structural dimensions.¹¹

The FHS was developed through a rigorous expert-led process to define and measure the key dimensions of family health. Drawing on an established theoretical framework and expert consensus, the measure was designed to capture the broad and multidimensional nature of family health. An interdisciplinary team used identified family health domains and evidence from existing validated family measures to develop the scale items.¹¹

The FHS short form (FHS-SF) was subsequently developed by selecting the most representative items from the larger scale, resulting in a concise 10-item measure that captures key aspects of family social and emotional health, family resources, healthy lifestyles and external social supports.¹¹

- The FHS is a 32-item self-report measure of family health and functioning, with responses rated on a 5-point Likert scale. Scores are calculated by reverse-scoring selected items and summing dichotomised item responses, with total scores categorised as poor family health (0 to 5), moderate family health (6 to 8), or excellent family health (9 to 10). A 10-item short form (FHS-SF) is also available.¹¹
- A 4-factor structure comprising family social and emotional health processes, family healthy lifestyle, family health resources, and family external social support is supported with adequate to good model fit.¹¹ The FHS-SF demonstrated good structural validity with adequate to excellent model fit.¹¹
- The FHS and the FHS-SF demonstrated strong internal consistency indicating good to excellent reliability.¹¹
- The FHS demonstrated good convergent validity, with all subscales showing significant correlations in the expected direction with established measures of family functioning and family health. The FHS-SF also showed strong correlations with these measures.¹¹
- Clinically meaningful cut-offs were established for the FHS-SF (poor, moderate and excellent family health). Higher family health scores were significantly associated with lower odds of moderate or severe depression, supporting the scale's ability to distinguish between groups with differing mental health outcomes, demonstrating criterion validity. These associations remained after adjustment for demographic and socioeconomic factors.¹¹

Family Resilience Scale for Veterans (FRS-V)

The Family Resilience Scale for Veterans (FRS-V) was developed to measure family resilience among veterans and their families, drawing on Walsh's model of family resilience.¹² The FRS-V was developed using a multi-stage process that drew on items from 3 established family resilience and coping measures. An expert panel of military mental health clinicians and researchers reviewed, revised and refined the item pool to ensure alignment with Walsh's family resilience framework and relevance to veteran families. The items were then further refined through cognitive testing with veterans and their partners.¹²

- The FRS-V is a brief, 6-items measure. Respondents rate each item on a 4-point Likert scale. Scores of all items are summed and divided by the number of responses, resulting in a possible range of 1 to 4.¹²
- The FRS-V demonstrated excellent internal consistency. A condensed 6-factor, 6-construct model was supported while retaining conceptual alignment with the original theory.¹²
- The FRS-V demonstrated strong construct validity, with correlations in the expected directions with related measures of resilience, family coping, reintegration, self-efficacy, post-traumatic stress disorder and depression. Moderate to strong associations with individual resilience measures and family coping measures supported its ability to assess family-level resilience while remaining distinct from related constructs.¹²
- Evidence of discriminant validity was demonstrated by the absence of significant differences in FRS-V scores across groups defined by employment status or military service component, suggesting the measure is not unduly influenced by these unrelated characteristics.¹²

Protective Factors Survey, 2nd Edition (PFS-2)

The Protective Factors Survey-2 (PFS-2) was developed as a revision of the original Protective Factors Survey (PFS) in response to practitioner and researcher feedback that highlighted limitations in sensitivity to change, ceiling effects, item wording, reverse-scored items and score interpretation. A multidisciplinary team reviewed the protective factors literature, existing measurement tools and emerging issues such as the role of social media in social support and used this evidence to develop a new pool of survey items.¹³

Development involved extensive stakeholder engagement, including consultation with researchers, evaluators, service providers, technical assistance coordinators, and parents with lived experience of family support and child maltreatment prevention services. Focus groups and pilot testing with panel respondents were used to refine items, improve clarity, reduce social desirability bias, and ensure the survey captured a broader range of protective attitudes and behaviours.¹³

The resulting PFS-2 retained 4 core domains from the original survey (family functioning and resilience, nurturing and attachment, social supports, and concrete supports), removed the child development/knowledge of parenting items, and introduced a new caregiver–practitioner relationship subscale. Exploratory factor analysis supported this revised structure, and the measure subsequently underwent field testing across a range of family support and prevention programs.¹³

- The PFS-2 is comprised of 24 items across 5 subscales. Respondents rate each item on a 5-point Likert-type scale.¹³
- The PFS-2 demonstrated good internal consistency.¹³

- Five-factor structure supported comprising family functioning and resilience, nurturing and attachment, social supports, concrete supports, and caregiver–practitioner relationship. The caregiver–practitioner relationship factor emerged from items initially intended to measure social support and was developed into a distinct subscale.¹³
- The PFS-2 has been translated into Spanish, and demonstrated reliability and validity.²⁵

Beavers Systems Model Self-Report Family Inventory (SFI)

The Beavers Systems Model Self-Report Family Inventory (SFI) was developed to classify a family's profile of functioning based on the Beavers Systems Model and centres on 2 constructs, family competence and family style, providing a cross-sectional perspective of family functioning.⁶

- The SFI is a 36-item self-report measure that uses a 3-point Likert scale.⁶ The SFI yields scores on 5 domains of family functioning: health/competence, conflict, cohesion, leadership and emotional expressiveness.⁶
- The SFI demonstrates good internal consistency. However, evidence for test-retest reliability is mixed, with some studies reporting adequate stability over time and others finding weaker reliability for the conflict, cohesion and leadership scales.⁶
- The SFI has demonstrated construct validity through moderate correlations with other self-report and clinician-rated measures of family functioning, and criterion validity through its ability to distinguish between clinical and non-clinical families.⁶
- While the SFI was originally conceptualised around five family functioning domains, research has reported differing factor structures (5-, 6-, and 2-factor models), indicating some uncertainty regarding the instrument's underlying dimensional structure despite overall support for its use as a family functioning measure.⁶
- The SFI has been translated into Persian, and demonstrated reliability and validity.²⁶

Systemic Clinical Outcome and Routine Evaluation (SCORE)

The Systemic Clinical Outcome and Routine Evaluation (SCORE) was developed in the UK and Ireland as a family functioning and outcome measure for routine use in family therapy and clinical services. Its development was informed by a review of family assessment theories and existing psychometric instruments, clinical expertise, and extensive consultation with family therapists, researchers, service users and practice research networks.^{6, 14}

The original SCORE-40 was created from a large pool of existing and newly developed items designed to assess key aspects of family functioning, including family atmosphere and mood, conflict, communication, roles and rules, and adaptability. The measure was refined through qualitative and quantitative studies, expert review, and testing with families attending therapy services, resulting in a psychometrically validated 40-item instrument.¹⁴

To improve practicality in routine clinical settings, shorter versions (SCORE-28 and SCORE-15) were subsequently developed through empirical item reduction and expert appraisal, retaining the most clinically useful and sensitive items.¹⁴ The instrument is designed for use by family members 12 years and older.⁶ Items from the 15- and 28-item versions were combined by this team into a 29-item version, primarily for use in a series of studies on the psychometric properties of the SCORE-15 and SCORE-28.¹⁴ Additional adaptations were created for specific populations, including the Child SCORE-15 for children aged 7 to 10 years and the Relational SCORE-15 for lesbian, gay and bisexual adults.^{6, 14}

The SCORE is designed to assess family functioning and monitor change over the course of therapy, with the brief versions suitable for routine outcome monitoring and the SCORE-40 more commonly used in research. It has been widely adopted by professional family therapy organisations and translated into numerous languages for international use.⁶

- The SCORE is self-report, with 6 main versions: SCORE-40 (40-item), SCORE-28 (28-item), SCORE-15 (15-item), SCORE-29 (29-item specifically developed to validate the SCORE 28 and SCORE 15), CHILD SCORE-15, and Relational SCORE-15.^{6, 14} It uses statements about family life rated on 5- or 6-point Likert scales. Scores are calculated as mean total and subscale scores, with higher scores indicating poorer family functioning.^{6, 14} In addition to these descriptive items, respondents identify their family's most significant current problem and rate its severity on a 10-point scale.^{6, 14}
- Web-based administration and scoring are available for the SCORE-15 and SCORE-28 versions.⁶
- The SCORE-40, SCORE-28, SCORE-15, Child SCORE-15, and Relational SCORE-15 demonstrated adequate to excellent internal consistency.^{6, 14}
- The SCORE-40, SCORE-28, and SCORE-15 demonstrated a generally stable 3-factor structure comprising family strengths, family difficulties and communication.^{6, 14} although some SCORE-28 studies have reported variation in factor structure across samples.
- The SCORE-40, SCORE-28, SCORE-15, and Child SCORE-15 demonstrated good test-retest reliability.^{6, 14} The SCORE 15- and 28-item versions have been shown to demonstrate responsiveness to change and sensitivity to therapeutic improvement, supporting their use for monitoring outcomes during family therapy.^{6, 14}
- The SCORE-40, SCORE-28, and SCORE-15 demonstrated construct validity through expected correlations with other measures of family functioning and wellbeing measures;^{6, 14} divergent validity through negligible associations with demographic characteristics and social desirability;¹⁴ and criterion validity through its ability to distinguish between clinical and non-clinical families.^{6, 14}
- Preliminary validation of the Child SCORE-15 indicated that the measure performed well in assessing family adjustment from a child's perspective and was appropriate for use with younger respondents through age-adapted language and response formats.^{6, 14}
- The Relational SCORE-15 showed evidence of content validity through adaptation and review by focus groups and experts, and criterion validity through its ability to distinguish between lesbian, gay and bisexual and non-lesbian, gay and bisexual samples on relationship adjustment measures.^{6, 14}
- The SCORE-15 has been translated into more than 30 languages.²⁷ The Portuguese, Swedish and Thai versions demonstrated reliability and validity.²⁸⁻³⁰

Systemic Therapy Inventory of Change (STIC)

The Systemic Therapy Inventory of Change (STIC) was specifically designed to track change in family and couple therapy, and to bring a multi-systemic perspective to the study of change in individual therapy.⁶ STIC is based on the Integrative Problem Centered Meta Frameworks which provides a multi-systemic perspective for family, couple and individual therapy.⁶ The STIC assesses and monitors client progress and therapeutic alliance over the course of therapy, and feeds this information to therapists and other stakeholders on demand.⁶

- The STIC system is comprised of 2 self-report assessments, the STIC Initial and the STIC Intersession.⁶

- The STIC Initial is completed by family members before the first therapy session and contains a demographic questionnaire and 6 system scales: individual problems and strengths; family of origin; relationship with partner; family/household; child problems and strengths; and relationship with child.⁶ Responses on each scale are given on 5-point Likert scales.⁶
- The STIC Intersession is completed within 24 hours of the next therapy session and contains brief versions of the 6 scales of the STIC Initial and an additional 3 scales to assess therapeutic alliance, which are administered based on the type of therapy a client is receiving: Individual therapy alliance scale, couple therapy alliance scale, and family therapy alliance scale.⁶ Responses on each of the therapeutic alliance scales are given on 7-point Likert scales.⁶
- The factor structure of the STIC is supported. The STIC demonstrated adequate to good internal consistency across its scales and subscales, with subsequent validation studies also supporting test-retest reliability, indicating that scores are stable and consistent over time.⁶
- Four of the 5 scales showed strong construct validity through substantial correlations with established assessment measures.⁶ The STIC demonstrated good criterion validity, including the ability to differentiate between clinical and non-clinical populations, and validation studies established clinically meaningful cut-off scores for interpreting results.⁶
- The STIC has been translated into German, and demonstrated validity and reliability.³¹

Western Australian Aboriginal Child Health Survey (WAACHS)

The Western Australian Aboriginal Child Health Survey (WAACHS) was developed as a one-factor family functioning scale for use in Aboriginal communities in Western Australia, to measure the extent to which families have established a climate of cooperation, emotional support and good communication.¹⁰

WAACHS was created using family and resilience protective factors identified in the literature review of international research on family resilience. The wording used in the WAACHS family functioning scale was developed in collaboration with Aboriginal health professionals to ensure it was able to be interpreted by respondents whose first language was Aboriginal English or an Aboriginal language.¹⁰

- The WAACHS comprises 9 items that are scored and summed to produce an overall family functioning score, which can be categorised into 4 levels: poor (9 to 34); fair (35 to 38); good (39 to 41); very good (42 to 45). These categories were noted by the developers to be somewhat arbitrary.¹⁰
- Limited psychometric testing has been conducted on the WAACHS family functioning scale.¹⁰
- The WAACHS demonstrated good convergent validity through significant associations with psychosocial factors known to be related to family functioning, divergent validity through the absence of associations with unrelated physical health conditions.¹⁰

Background

The NSW Health Whole Family Team program supports families with complex needs where parental mental health concerns, alcohol and other drug use, trauma, domestic and family violence, and child safety concerns intersect. Whole Family Teams work from a systemic, family-centred approach that recognises the interconnected nature of family relationships and the influence of broader service systems, including health, child protection, education and community services. The model aims to

strengthen family functioning, enhance parenting capacity, improve child wellbeing and safety, and support families to achieve sustainable positive outcomes.³⁴

Family functioning is an important determinant of child wellbeing, safety and developmental outcomes. Assessing family functioning helps identify family strengths, support needs, and areas of risk across domains such as communication, relationships, parenting, emotional wellbeing, safety and problem-solving.³⁴ Measuring family functioning can inform care planning, monitor progress over time and evaluate the impact of interventions.^{2, 6} The contextual differences in the relational caregiving ecosystem or environment of Aboriginal and Torres Strait Islander populations highlights a need for adequately validated family functioning measures for Aboriginal and Torres Strait Islander populations.³⁵ For services working with families experiencing multiple and complex challenges, understanding family functioning is critical to identifying factors that may contribute to child safety concerns, family resilience, and long-term outcomes for children and their caregivers.³⁴

Method

PubMed and Google searches were conducted on 15 June 2026. A total of 461 peer-reviewed studies (after removing duplicates) returned from PubMed search were screened using TERA and Covidence software. See Appendix 1 for the search strategy and inclusion criteria.

Limitations

A key limitation of this rapid review is that it was restricted to English-language publications published from 2016 onwards, which may have resulted in the exclusion of relevant evidence, validation studies, or tool development research published earlier or in other languages. By excluding studies that used family functioning assessment tools solely as outcome measures without reporting psychometric properties, as well as studies describing family functioning theories or concepts, some potentially useful contextual information may have been omitted.

Evidence tables

Table 1: Structure and scoring of family functioning tools

Family functioning measurement tools: structure and scoring				
Source	Items	Response	Scoring	Categories
North Carolina Family Assessment Scale (NCFAS) Reed-Ashcraft, Kirk (2001) Kirk (2015)	25 items across 5 domains: environment, parental capabilities, family interactions, family safety, child wellbeing.	6-point ordinal scale (serious problem, moderate problem, mild problem, baseline/adequate, mild strength, and clear strength) ranging from +2 (clear strength) to -3 (serious problem).	Domain and item ratings used to assess strengths and problems.	No formal clinical cut-offs reported. Ratings interpreted along the continuum from serious problem (-3) to clear strength (+2).
North Carolina Family Assessment Scale – Reunification (NCFAS-R) Reed-Ashcraft, Kirk (2001) Kirk (2015) Kirk (2002)	48 items across 7 domains: environment, parental capabilities, family interactions, family safety, child wellbeing, ambivalence, readiness for reunification.	6-point ordinal scale (serious problem, moderate problem, mild problem, baseline/adequate, mild strength, and clear strength) ranging from +2 (clear strength) to -3 (serious problem).	Domain and item ratings used to assess strengths and problems.	No formal clinical cut-offs reported. Ratings interpreted along the continuum from serious problem (-3) to clear strength (+2).
North Carolina Family Assessment Scale – General (NCFAS-G) Reed-Ashcraft, Kirk (2001) Kirk (2015)	58 items across 9 domains: environment, parental capabilities, family interactions, family safety, child wellbeing, youth well-being, social/community life, self-sufficiency, family health.	6-point ordinal scale (serious problem, moderate problem, mild problem, baseline/adequate, mild strength, and clear strength) ranging from +2 (clear strength) to -3 (serious problem).	Domain and item ratings used to assess strengths and problems.	No formal clinical cut-offs reported. Ratings interpreted along the continuum from serious problem (-3) to clear strength (+2).

Family functioning measurement tools: structure and scoring				
Source	Items	Response	Scoring	Categories
North Carolina Family Assessment Scale – General + Reunification (NCFAS-G+R) California Evidence-Based Clearinghouse (2026)	70 items across 10 domains: environment, parental Capabilities, family interactions, family safety, child wellbeing, youth well-being, social/community life, self-sufficiency, family health, caregiver/child ambivalence, and readiness for reunification.	6-point ordinal scale (serious problem, moderate problem, mild problem, baseline/adequate, mild strength, and clear strength) ranging from +2 (clear strength) to –3 (serious problem).	Domain and item ratings used to assess strengths and problems.	No formal clinical cut-offs reported. Ratings interpreted along the continuum from serious problem (–3) to clear strength (+2).
Family Star Plus Good and MacKeith (2021)	10 outcome areas.	Collaborative assessment completed jointly by the practitioner and parent. Each domain is rated on a 10-point scale reflecting the family's position on a journey of change.	Each domain receives a score from 1 to 10. Scores are plotted on a star diagram and can be compared over time to assess change. Ratings may be agreed collaboratively or recorded separately when agreement cannot be reached.	Higher scores indicate more effective functioning and parenting. Journey of change stages: stuck (1–2), accepting help (3–4), trying/believing and trying (5–6), finding what works (7–8), and effective parenting (9–10).
Emotional Tone Index for Families (ETIF) Kerr, Hudenko (2021)	14 items (7 positive, 7 negative).	5-point Likert scale from strongly disagree (1) to strongly agree (5). Each member rates closeness with every other family member, yielding unidirectional (e.g. mother to father), bidirectional (e.g., average of mother to father unidirectional and father to mother unidirectional score), and	Produces unidirectional, bidirectional, and overall family-level scores. Unidirectional score for each member regarding each of the other family members are obtained. The positive affect (PA) score is calculated by summing the 7 positive items and the negative affect (NA)	Not reported.

Family functioning measurement tools: structure and scoring				
Source	Items	Response	Scoring	Categories
		total family ETIF scores (average of every unidirectional score).	score is calculated by summing the 7 negative items. Since each item is rated on a scale from 1 to 5, PA and NA scores range between 7 and 35. A unidirectional total ETIF score is then calculated by subtracting the NA scale score from the PA scale score and adding 28 so that the values cannot be negative (0–56). Bidirectional scores are calculated for each dyad by averaging each unidirectional score. Total family scores are calculated by averaging all the unidirectional scores.	
Circumplex Model Family Adaptability and Cohesion Evaluation Scales (FACES-IV) Hamilton and Carr (2016)	42 items (full version); 24-item short version. 6 scales: balanced cohesion, balanced flexibility, disengaged, enmeshed, rigid, chaotic.	5-point response formats.	Along with the balanced and unbalanced total scores, FACES-IV yields a circumplex total ratio figure, which provides a summary of a family's balanced and unbalanced characteristics in a single score. Circumplex ratio score can be calculated.	Higher scores on the balanced scales are associated with better family functioning. Higher scores on the unbalanced scales are associated with problematic family functioning.

Family functioning measurement tools: structure and scoring				
Source	Items	Response	Scoring	Categories
The McMaster Family Assessment Device (FAD) Hamilton and Carr (2016)	60 items across 7 domains: problem solving, communication, roles, affective responsiveness, affective involvement, behaviour control.	4-point Likert scale: strongly agree, agree, disagree and strongly disagree.	The measure yields a general functioning score and a score for each domain.	Scores interpreted on a healthy to unhealthy continuum.
General Functioning – Family Assessment Device (GF-FAD) Sianko and McDonell (2020)	12 items.	4-point Likert scale: strongly agree, agree, disagree and strongly disagree.	GF-FAD items averaged.	Scores interpreted on a healthy to unhealthy continuum. Higher scores reflect more positive family functioning.
Brief Assessment of Family Functioning Scale (BAFFS) Mansfield, Keitner (2019)	3 items (derived from GF-FAD).	4-point Likert scale: strongly agree, agree, disagree and strongly disagree.	Brief global family functioning score.	Not reported.
Family Assessment Measure III (FAM III) Hamilton and Carr (2016)	General scale: 50 items, 9 subscales. Dyadic scale: 42 items, 7 subscales. Self-rating scale: 42 items, 7 subscales. 14-item version of the FAM scales.	4-point Likert scales ranging from strongly agree to strongly disagree.	Total and subscale scores generated for family functioning domains. 14-item version of the FAM scales (general, dyadic relationships, and self-rating), which each yield an overall score of family functioning.	Not reported.
Family Environment Scale (FES) Hamilton and Carr (2016)	90 items (full version); 40-item (short form-version).	Full version: true/false. Short form: 4-point Likert scale.	Scores generated across 10 subscales and 3 broad dimensions.	Not reported.

Family functioning measurement tools: structure and scoring				
Source	Items	Response	Scoring	Categories
Family Functioning and Cohesion Scale (FFCS) Tsoutsi and Dikeos (2024)	14 items.	4-point response scale: completely true, more true than false, more false than true, completely false.	Positive items scored 0–3; negative items reverse scored. For the items that contribute positively to a higher family cohesion (e.g. We always support each other) the highest points (i.e. a score of 3) are given to the completely true statement (and a score of 0 to the completely false), while for the negative items (e.g. There are episodes of physical violence in the family) the score is reversed.	Higher scores indicate greater cohesion/functioning.
Mayi Kuwayu Family Functioning Measure (FFM) Brinckley, Jones (2022)	9 items.	4-point Likert scale: not at all (score = 1), a little bit (score = 2), a fair bit (score = 3), a lot (score = 4), or unsure (recoded to missing).	A total family functioning score is created by summing responses to the nine items (range 9–36). One missing response can be replaced by the mean of remaining items; more than 1 missing item results in missing total score.	Low family functioning (scores: 9 to ≤ 24), moderate family functioning (> 24 to ≤ 29), high family functioning (> 29 to ≤ 33) and very high family functioning (> 33 to 36).
Family Health Scale (FHS) Crandall, Weiss-Laxer (2020)	32 items (long-form version); 10 items (short-form version).	5-point Likert scale from strongly disagree to strongly agree.	Negatively worded items are reverse scored. After reverse coding items create binary variables for each item by coding items with a	Higher scores indicate better family health. Poor family health: 0–5; moderate: 6–8; excellent: 9–10.

Family functioning measurement tools: structure and scoring				
Source	Items	Response	Scoring	Categories
			score of 4 or 5 as one and coding those with a score less than 4 as 0. Sum the scores.	
Family Resilience Scale for Veterans (FRS-V) Finley, Pugh (2016)	6 items.	4-point Likert scale from strongly disagree (1) to strongly agree (4).	Total score calculated as sum of items divided by number of responses (range 1–4).	Not reported.
Protective Factors Survey, 2nd Edition (PFS-2) Sprague-Jones, Counts (2019)	24 items across 5 subscales: family functioning and resilience, nurturing and attachment, social supports, concrete supports, and caregiver–practitioner relationship.	5-category responses; wording varies by subscale. Social supports, family functioning and resilience, nurturing and attachment: not at all like my life, not much like my life, somewhat like my life, quite a lot like my life, just like my life. Concrete supports: never, rarely, sometimes, often, almost always. Caregiver–practitioner relationship: strongly agree, agree, neither, agree nor disagree, disagree, strongly disagree.	Subscale scores calculated separately.	Not reported.

Family functioning measurement tools: structure and scoring				
Source	Items	Response	Scoring	Categories
Beavers Systems Model Self-Report Family Inventory (SFI) Hamilton and Carr (2016)	36 items.	3-point Likert scale.	The SFI yields scores on 5 domains of family functioning: health/competence, conflict, cohesion, leadership, and emotional expressiveness.	Not reported.
Systemic Clinical Outcome and Routine Evaluation (SCORE) Hamilton and Carr (2016) Carr and Stratton (2017)	Multiple versions: SCORE-40 (40-items), SCORE-29 (29-items), SCORE-28 (28-items), SCORE-15 (15-items), Child SCORE-15 (15-items), R-SCORE-15 (15-items). In addition to these descriptive items, all versions contain an item which asks respondents to write down what they think is the biggest problem or challenge in their family at the moment.	5- or 6-point Likert scale from "describes my family extremely well" to "does not describe my family at all". The additional item regarding the biggest problem or challenge in their family at the moment is rated on a 10-point scale of severity.	Negatively worded items reverse-scored. Total and subscale scores are the mean for items in the scale, so that scores range from 1–5 or 1–6 depending on whether response to items are given on a 5- or 6-point Likert scale.	Higher scores indicate poorer functioning. Low scores indicate better adjustment. Child SCORE guidance: scores near 1–2 indicate few problems; scores of 4–5 indicate clinically significant problems.
Systemic Therapy Inventory of Change (STIC) Hamilton and Carr (2016) Sprague-Jones, Counts (2019)	The STIC system is comprised of 2 self-report assessments, the STIC Initial and the STIC Intersession. The STIC Initial is completed before the first therapy session and contains a demographic questionnaire and 6 system scales: individual problems and strengths; family of origin; relationship with partner;	STIC Initial scales: 5-point frequency Likert scale. STIC Intersession scales: 7-point agreement Likert scale.	Not reported.	Cut-offs established but not reported/detailed.

Family functioning measurement tools: structure and scoring				
Source	Items	Response	Scoring	Categories
	family/household; child problems and strengths; and relationship with child. The STIC Intersession is completed within 24 hours of the next therapy session and contains brief versions of the 6 scales of the STIC Initial and an additional three scales to assess therapeutic alliance, which are administered based on the type of therapy a client is receiving: Individual Therapy Alliance Scale, Couple Therapy Alliance Scale, and Family Therapy Alliance Scale.			
Western Australian Aboriginal Child Health Survey (WAACHS) Brinckley, Jones (2022)	9 items.	Not reported.	WAACHS family functioning scale items are summed to total score.	Poor (9–34), fair (35–38), good (39–41) and very good (42–45).

Table 2: Administration and feasibility characteristics of identified family functioning tools

Family functioning measurement tools: administration and feasibility					
Source	Respondent type	Modalities	Completion time	Training requirements	Cost/licensing
North Carolina Family Assessment Scale: NCFAS National Family Preservation Network (2026) California Evidence-Based Clearinghouse (2026)	Clinician rated	Paper and pen Online	30–40 minutes	No minimum degree or experience required to administer the scale. The scales are designed to be intuitive. Agencies are strongly encouraged to complete the training prior to using the scale in practice. The training materials in the package are self-administered, can be done individually or in groups, and require several hours to complete. Live or online training. National Family Preservation Network (NFPN) offers supplemental training for a fee.	The cost of the tools is based on the number of users. License for the number of staff who are using the tool at any given time.
North Carolina Family Assessment Scale for Reunification: NCFAS-R Reed-Ashcraft, Kirk (2001) Kirk (2015) Kirk (2002)	Clinician rated	Paper and pen Online	30–40 minutes	No minimum degree or experience required to administer the scale. The scales are designed to be intuitive. Agencies are strongly encouraged to complete the training prior to using the scale in practice. The training materials in the package are self-administered, can be done individually or in groups, and require several hours to complete.	The cost of the tools is based on the number of users. License for the number of staff who are using the tool at any given time.

Family functioning measurement tools: administration and feasibility					
Source	Respondent type	Modalities	Completion time	Training requirements	Cost/licensing
				Live or online training. NFPN offers supplemental training for a fee.	
North Carolina Family Assessment Scale: NCFAS-G Reed-Ashcraft, Kirk (2001) Kirk (2015)	Clinician	Paper and pen Online	30–40 minutes	No minimum degree or experience required to administer the scale. The scales are designed to be intuitive. Agencies are strongly encouraged to complete the training prior to using the scale in practice. The training materials in the package are self-administered, can be done individually or in groups, and require several hours to complete. Live or online training. NFPN offers supplemental training for a fee.	The cost of the tools is based on the number of users. License for the number of staff who are using the tool at any given time.
North Carolina Family Assessment Scale – General + Reunification: NCFAS-G+R California Evidence-Based Clearinghouse (2026)	Clinician	Paper and pen Online	30–40 minutes	No minimum degree or experience required to administer the scale. The scales are designed to be intuitive. Agencies are strongly encouraged to complete the training prior to using the scale in practice. The training materials in the package are self-administered, can be done	The cost of the tools is based on the number of users. License for the number of staff who are using the tool at any given time.

Family functioning measurement tools: administration and feasibility					
Source	Respondent type	Modalities	Completion time	Training requirements	Cost/licensing
				individually or in groups, and require several hours to complete. Live or online training. NFPN offers supplemental training for a fee.	
Family Star Plus Good and MacKeith (2021) Outcomes Star (2026)	Clinician	Paper and pen Online	Not reported.	All license holders must complete the training. Training options: • In-house training is priced according to number of participants: £980 for 4, £1,200 for 5–8, £1,320 for 9–12, £1,630 for 13–16. • Open training is £245 per person for groups of 3 or less. • To develop your own in-house trainer is £2,500 for initial training and a one-year license with ongoing support. Once completed, licence renewal, ongoing support and CPD programme is £1,250 per year. • Refresher courses or practice reviews for £1,320 for a full-day programme or £665 for a half-day.	The cost of the tool depends on the number of users. Account: £255/€309 per annum (includes 1 Account Lead licence/login); Licence and login per annum: £41/€50 per practitioner. Practitioners need to be licensed to use Outcomes Star.
Emotional Tone Index for Families (ETIF) Kerr, Hudenko (2021)	Self-rated	Online	Less than 1 hour	Not reported.	Not reported.

Family functioning measurement tools: administration and feasibility					
Source	Respondent type	Modalities	Completion time	Training requirements	Cost/licensing
Circumplex Model Family Adaptability and Cohesion Evaluation Scales (FACES-IV) Hamilton and Carr (2016) The National Child Traumatic Stress Network (2026)	Self-rated	Paper and pen	Not reported.	Not reported.	\$95 FACES-IV package with unlimited use; \$75 for students. Copyrighted.
The McMaster Family Assessment Device (FAD) Hamilton and Carr (2016) The National Child Traumatic Stress Network (2026)	Self-rated	Paper and pen	10–15 minutes	Not reported.	Not reported.
General Functioning - Family Assessment Device (GF-FAD) Sianko and McDonell (2020) The National Child Traumatic Stress Network (2026) Assaf, Hallit (2025)	Self-rated	Paper and pen	Not reported.	Not reported.	Not reported.
Brief Assessment of Family Functioning Scale (BAFFS) Mansfield, Keitner (2019) The National Child Traumatic Stress Network (2026) Assaf, Hallit (2025)	Self-rated	Paper and pen	Not reported.	Not reported.	Not reported.

Family functioning measurement tools: administration and feasibility					
Source	Respondent type	Modalities	Completion time	Training requirements	Cost/licensing
Family Assessment Measure III (FAM III) Hamilton and Carr (2016) Skinner, Steinhauer (2016) Australian Institute of Family Studies (2024)	Self-rated	Not reported.	30–40 minutes	Not reported.	Available for purchase \$90+
Family Environment Scale (FES) Hamilton and Carr (2016) Australian Institute of Family Studies (2024) Mind Garden (2026)	Self-rated	Online	15–20 minutes	Not reported.	\$150+ Can purchase a license to administer the survey via a non-Mind Garden survey system. License is \$137.50+
Family Functioning and Cohesion Scale (FFCS) Tsoutsis and Dikeos (2024)	Self-rated	Not reported.	Not reported.	Not reported.	Available in research article.
Mayi Kuwayu Family Functioning Measure (FFM) Brinckley, Jones (2022)	Self-rated	Paper and pen Online	Not reported.	Not reported.	Available in research article supplementary material.
Family Health Scale (FHS) Crandall, Weiss-Laxer (2020) Family Well-Being Research Network (2026)	Self-rated	Not reported.	Not reported.	Not reported.	Free. Available in research article.

Family functioning measurement tools: administration and feasibility					
Source	Respondent type	Modalities	Completion time	Training requirements	Cost/licensing
Family Resilience Scale for Veterans (FRS-V) Finley, Pugh (2016) Family Well-Being Research Network (2026)	Self-rated	Not reported.	Not reported.	Not reported.	Free. Available with permission from author.
Protective Factors Survey, 2nd Edition (PFS-2) Sprague-Jones, Counts (2019) California Evidence-Based Clearinghouse (2026)	Self-rated	Paper and pen Online	10–15 minutes	A detailed user manual is available which includes all necessary information required to administer the instrument.	There are no licensing requirements to use the tool. All materials, including the user manual and supporting guides, are available for download (https://friendsnrc.org/protective-factors-survey/pfs-2/) and are free of charge.
Beavers Systems Model Self-Report Family Inventory (SFI) Hamilton and Carr (2016)	Self-rated	Not reported.	Not reported.	Not reported.	Not reported.
Systemic Clinical Outcome and Routine Evaluation (SCORE) Hamilton and Carr (2016), Carr and Stratton (2017) Association for Family and Systemic Psychotherapy (2026)	Self-rated	Paper and pen App	Less than 10 minutes	Guides to using SCORE-15 are available at: https://www.theafsp.org/publications-resources/evaluating-family-therapy-score-15/	Free. Available at: https://www.theafsp.org/publications-resources/evaluating-family-therapy-score-15/

Family functioning measurement tools: administration and feasibility					
Source	Respondent type	Modalities	Completion time	Training requirements	Cost/licensing
Systemic Therapy Inventory of Change (STIC) Hamilton and Carr (2016) He, Zinbarg (2024) The Family Institute at Northwestern University (2017)	Self-rated	Paper and pen Online	45 minutes	Not reported.	Access through The Family Institute at Northwestern University.
Western Australian Aboriginal Child Health Survey (WAACHS) Brinckley, Jones (2022)	Self-rated	Not reported.	Not reported.	Not reported.	Not reported.

Table 3: Psychometric properties of identified family functioning tools

Family functioning tools: psychometrics							
Source	Reliability/ internal consistency	Factor structure/ construct validity	Convergent/ concurrent validity	Criterion validity/ clinical discrimination	Test-retest/ responsiveness	Cultural validation	Key limitations
North Carolina Family Assessment Scale: NCFAS Reed-Ashcraft, Kirk (2001) Kirk (2015)	Good to excellent internal consistency ($\alpha=.70-.94$; intake $\alpha=.72-.90$; closure $\alpha=.79-.91$).	Five-factor structure supported: environment, parental capabilities, family interactions, family safety, child wellbeing.	Significant correlations with related measures: child wellbeing scales, Index of Family Relations, Family Index of Resources for Management	Closure strengths associated with non-placement; problem ratings associated with placement.	Sensitive to change between intake and closure.	Translated into Spanish.	Not reported.

Family functioning tools: psychometrics							
Source	Reliability/ internal consistency	Factor structure/ construct validity	Convergent/ concurrent validity	Criterion validity/ clinical discrimination	Test-retest/ responsiveness	Cultural validation	Key limitations
			($r=.261-.708$, $p<.01$).				
North Carolina Family Assessment Scale for Reunification: NCFAS-R Reed-Ashcraft, Kirk (2001) Kirk (2015) Kirk (2002)	Excellent internal consistency (intake $\alpha=.82-.92$; closure $\alpha=.81-.94$).	Original NCFAS structure plus reunification domains supported.	Significant associations between treatment outcomes and: overall aggregate domain ratings at intake and closure; improvement or deterioration on domain ratings during intervention; and being at or above the baseline threshold at intake and closure.	Threshold ratings associated with reunification outcomes.	Sensitive to improvement and deterioration over intervention.	Translated into Spanish.	Less psychometric research than original NCFAS.
North Carolina Family Assessment Scale: NCFAS-G Reed-Ashcraft, Kirk (2001) Kirk (2015)	Excellent internal consistency (intake $\alpha=.83-.95$; closure $\alpha=.88-.98$).	Structure supported in differential response populations.	Ratings consistent with low-moderate risk populations and expected service outcomes.	Significant increase in families functioning at or above adequate levels at closure.	Demonstrated change following service delivery.	Translated into Spanish.	Evidence mainly from differential- response settings.

Family functioning tools: psychometrics							
Source	Reliability/ internal consistency	Factor structure/ construct validity	Convergent/ concurrent validity	Criterion validity/ clinical discrimination	Test-retest/ responsiveness	Cultural validation	Key limitations
North Carolina Family Assessment Scale – General + Reunification: NCFAS-G+R California Evidence-Based Clearinghouse (2026) Kirk (2015) Kirk and Martens (2012)	Good internal consistency ($\alpha=.87-.93$ intake; $\alpha=.88-.96$ closure).	Factor structure was confirmed.	New scale items and domains correlated predictably and significantly with other NCFAS domains.	Not reported.	Not reported.	Translated into Spanish.	Less psychometric research than original NCFAS.
Family Star Plus Good and MacKeith (2021) Outcomes Star (2026)	Good inter-rater reliability, with a mean Krippendorff's $\alpha = .83$ when practitioners applied ratings to standardised case studies. Consistency between raters was also good ($\alpha = .75$).	Single-factor (unidimensional) structure supported.	Not reported.	Higher scores on the education and learning domain predicted lower subsequent unauthorised school absence ($B = -2.24, p < .001$), and higher Boundaries and Behaviour scores also predicted	Good responsiveness to change: significant improvements ($p < .001$) were observed across all 10 domains between assessments.	Translated into Romanian (short version), Czech (full), French (resources), and Welsh (resources). No cultural or language validations reported.	Convergent validity and formal test-retest reliability have not been established.

Family functioning tools: psychometrics							
Source	Reliability/ internal consistency	Factor structure/ construct validity	Convergent/ concurrent validity	Criterion validity/ clinical discrimination	Test-retest/ responsiveness	Cultural validation	Key limitations
	Good internal consistency (Cronbach's $\alpha = .86$).			lower absence ($B = -1.87, p = .004$).			
Emotional Tone Index for Families (ETIF) Kerr, Hudenko (2021)	Good internal consistency (Positive Affect $\alpha=.87$; Negative Affect $\alpha=.85$).	Structure reflects positive and negative affect dimensions.	Moderate to strong correlations with GF-FAD, BAFFS, and FACES-IV.	Not reported.	Adequate test-retest reliability (ICC=.57 for negative affect scale, .75 for positive affect scale, .71 for total ETIF, and .49 for the lie scale).	No cultural or language validations reported.	Potential response bias (27% endorsed lie-scale item).
Circumplex Model Family Adaptability and Cohesion Evaluation Scales (FACES-IV; 42-items) Hamilton and Carr (2016) The National Child Traumatic Stress Network (2026)	Acceptable-good reliability for most scales; enmeshed scale weaker.	Circumplex model supported.	Strong correlations with family functioning measures, except enmeshed and rigid scales.	Distinguishes problematic from non-problematic families.	Not reported.	Has been translated into Spanish, Portuguese, Italian, Greek, Hungarian and Turkish. Has demonstrated good psychometrics and similar factor structure in these languages. The Indigenous-adapted FACES-II	Enmeshed and rigid scales have weaker psychometric performance. May not reflect Indigenous concepts of family and parenting.

Family functioning tools: psychometrics							
Source	Reliability/ internal consistency	Factor structure/ construct validity	Convergent/ concurrent validity	Criterion validity/ clinical discrimination	Test-retest/ responsiveness	Cultural validation	Key limitations
Martínez-Pampliega, Merino (2017) Gomes, Peixoto (2019) Sequeira, Vicente (2021) Türkdoğan, Duru (2018) Waterman, Wheeler (2026)						showed limited validity for family cohesion and caregiver-child communication.	
Circumplex Model Family Adaptability and Cohesion Evaluation Scales (FACES-IV; 24-items) Hamilton and Carr (2016)	Acceptable reliability for most scales; rigid scale weaker.	Circumplex model supported.	Not reported.	Not reported.	Not reported.	Not reported.	Sparse validation data compared with full version.
The McMaster Family Assessment Device (FAD)	Good to excellent reliability (General functioning subscale $\alpha=.95$;	Six-domain model supported, although overlap between domains has been reported.	Moderate expected correlations with family functioning measures: MCRS General	Distinguishes clinical and non-clinical families.	Strong evidence overall.	Translated into Spanish and French. Spanish adaption was reliable.	Ongoing concerns regarding factorial validity.

Family functioning tools: psychometrics							
Source	Reliability/ internal consistency	Factor structure/ construct validity	Convergent/ concurrent validity	Criterion validity/ clinical discrimination	Test-retest/ responsiveness	Cultural validation	Key limitations
Hamilton and Carr (2016) The National Child Traumatic Stress Network (2026)	dimension scales $\alpha=.70-.81$).		Functioning Scale ($r= -.40$ to $-.60$).	Low associations with social desirability.			
General Functioning – Family Assessment Device (GF-FAD) Sianko and McDonell (2020) The National Child Traumatic Stress Network (2026) Assaf, Hallit (2025)	Good internal consistency ($\alpha=.78-.88$).	One-factor structure and measurement invariance supported across caregivers and adolescents.	Correlated with clinician ratings ($r\approx .56$) and marital functioning ($r\approx .69$); adolescent–caregiver agreement weak ($r=.11-.15$, $p < .01$).	Differentiates clinical and non-clinical groups.	Detects change over time ($t=3.37$, $p<.01$; $d=.25$).	Translated into Spanish, French, and Arabic. Spanish adaption was reliable. Arabic version demonstrated excellent internal consistency, acceptable convergent validity, and concurrent validity.	Weak cross-informant agreement.
Brief Assessment of Family Functioning Scale (BAFFS) Mansfield, Keitner (2019)	Acceptable reliability ($\alpha=.71$).	Factorial invariance demonstrated across spouses and families.	Strong correlation with GF-FAD ($r=.87$); moderate correlation with the Dyadic Adjustment Scale-4 ($r=.66$); moderate	Differentiates clinical and non-clinical groups.	Detects change over time ($t=3.06$, $p<.01$; $d=.23$).	Translated into Spanish and French. Spanish adaption was reliable.	Brevity may reduce depth of assessment.

Family functioning tools: psychometrics							
Source	Reliability/ internal consistency	Factor structure/ construct validity	Convergent/ concurrent validity	Criterion validity/ clinical discrimination	Test-retest/ responsiveness	Cultural validation	Key limitations
The National Child Traumatic Stress Network (2026)			correlation with clinician ratings on MCRS General Functioning ($r=-.50$).				
Family Assessment Measure III (FAM III) Hamilton and Carr (2016) Family Well-Being Research Network (2026)	Excellent reliability for overall scales; weaker for some subscales.	Original 7-factor model inconsistently replicated; one factor solution also reported.	Moderate to strong correlations with family functioning measures.	Distinguishes problem and non-problem families.	Less than adequate test-retest reliability.	Translated into Danish, Italian, French and Spanish.	Unstable factor structure; limited evidence of responsiveness to change.
Family Environment Scale (FES) Hamilton and Carr (2016) Mind Garden (2026)	Generally poor internal consistency for full and short versions.	Original 10-scale structure inconsistently replicated. The factor structure of the shorter version did not fit the original model.	Not reported.	Five subscales differentiate clinical and non-clinical groups.	Not reported.	Translated into Arabic, Chinese Cantonese, Chinese Mandarin, Creole Haiti, Danish, Dutch, Farsi, Finnish, French, Greek, Hebrew, Hindi, Italian, Japanese, Korean, Malay, Norwegian, Polish, Portuguese,	Poor reliability and unstable factor structure.

Family functioning tools: psychometrics							
Source	Reliability/ internal consistency	Factor structure/ construct validity	Convergent/ concurrent validity	Criterion validity/ clinical discrimination	Test-retest/ responsiveness	Cultural validation	Key limitations
						Slovenian, Spanish, Swedish, Tagalog and Thai. No validation studies found.	
Family Functioning and Cohesion Scale (FFCS) Tsoutsi and Dikeos (2024)	Good internal consistency ($\alpha=.85-.88$).	Three-factor structure: communication; anger/resentment/ aggression; values/beliefs.	Strong correlation with GF-FAD ($r=.83$, $p<.001$).	Not reported.	Excellent test- retest reliability ($r=.88$).	No translations or cultural validations reported.	Validation sample predominantly female and highly educated.
Mayi Kuwayu Family Functioning Measure (FFM) Brinckley, Jones (2022)	Excellent internal consistency ($\alpha=.905-.906$).	Unidimensional structure supported. All items showed significant discrimination ($p<.001$).	Better family functioning associated with greater financial security (PR=1.34- 1.61) and less pain (PR=0.77-0.91).	No association with heart disease.	Not reported.	Developed for and validated in Aboriginal and Torres Strait Islander peoples.	Some evidence of item redundancy.
Family Health Scale (FHS) Crandall, Weiss- Laxer (2020)	Long form $\alpha=.82$ - .92. Short form $\alpha=.80$ - .84.	Four-factor structure supported.	Strong correlations with FAD ($r=.89$) and Family Health Climate Scale ($r=.66$), $p < 0.001$.	Better family health associated with lower odds of depression (OR=.07-.28, $p<.001$).	Not reported.	No translations or cultural validations reported.	Developed and validated in U.S.A. populations.

Family functioning tools: psychometrics							
Source	Reliability/ internal consistency	Factor structure/ construct validity	Convergent/ concurrent validity	Criterion validity/ clinical discrimination	Test-retest/ responsiveness	Cultural validation	Key limitations
Family Resilience Scale for Veterans (FRS-V) Finley, Pugh (2016)	Excellent reliability ($\alpha=.90$).	Condensed three-factor/six-construct model supported.	Moderate to strong correlations with resilience, coping, social support, PTSD and depression measures ($r=.45-.67$, $p < .01$).	No differences by employment/service component.	Not reported.	No translations or cultural validations reported.	Developed specifically for veteran populations.
Protective Factors Survey, 2nd Edition (PFS-2) Sprague-Jones, Counts (2019) Sprague-Jones, Bash (2023)	Good reliability across domains ($\alpha=.77-.88$).	Five-factor structure supported.	Evidence from stakeholder input and factor analysis.	Not reported.	Designed to be sensitive to change.	Translated into Spanish. Spanish version demonstrated good reliability and validity, and psychometrically similar to the English version.	Limited validation.
Beavers Systems Model Self-Report Family Inventory (SFI) Hamilton and Carr (2016) Saleh (2026)	Good internal consistency.	Mixed evidence; proposed 5-, 6-, and 2-factor solutions.	Moderate correlations with self-report and clinician-rated family functioning measures.	Distinguishes clinical populations.	Mixed test-retest findings.	Translated into Persian. The Persian version demonstrated good reliability and validity.	Unstable factor structure; mixed test-retest evidence.

Family functioning tools: psychometrics							
Source	Reliability/ internal consistency	Factor structure/ construct validity	Convergent/ concurrent validity	Criterion validity/ clinical discrimination	Test-retest/ responsiveness	Cultural validation	Key limitations
Systemic Clinical Outcome and Routine Evaluation (SCORE) Hamilton and Carr (2016) Carr and Stratton (2017) Association for Family and Systemic Psychotherapy (2026) Vilaça, Relvas (2018) Zetterqvist, Hånell (2020) Limsuwan and Prachason (2020)	Adequate to excellent reliability across versions.	Stable 3-factor structure for SCORE-40, SCORE-28, SCORE-15: strengths, difficulties, communication.	The 15-item, 28-item and 40-item versions of the SCORE have demonstrated moderate correlations with family functioning, mental health, quality of life and maltreatment measures; low to moderate correlations with clinician-rated measures of family functioning. SCORE-28 and SCORE-15 correlated with indices of global functioning, service needs, and recollections of child maltreatment.	Differentiates clinical and non-clinical populations. SCORE-15 and SCORE-28 were not correlated with demographic characteristics or social desirability response set.	Good test-retest reliability has been established for the 15-item and 28-item, 40-item and child versions. SCORE 15- and 28-item versions demonstrate responsiveness to therapeutic change.	SCORE-15 has been translated into Albanian, Amharic, Arabic, Bosnian, Chinese, Croatian, Czech, Dari, Farsi, Finnish, French, German, Greek, Hindi, Hungarian, Icelandic, Italian, Lingala, Maltese, Netherlands and Flemish Belgian, Norwegian, Pashto, Polish, Portuguese, Romanian, Serbian, Somali, Sorani, Spanish, Swedish, Tamil, Thai, Transylvanian Hungarian, Turkish, and Welsh. Portugues, Swedish and Thai versions demonstrated good	Some inconsistency in adult mental health samples.

Family functioning tools: psychometrics							
Source	Reliability/ internal consistency	Factor structure/ construct validity	Convergent/ concurrent validity	Criterion validity/ clinical discrimination	Test-retest/ responsiveness	Cultural validation	Key limitations
						reliability and validity.	
Systemic Therapy Inventory of Change (STIC) Hamilton and Carr (2016) Brendel, Hartmann (2026)	Adequate to good reliability across scales.	Factor structure supported.	Strong correlations with established measures on four of five scales.	Distinguishes clinical and non- clinical populations; cut-offs established.	Good test-retest reliability.	Translated into German. German version demonstrated good validity and reliability.	Limited evidence of responsiveness to change.
Western Australian Aboriginal Child Health Survey (WAACHS) Brinckley, Jones (2022)	Not reported.	One-factor structure supported.	Strong correlations with associations with poor family functioning and financial strain, alcohol use, not having someone to yarn to (converse respectfully in a safe space), and carer relationship issues.	No significant associations with physical health conditions.	Not reported.	Developed for and validated in Aboriginal peoples.	Limited validation studies.

Appendix

Methods

Critical Intelligence Unit (CIU) evidence checks are not intended to be exhaustive systematic reviews (multiple databases, formal critical appraisal, etc) but instead rapid, responsive evidence summaries:⁴⁶

- Search terms for PubMed are developed by CIU team and checked by requesting team.
- Literature is restricted to the highest levels of evidence available for a particular topic.
- Single reviewer screening and data extraction, with consultation in case of any uncertainty.
- Review of evidence check by: CIU Manager and Executive Lead, requesting team, clinical expert advisory group (as appropriate), and at least one external peer reviewer.

Data extraction was completed in a separate Excel file which is available upon request.

CIU evidence checks include searching for literature specific to Aboriginal and Torres Strait Islander people to highlight any relevant literature or gaps in the literature as one means to work towards reducing the gap between Aboriginal and non-Aboriginal people.

PubMed search terms

("family function"[Title/Abstract] OR "family cohes"[Title/Abstract] OR "family assessment"[Title/Abstract] OR "family needs assessment"[Title/Abstract] OR "parenting capacity assessment"[Title/Abstract]) AND ("measure"[Title/Abstract] OR "tool"[Title/Abstract] OR "scale"[Title/Abstract] OR "instrument"[Title/Abstract] OR "questionnaire"[Title/Abstract]) AND ("valid"[Title/Abstract] OR "reliab"[Title/Abstract] OR "psychometric"[Title/Abstract]) AND ("english"[Language] AND 2016/01/01:2026/12/31[Date - Publication]) AND "english"[Language]

461 hits on 15 June 2026

"North Carolina Family Assessment Scale"[Title/Abstract] –

0 hits – Need to find a validations study through targeted google search

Aboriginal health lens search terms

Text AND ("indigenous"[Title/Abstract] OR "aboriginal"[Title/Abstract] OR "first nation"[Title/Abstract] OR "native"[Title/Abstract])

7 hits on 15 June 2026

Google search terms

Family functioning, psychometric, validation.

Inclusion and exclusion criteria

Inclusion	Exclusion
• Published in English • Published since 2016 • Population: Families with children and complex parental mental health and drug & alcohol needs and coexisting child safety concerns, often in the context of: Trauma; Domestic and family	• Published prior to 2016 • Studies that do not meet PICOS criteria • Specifically developed for families impacted by chronic

Inclusion	Exclusion
violence; Parental drug and alcohol misuse; Mental health problems; Parents with significant learning difficulties and/or intellectual disabilities • Intervention: Family functioning assessment tools measuring how well parents/carers can meet their children's developmental, emotional, behavioural, and safety needs. Includes relationships, impact of trauma, violence and substance use across the family system, and relationships with external systems such as DCJ, health, NGO's. – Tools intended for parents that have part/s that are rated/completed by children/adolescents can be considered for inclusion – Studies on developing or adapting the tools in languages other than English in English speaking countries • Comparison: other tools • Outcomes: key features and implementation considerations including: – target population – domains or dimensions – psychometric properties – strengths and limitations – practical implementation considerations such as the time required to complete the tool and whether it can be administered online or on paper – cultural or language adaptations including validity and appropriateness • Study types: – Studies reporting on psychometric evaluations of measurement tools – Review studies – Grey literature such as guidelines and consensus statements • Setting: – NSW, Australia and international jurisdictions with healthcare systems similar to NSW Health (UK, Canada, New Zealand, European/Asian high-income countries with public healthcare systems), US – Family assessment and child welfare contexts	diseases other than mental health conditions or cancer, affecting either the parent or the child • Studies that employ family functioning assessment tools to measure outcomes but do not report on the tools' development, measurement properties, or validation • Tools that measure marital functioning • Tools that are exclusively intended for data collection from children rather than parents • Studies that describe family functioning concepts, models or theories only • Studies on developing or adapting the tools in languages other than English in non-English speaking countries • Letters, comments, editorials, study protocols, conference abstracts

References

1. Reed-Ashcraft K, Kirk RS, Fraser MW. The reliability and validity of the North Carolina Family Assessment Scale. *Research on Social Work Practice* 2001;11(4):503-20. DOI: 10.1177/104973150101100406
2. Kirk RS. Development, intent, and use of the North Carolina Family Assessment Scales, and their relation to reliability and validity of the scales. Buhl: National Family Preservation Network; 2015 [cited 19 Jun 2026]. Available from: https://www.nfpn.org/media/gp2nc5bd/ncfas_scale_development.pdf
3. Kirk RS. North Carolina Family Assessment Scale for Reunification: research report. Buhl: National Family Preservation Network; 2002 [cited 19 Jun 2026]. Available from: https://www.nfpn.org/media/dkrkukvv/ncfas-r_research_report.pdf
4. Good A, MacKeith J. Assessing family functioning: psychometric evaluation of the Family Star Plus. *Family Relations*. 2021;70(2):529-39. DOI: 10.1111/fare.12488
5. Kerr S, Hudenko WJ, Godfrey DA, et al. Validation of the Emotional Tone Index for Families (ETIF): a multi-Informant measure of emotional closeness. *Family Process*. 2021;60(3):935-49. DOI: 10.1111/famp.12608
6. Hamilton E, Carr A. Systematic review of self-report family assessment measures. *Family Process*. 2016;55(1):16-30. DOI: 10.1111/famp.12200
7. Sianko N, McDonell JR. Is family functioning stable and consistent over time and stakeholders? A comparison of adolescents' and caregivers' views. *Children and Youth Services Review*. 2020;113(1):104996. DOI: 10.1016/j.childyouth.2020.104996
8. Mansfield AK, Keitner GI, Sheeran T. The Brief Assessment of Family Functioning Scale (BAFFS): a three-item version of the General Functioning Scale of the Family Assessment Device. *Psychotherapy Research*. 2019;29(6):824-31. DOI: 10.1080/10503307.2017.1422213
9. Tsoutsis V, Dikeos D. Family Functioning and Cohesion Scale: validation of a short instrument for the assessment of intrafamily relations. *Behavioral Sciences*. 2024;14(10):969. DOI: 10.3390/bs14100969
10. Brinckley MM, Jones R, Batterham PJ, et al. The development and validation of a family functioning measure for Aboriginal and Torres Strait Islander adults. *BMC Public Health*. 2022;22(1):1976. DOI: 10.1186/s12889-022-14363-7
11. Crandall A, Weiss-Laxer NS, Broadbent E, et al. The Family Health Scale: reliability and validity of a short- and long-form. *Frontiers in Public Health*. 2020;8(1):587125. DOI: 10.3389/fpubh.2020.587125
12. Finley EP, Pugh MJ, Palmer RF. Validation of a measure of family resilience among Iraq and Afghanistan veterans. *Military Behavioral Health*. 2016;4(3):205-19. DOI: 10.1080/21635781.2016.1153530
13. Sprague-Jones J, Counts J, Rousseau M, et al. The development of the protective factors survey, 2nd edition: a self-report measure of protective factors against child maltreatment. *Child Abuse & Neglect*. 2019;89(1):122-34. DOI: 10.1016/j.chiabu.2019.01.008
14. Carr A, Stratton P. The Score Family Assessment Questionnaire: a decade of progress. *Family Process*. 2017;56(2):285-301. DOI: 10.1111/famp.12280
15. Outcomes Star. Using the stars. Brighton: Outcomes Star; 2026 [cited 17 Aug 2026]. Available from: <https://outcomesstar.org/using-the-stars/>
16. The National Child Traumatic Stress Network. Family adaptability and cohesion scale. Los Angeles: The National Child Traumatic Stress Network; 2026 [cited 17 Aug 2026]. Available from: <https://www.nctsn.org/measures/family-adaptability-and-cohesion-scale>
17. Martínez-Pampliega A, Merino L, Iriarte L, et al. Psychometric properties of the Spanish version of the Family Adaptability and Cohesion Evaluation Scale IV. *Psicothema*. 2017;29(3):414-20. DOI: 10.7334/psicothema2016.21
18. Gomes HMS, Peixoto F, Gouveia-Pereira M. Portuguese validation of the Family Adaptability and Cohesion Evaluation Scale – FACES IV. *Journal of Family Studies*. 2019;25(4):477-94. DOI: 10.1080/13229400.2017.1386121

19. Sequeira J, Vicente HT, Daniel F, et al. Family Adaptability and Cohesion Evaluation Scale – Version IV (FACES IV): validation study in the Portuguese population. *Journal of Child and Family Studies*. 2021;30(7):1650-63. DOI: 10.1007/s10826-021-01941-3
20. Türkdoğan T, Duru E, Balkıs M. Turkish adaptation of the Family Adaptability and Cohesion Scale IV. *International Journal of Assessment Tools in Education*. 2018;5(4):631-44. DOI: 10.21449/ijate.409110
21. The National Child Traumatic Stress Network. Family Assessment Device. Los Angeles: The National Child Traumatic Stress Network; 2026 [cited 17 Aug 2026]. Available from: <https://www.nctsn.org/measures/family-assessment-device>
22. Assaf M, Hallit R, Malaeb D, et al. Psychometric properties of the Arabic version of the McMaster family assessment device–general functioning subscale (FAD-GF) among a sample of Lebanese adults. *Cogent Psychology*. 2025;12(1):2540065. DOI: 10.1080/23311908.2025.2540065
23. Family Well-Being Research Network. Family Assessment Measure Version III (FAM III). Cambridge: Family Well-Being Research Network; 2026 [cited 17 Aug 2026]. Available from: <https://fam-net.org/product/family-assessment-measure-iii/>
24. Mind Garden. Family Environment Scale. Menlo Park: Mind Garden Inc; 2026 [cited 18 Aug 2026]. Available from: <https://www.mindgarden.com/96-family-environment-scale#horizontalTab1>
25. Sprague-Jones J, Bash K, Singh P, et al. The Spanish protective factors survey, second edition: initial translation and preliminary validation. *Children and Youth Services Review*. 2023;155(1):107204. DOI: 10.1016/j.childyouth.2023.107204
26. Saleh B. Reconstruction, validation, making reliable and normalization of Self-Reported Family Inventory (SFI). *Educational Measurement*. 2026;17(64):37-64.
27. Association for Family and Systemic Psychotherapy. Evaluating family therapy: SCORE-15. Cheshire: Association for Family and Systemic Psychotherapy; 2026 [cited 18 Aug 2026]. Available from: <https://www.theafsp.org/publications-resources/evaluating-family-therapy-score-15/>
28. Vilaça M, Relvas AP, Stratton P. A Portuguese translation of the Systemic Clinical Outcome and Routine Evaluation (SCORE): the psychometric properties of the 15- and 28-item versions. *Journal of Family Therapy*. 2018;40(4):537-56. DOI: 10.1111/1467-6427.12197
29. Zetterqvist M, Hånell HE, Wadsby M, et al. Validation of the Systemic Clinical Outcome and Routine Evaluation (SCORE-15) self-report questionnaire: index of family functioning and change in Swedish families. *Journal of Family Therapy*. 2020;42(1):129-48. DOI: 10.1111/1467-6427.12255
30. Limsuwan N, Prachason T. The reliability and validity of the 15-item Systemic Clinical Outcome and Routine Evaluation (SCORE-15) Thai version. *Journal of Family Therapy*. 2020;42(1):119-28. DOI: 10.1111/1467-6427.12248
31. Brendel A, Hartmann M, Haun MW, et al. Translation and validation of the German version of the Systemic Inventory of Change. *Frontiers in Psychiatry*. 2026;16(1):1686468. DOI: 10.3389/fpsyt.2025.1686468
32. Waterman EA, Wheeler L, Edwards KM, et al. Measures to assess efficacy of family-based programs for Indigenous populations: adaptation, psychometric properties, and congruence. *Family Process*. 2026;65(1):e70118. DOI: 10.1111/famp.70118
33. California Evidence-Based Clearinghouse. North Carolina Family Assessment Scale - General + Reunification (NCFAS-G+R). San Diego: California Evidence-Based Clearinghouse; 2026 [cited 17 Aug 2026]. Available from: <https://www.cebc4cw.org/measurement-tools/north-carolina-family-assessment-scale-general-reunification/>
34. NSW Ministry of Health. Whole Family Team guideline. St Leonards: NSW Ministry of Health; 2020.
35. Westerman T. Arguing the need for Aboriginal-specific testing of attachment theory as a foundation of culturally informed child protection and attachment best practice. *Clinical Psychologist*. 2026;1-18. DOI: 10.1080/13284207.2026.2682255
36. National Family Preservation Network. Selecting and assessment tool. Asheville: National Family Preservation Network; 2026 [cited 17 Aug 2026]. Available from: <https://www.nfpn.org/assessment-tools/how-to-choose/#step1>

37. Skinner H, Steinhauer P, Santa-Barbara J. Family Assessment Measure–III (FAM-III. Dinas Powys: The Cognitive Centre; 2016 [cited 17 Aug 2026]. Available from: <https://www.cognitivecentre.com/assessment/family-assessment-measure-iii-fam-iii/>
38. Australian Institute of Family Studies. Families and children outcomes measurement matrix. Southbank: Australian Institute of Family Studies; 2024 [cited 17 Aug 2026]. Available from: <https://aifs.gov.au/resources/resource-sheets/families-and-children-outcomes-measurement-matrix-explanatory-notes>
39. Family Well-Being Research Network. Family Health Scale. Cambridge: Family Well-Being Research Network; 2026 [cited 17 Aug 2026]. Available from: <https://fam-net.org/product/family-health-scale/>
40. Family Well-Being Research Network. Family Resilience Scale for Veterans (FRS-V). Cambridge: Family Well-Being Research Network; 2026 [cited 17 Aug 2026]. Available from: <https://fam-net.org/product/family-resilience-scale-for-veterans-frs-v/>
41. California Evidence-Based Clearinghouse. Protective Factors Survey, 2nd Edition (PFS-2). San Diego: California Evidence-Based Clearinghouse; 2026 [cited 18 Aug 2026]. Available from: <https://www.cebc4cw.org/measurement-tools/protective-factors-survey-2nd-edition/>
42. He Y, Zinbarg RE, Goldsmith JZ, et al. Sensitivity to change of the Systemic Therapy Inventory of Change (STIC) intersession scales. *Psychotherapy Research*. 2024;34(4):461-74. DOI: 10.1080/10503307.2023.2244150
43. The Family Institute at Northwestern University. Report on the Systemic Therapy Inventory of Change (STIC) feedback system. Evanston: The Family Institute at Northwestern University; 2017 [cited 19 Aug 2026]. Available from: https://www.family-institute.org/sites/default/files/report_on_stic_feedback_system_dec_2017.pdf
44. Kirk RS. Psychometric properties of the trauma and post-trauma well-being assessment domains of the North Carolina Family Assessment Scale for General and Reunification Services (NCFAS G+R). *Journal of Public Child Welfare*. 2015;9(5):444-62. DOI: 10.1080/15548732.2015.1090364
45. Kirk RS, Martens P. An examination of the cultural competence and inter-rater reliability of the North Carolina Family Assessment Scales (NCFAS-R and NCFAS-G+R). Asheville: National Family Preservation Network; 2012 [cited 11 Sep 2026]. Available from: <https://www.nfpn.org/resources/>
46. Williamson L, McArthur E, Dolan H, et al. Horizon scanning, rapid reviews and living evidence to support decision-making: lessons from the work of the Critical Intelligence Unit in New South Wales, Australia during the COVID-19 pandemic. *BMJ Open*. 2023;13(5):e071003. DOI: 10.1136/bmjopen-2022-071003