



Capturing Capacity for Self-Sufficiency

A Multidimensional Framework and Measurement Tools



March 2024

March 2024

**Authors: Kate Stepleton, Stephen Nuñez,
Sophie Hearn, and Claire McMahon-Fishman**

Submitted to:

Tennessee Department of Human Services

Tennessee Tower
312 Rosa Parks Boulevard, 17th Floor
Nashville, TN 37243

Project Director: Mike Fishman

MEF Associates
1330 Braddock Place, Suite 220
Alexandria, VA 22314

Suggested citation: Stepleton, K., Nuñez, S., Hearn, S., & McMahon-Fishman, C. (2024). Capturing Capacity for Self-Sufficiency: A Multidimensional Framework and Measurement Tools. Alexandria, VA: MEF Associates.

Disclaimer: The views expressed in this publication do not necessarily reflect the views or policies of the Tennessee Department of Human Services.

Acknowledgements: The authors thank the staff who contributed to this report, including Olivia Mirek and Kia Heryadi. Additionally, we thank the many staff who provided valuable feedback on drafts of the report: Mike Fishman, Kimberly Foley, Carly Morrison, Ashweeta Patnaik, Valerie Benson, Erin Bumgarner, LaDonna Pavetti, and Laura Peck.



Table of Contents

Table of Contents	3
Executive Summary	i
Development of a Conceptual Framework of Capacity for Self-Sufficiency	i
Identifying Methods to Measure Capacity and Use of Safety Net Programs	ii
Recommendations	iii
Next Steps	iii
Overview of Recommendations	iv
Introduction	1
Overview of Approach to the Outcomes Review	2
Review of Traditional Measures of Self-Sufficiency and Use of Safety Net Programs.....	2
Critiques of Traditional Measures of Self-Sufficiency and Dependency	5
Methods	7
Development of the Conceptual Framework.....	8
Key Informant Interviews.....	8
Administrative Data	8
Identifying Surveys, Scales, and Indices	8
Findings.....	9
Conceptual Framework	9
Insights from Key Informant Interviews	11
Measures Derived from Administrative Data	12
Surveys, Scales, and Indices.....	12
Recommendations	13
Conclusion and Next Steps.....	24
Appendix A: References Informing the Conceptual Framework	25

Appendix B: Glossary of Key Terms	29
Appendix C: List of Search Terms for Surveys, Scales, and Indices	30
Appendix D: Literature Review Process Graphic	31
Appendix E: Tables of Surveys, Scales, and Indices	32
Economic Well-Being Measures.....	32
Social Capital Measures	34
Skills, Knowledge, and Experience Measures	38
Mental Well-Being Measures	39
Physical Well-Being Measures	56
Family Well-Being Measures	61
Cross-Domain Measures	87

Executive Summary

The TANF Opportunity Act (TOA) authorizes seven pilot programs to support families with low incomes to achieve and sustain self-sufficiency. All seven pilots are providing an array of services with the aim of improving individual and family outcomes. Each program is operating with a broad theory of change: if services can increase participants' capacity to be self-sufficient, participants will be less likely to need government assistance from safety net programs like TANF. Each pilot is participating in a rigorous, independent evaluation of its approach.

Many studies of self-sufficiency (or the capacity for self-sufficiency) and use of safety net programs focus on a narrow set of measures related to income and benefits receipt. Scholars have argued that this approach neglects the role that individual factors play in achieving self-sufficiency, like mental health, executive functioning, skills, and education. It also discounts the influence of interpersonal, community, and institutional factors that exist outside the individual. To understand how services can support self-sufficiency and reduce individual and family need for safety net programs, a more expansive set of outcomes is needed.

In support of the TOA pilot evaluations, MEF Associates conducted a literature review to identify potential outcome measures to tell the stories of the pilots and their participants, capturing the complexity of their efforts and impacts. This effort was twofold: it involved exploring outcomes related to capacity for self-sufficiency and outcomes related to use of safety net programs. In both cases, the team sought to expand beyond traditional, narrow measures of self-sufficiency and benefits receipt, in order to encompass the range of impacts programs might achieve as they promote participants' economic stability and well-being.

MEF's team developed recommendations for a set of outcome measures to more fully describe the impact of services on individual and family capacity for self-sufficiency, as well as use of safety net programs. The team scanned nearly 300 sources and gathered input from five experts from the field. While these recommendations were developed to support the TOA pilot evaluations, findings may be applied more broadly to efforts to promote capacity for self-sufficiency among families with low income.

Development of a Conceptual Framework of Capacity for Self-Sufficiency

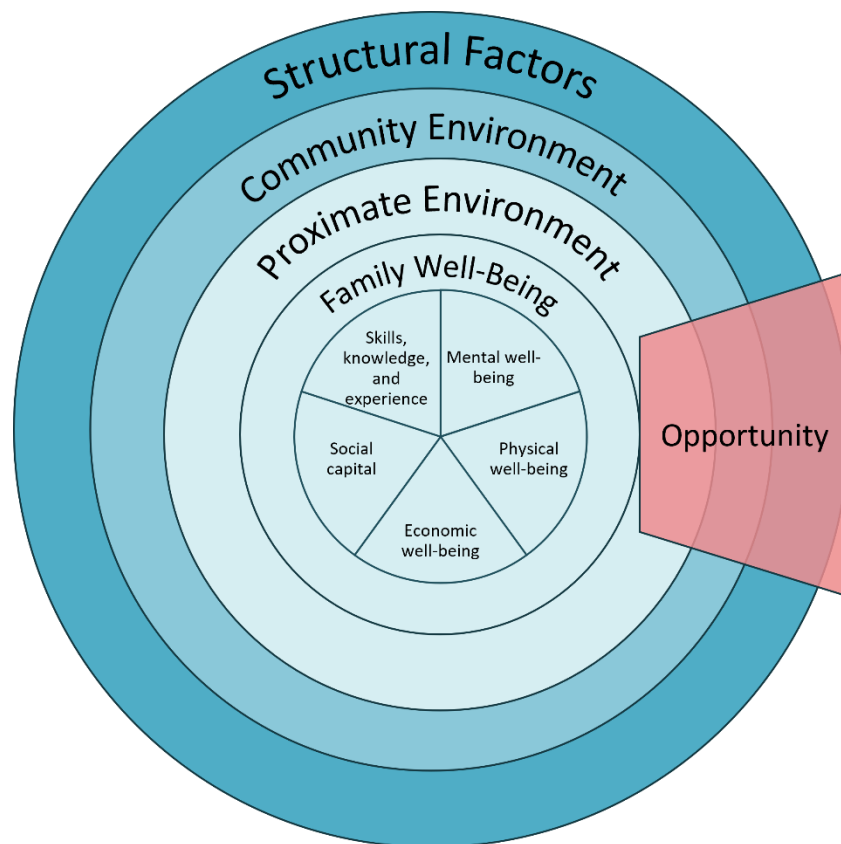
“Capacity for self-sufficiency” can be broadly understood as a combination of factors both internal and external to individuals that position them to successfully support themselves and their dependents. To organize the information in the literature review, the team developed a conceptual framework describing the factors that make up capacity for self-sufficiency, as described in scholarly literature. This literature highlights several key concepts about self-sufficiency:

A note on language

This report concerns individual and family needs for safety net programs like TANF, SNAP, or SSI to make ends meet. While the term “dependency” has long been used in policy and appears in legislation, including the TANF Opportunity Act, this report uses the more precise phrase, “use of safety net programs” or “benefits receipt.”

- Capacity for self-sufficiency is a multidimensional construct, made up of factors in multiple domains.
- Capacity for self-sufficiency is influenced by factors at multiple levels, including the individual, family, community, and structural levels.
- There are dynamic relationships between and among domains of capacity for self-sufficiency, both within individuals and across levels of influence.
- An essential element of capacity is opportunity; without opportunities in one’s environment to access, factors like skills, knowledge, and experience cannot be translated into self-sufficiency.

The conceptual framework incorporates these concepts and shows that capacity for self-sufficiency is a complex construct; the pilots’ goals are therefore similarly complex. Efforts to increase capacity will look different depending on which domains the pilots target, at which levels. For example, one program might provide evidence-based counseling services to improve participants’ mental health; another might focus on expanding a training program at a local community college so clients can gain skills certifications. Both could positively affect participants’ capacity to achieve self-sufficiency.



Conceptual Framework of Capacity for Self-Sufficiency

Identifying Methods to Measure Capacity and Use of Safety Net Programs

There is no singular measure of capacity that captures all of the different aspects shown in the conceptual framework. Likewise, each pilot provides a unique set of services that focus on different aspects of capacity, making it challenging to capture impacts of these services using a single measure. As such, it is necessary to consider multiple measurement options that can be used together, tailored

to the pilots' designs, to understand how each of the seven pilots are aiming to influence participants' capacity for self-sufficiency. To that end, the team systematically searched for and identified a selection of measures drawn from surveys and administrative data aligned with the varied elements of the conceptual framework.

The team also explored measures of use of safety net programs that might inform understanding of the effects of the pilots. For these measures, the team used the scholarly literature and the codebook for the P20 Connect TN (P20) database as sources for what could be measured regarding use of safety net programs.¹

Recommendations

Based on its assessment of the literature, the team developed five broad recommendations for TDHS to inform evaluation of the TOA pilots, as well as future assessment of the effects of TANF-funded programs on participants' well-being. Recommendations are listed in the following text box, and a more detailed overview is provided on the next page.

Recommendation 1: Use administrative data sources to track characteristics of benefits receipt, educational attainment, and employment.

Recommendation 2: Capture additional information about family finances and job quality through surveys.

Recommendation 3: Use well-tested scales with demonstrated reliability and validity among similar populations to capture participants' subjective assessments of well-being in multiple individual and family domains.

Recommendation 4: Examine variability in opportunity in the communities where participants live and work and, if necessary, include indicators of opportunity as covariates in outcome or impact analyses.

Recommendation 5: Develop logic models to explore the relationship between program services and participant outcomes that account for the complex, dynamic nature of capacity for self-sufficiency.

Next Steps

MEF has worked with the TOA pilots to develop evaluation plans and is mapping outcomes tailored to each design. The evaluation team will use findings from this report to shape the selection of outcomes that can be accessed either in administrative data or with participant surveys. Additionally, MEF will use recommendations from this report to inform its approach to assessing multidimensional poverty among study participants.

¹ P20 Connect TN is Tennessee's integrated, longitudinal database that links data from K-12 education, higher education, social services, and workforce development.

Overview of Recommendations

Recommendation 1: Use administrative data sources to track characteristics of benefits receipt, educational attainment, and employment.

TDHS has access to a rich, longitudinal database of linked administrative records from seven state agencies (P20). The Department can use this information to construct indicators related to individuals' economic well-being and their skills, knowledge, and experience. P20 data can also be used to create measures describing individuals' use of safety net programs.

Recommendation 2: Capture additional information about family finances and job quality through surveys.

A participant survey can complement administrative data to provide a more complete picture of a family's financial situation and individual employment circumstances. Several well-known, large-sample surveys include questions that can be used or adapted to capture this information.

Recommendation 3: Use well-tested scales with demonstrated reliability and validity among similar populations to capture participants' subjective assessments of well-being in multiple individual and family domains.

A participant survey should also include indicators of how participants are faring across multiple domains. Drawing from a list of recommended surveys, scales, and indices, follow-up surveys should be tailored to each program's design, incorporating measures from domains most likely to be affected by the intervention.

Recommendation 4: Examine variability in opportunity in the communities where participants live and work and, if necessary, include indicators of opportunity as covariates in outcome or impact analyses.

The context in which individuals and families live has an enormous impact on their ability to build capacity for self-sufficiency. It is important to track indicators of opportunity in participants' communities and utilize those measures in evaluations, both to assure comparability across study groups and to contextualize findings.

Recommendation 5: Develop logic models to explore the relationship between program services and participant outcomes that account for the complex, dynamic nature of capacity for self-sufficiency.

The conceptual framework guiding this review highlights the complex, dynamic nature of capacity to achieve self-sufficiency. While each pilot aims to increase this capacity, every program has its own approach, targeting different domains of capacity in distinct ways. Developing a logic model that clearly describes how pilot services affect different aspects of capacity will help pilots identify appropriate outcomes. This in turn will allow the evaluations to detect the range of impacts the pilots achieve.

Capturing Capacity for Self-Sufficiency: A Multidimensional Framework and Measurement Tools

Introduction

Tennessee’s TANF Opportunity Act (TOA) aims to enhance self-sufficiency among Tennesseans with low incomes through increased cash transfers to Temporary Assistance for Needy Families (TANF) participants and expanded investments in community partnerships. TOA funding comes from the state’s substantial reserve of TANF dollars. In addition to providing more and better support to Tennesseans with low incomes, the Act is also an opportunity to build evidence about how programs can help families achieve and maintain self-sufficiency.

TOA authorizes seven pilot programs providing “comprehensive supports to families as they move from crisis to career pathways.”² Each pilot involves multiple community agencies collaborating to assess individual and family needs and provide a tailored array of services with the shared goal of promoting self-sufficiency. The seven pilots are participating in rigorous evaluations of their approaches conducted by MEF Associates, an independent evaluator.

While each pilot’s approach is unique, all seven pilots are delivering services intended to support families with low incomes to achieve self-sufficiency. The construct of self-sufficiency is often defined in research literature in opposition to dependency. In the context of public programs, dependency is often defined narrowly as receipt of public benefits that allow families to reach a minimum level of subsistence that is not attainable without that assistance. While the term “dependency” appears in legislation, including the TOA, this report uses more precise phrases: “use of safety net programs” or “benefits receipt.” Self-sufficiency is more broadly defined as the sustained ability to meet one’s needs and the needs of one’s dependents without public assistance. Self-sufficiency is based on a range of factors. These include individual factors, like skills, education,

Capacity, self-sufficiency, and use of safety net programs: A note about the terms used in this report

In this report, the terms **capacity**, **self-sufficiency**, and **use of safety net programs** refer to separate but related concepts, all connected to individual and family **well-being**. For the purposes of this report, capacity is understood as **capacity for self-sufficiency**, and **use of safety net program** (and **benefits receipt**) refer to use of TANF, SNAP, and SSI to meet minimum subsistence standards. **Well-being** is an umbrella term encompassing capacity for self-sufficiency, use (or non-use) of safety net programs, and other elements of thriving. These and other definitions are included in this report’s Glossary.

² Tennessee Department of Human Services (n.d.). *TANF Opportunity Act*. Website of the Tennessee Department of Human Services. Retrieved March 14, 2023, from <https://www.tn.gov/humanservices/tanf-opportunity-act.html>

and training; physical health; executive functioning; and social connections. Additionally, contextual factors, like the strength of the labor market, housing costs, transportation infrastructure, and structural racism, play a role in the degree to which individuals can support themselves and their families without government assistance. Typically, when human service programs aim to boost self-sufficiency to reduce use of safety net programs, they evaluate their success by examining process measures that capture information on the delivery of services or benefits (e.g., number of clients served) or a narrow set of outcomes related to income and employment. They may incorporate some outcomes related to individual well-being, such as mental health. Programs may neglect the pivotal role that individual and contextual factors play in making it possible for people to obtain stable employment and/or sustain sufficient income.

Overview of Approach to the Outcomes Review

In its role as evaluator for the TOA pilots, MEF Associates explored potential outcome measures for the evaluations. Tennessee's Department of Human Services (TDHS) charged the project team with identifying outcomes that could capture the range of impacts the pilots might have that could contribute to reducing families' need for safety net programs. This involved examining existing administrative data sources and exploring the research literature for scales, indices, and survey questions that could be administered directly to participants. To guide the literature review, the team developed a conceptual framework illustrating the individual, family, community, and structural factors that contribute to one's capacity to gain self-sufficiency. The team identified constructs and associated measurement tools within the domains in the conceptual framework, concentrating efforts on individual- and family-level domains. Concurrently, the team sought feedback and insights from several key informants who are familiar with the project or who have experience with multidimensional assessments of poverty, self-sufficiency, and well-being. The team then crafted a set of recommendations for TDHS. These recommendations apply specifically to the TOA evaluations, and some may apply more generally to TDHS's work in the state.

The review covered several areas of extensive scholarship, including those related to well-being, mental health assessment, and two-generation programs. Within project time and resource constraints, the team took a systematic approach and developed recommendations that are sound and grounded in rigorous research.

This report begins with a discussion of traditional measures of self-sufficiency and use of safety net programs, presenting critiques of those measures and suggesting proposed remedies. Next, the report describes the outcome review's methods. The results section presents the conceptual framework, shares insights from key informant interviews, and outlines, at a high level, what the project team found. Finally, this report provides recommendations for TDHS, explaining the reasoning for each. An appendix includes tables listing measurement tools that were considered in the search and details about their use and psychometrics (where appropriate).

Review of Traditional Measures of Self-Sufficiency and Use of Safety Net Programs

N.B. This section describes efforts to define welfare dependence and critiques of those approaches. As such, the term "dependency" appears throughout this section to accurately reflect the sources and debates that are represented. Outside of this section, the report uses the more precise phrase, "use of safety net programs," to refer to individuals' and families' receipt of certain government benefits to make ends meet.

The concepts of self-sufficiency and dependency are commonly referenced in the welfare policy sector. These discussions focus on goals for individuals with low incomes engaging in public assistance programs to increase their economic independence and decrease their dependency on government support. The Personal Responsibility and Work Reconciliation Act of 1996 (PRWORA) played an important role in popularizing these concepts. In establishing the Temporary Assistance to Needy Families (TANF) program, PRWORA created mandatory work requirements as a strategy for reducing participants' dependency on welfare.³ As of 2024, one of the four core goals of TANF is “to end the dependency of needy parents on government benefits.”⁴

Policymakers and scholars in the welfare policy domain have traditionally defined self-sufficiency broadly in terms of economic outcomes (i.e., finding a job and leaving welfare).⁵ Federal policy's definition of self-sufficiency has been “that it is obtained through work, that it includes freedom from dependence on government support, and that it strengthens families.”⁶ PRWORA defines self-sufficiency as “the absence of public welfare dependency,” while the Workforce Investment Act of 1998 defines self-sufficiency as “having enough money to meet basic needs,” leaving the specific operationalization of this up to the discretion of the State or Local Workforce Investment Boards.⁷ Experts often use self-sufficiency interchangeably with or alongside other similar terms such as independence, self-reliance, well-being, and capacity, without making clear distinctions between them.⁸ Although some researchers acknowledge that external and non-economic factors contribute to self-sufficiency, discussions of self-sufficiency are often deficit-based, focusing on individuals' psychological and cultural barriers to employment and economic independence.⁹

The related concept of welfare dependency has been defined by the federal government in terms of how long individuals receive public assistance, the amounts they receive, and the number or proportion of individuals receiving assistance. As mandated by the Welfare Indicators Act of 1994, the Secretary of the U.S. Department of Health and Human Services (HHS) submits an annual report to Congress on indicators and risk factors for welfare dependence. The report defines welfare dependency as when a family receives more than 50 percent of their total income in a one-year period from TANF, Supplemental Nutrition Assistance Program (SNAP), and Supplemental Security Income (SSI).¹⁰ This definition allows for variation in families' experiences with safety net programs, including variation in the amount of assistance received and the duration of receipt. The indicators and risk factors from the 2022 report are listed in Table 1 and Table 2.¹¹ It should be noted that these are aggregate, not individual, measures.

³ Robbins, S. M., & Barcus, H. R. (2004). Welfare Reform and Economic and Housing Capacity for Low-Income Households, 1997–1999. *Policy Studies Journal*, 32(3), 439–460.

⁴ Administration of Children and Families, U.S. Department of Health and Human Services (n.d.). *About TANF*. Website of the Administration for Children and Families (USDHHS). Retrieved March 14, 2023, from <https://www.acf.hhs.gov/ofa/programs/tanf/about>

⁵ Leibson Hawkins, R. (2005). From Self-Sufficiency to Personal and Family Sustainability: A New Paradigm for Social Policy. *Journal of Sociology and Social Welfare*, 32, 77., Robbins, S. M., & Barcus, H. R. (2004). Welfare Reform and Economic and Housing Capacity for Low-Income Households, 1997–1999. *Policy Studies Journal*, 32(3), 439–460.

⁶ Gowdy, E. A., & Pearlmuter, S. (1993). Economic Self-Sufficiency: It's Not Just Money. *Affilia*, 8(4), 368–387.

⁷ Hong, P. Y. P. (2013). Toward A Client-Centered Benchmark for Self-Sufficiency: Evaluating the ‘Process’ of Becoming Job Ready. *Journal of Community Practice*, 21(4), 356–378.

⁸ Leibson Hawkins, R. (2005). From Self-Sufficiency to Personal and Family Sustainability: A New Paradigm for Social Policy. *Journal of Sociology and Social Welfare*, 32, 77.

⁹ Daugherty, R. H., & Barber, G. M. (2001). Self-Sufficiency, Ecology of Work, and Welfare Reform. *Social Service Review*, 75(4), 662–675. , Hong, P. Y. P. (2013). Toward A Client-Centered Benchmark for Self-Sufficiency: Evaluating the ‘Process’ of Becoming Job Ready. *Journal of Community Practice*, 21(4), 356–378., Leibson Hawkins, R. (2005). From Self-Sufficiency to Personal and Family Sustainability: A New Paradigm for Social Policy. *Journal of Sociology and Social Welfare*, 32, 77.

¹⁰ United States. Dept. of Health. (1997). Indicators of Welfare Dependence: Annual Report to Congress. US Department of Health and Human Services. This definition was established in 1997 by a bipartisan Advisory Board authorized by the 1994 Act.

¹¹ Crouse, G. (2022). Welfare Indicators and Risk Factors, 21st Report to Congress. U.S. Department of Health and Human Services, Office of Human Services Policy, Office of the Assistant Secretary for Planning and Evaluation.

Table 1. U.S. Department of Human Services 2022 Indicators of Dependence

Indicator	Definition
Degree of Dependence	The proportion of individuals nationally who receive more than half of their income from TANF, SNAP, or SSI.
Program Receipt and Family Labor Force Attachment	In an average month, the percentage of recipients of TANF, SNAP, or SSI living in families with someone participating in the labor force, which includes people working, or unemployed and looking for work.
Program Reciprocity	The percentage of the U.S. population who received or lived with a family member who received cash benefits during the year from TANF or SSI, or benefits from SNAP.
Program Participation Among Those Eligible	The percentage of people who receive a benefit among those that are eligible in an average month.
Multiple Program Receipt	This indicator is measured in two ways: 1) the percentage of the <i>total population</i> that received benefits from more than one of TANF, SNAP or SSI in an average month; and 2) the percentage of <i>recipients</i> of TANF, SNAP, or SSI who received benefits from one or more than one program in an average month.
Program Receipt Duration	The percentage of program recipients getting assistance for a given number of months.

Source: Crouse, G. (2023). Welfare Indicators and Risk Factors, 22nd Report to Congress. U.S. Department of Health and Human Services, Office of Human Services Policy, Office of the Assistant Secretary for Planning and Evaluation.

Table 2. U.S. Department of Human Services 2022 Risk Factors Associated with Benefits Receipt

Risk Factor	Definition
Employment Among the Low-Skilled ¹² Population	The percentage of working age adults (18-65) with a high school education or less who were employed at any time over the calendar year, in either full time or part time work.
Poverty Rates	The percentage of the civilian non-institutionalized population living in families whose income is below the official poverty thresholds which are dependent upon the size of the family. If a family's total income is less than the family's threshold, then that family and every individual in it is considered in poverty. The official poverty definition uses money income before taxes, cash income from TANF and SSI, and excludes noncash benefits.
Food Insecurity	The percentage of households by food security status. Food secure households have consistent, dependable access to enough food for active, healthy living while food-insecure households are those that have difficulty at some time during the year providing enough food for all their members due to a lack of money and other resources. Food insecure households can be further broken down between those having low and very low food security. Very low food security, the more severe situation, applies to households in which food intake for one or more members was reduced and eating patterns were disrupted due to a lack of money or resources.
Nonmarital Teen Births	Birthrates by age for unmarried teens are the ratio of the number of births per 1,000 unmarried teens.

Source: Crouse, G. (2023). Welfare Indicators and Risk Factors, 22nd Report to Congress. U.S. Department of Health and Human Services, Office of Human Services Policy, Office of the Assistant Secretary for Planning and Evaluation.

¹² We acknowledge that the term “low-skilled” is a problematic and deficits-based term, and that “limited work experience” is a more appropriate way to refer to this status. However, we reference it here to accurately reflect the literature.

In the literature on measures of welfare dependency, there has been substantial attention to the measurement of duration of welfare receipt, or how long individuals receive assistance. It should be noted, however, that much of this literature predates PRWORA and its imposition of time limits of TANF receipt. The emphasis on duration of welfare receipt as a measure of dependency connects to the ideology that when individuals spend more time receiving welfare, they are less able to leave welfare. However, critics have questioned the idea that receiving benefits has negative effects on motivation. More broadly, they have challenged the emphasis on economic outcomes and individuals' behavior in traditional definitions of self-sufficiency and dependency, proposing more expansive and possibly more meaningful measures for these concepts.

Critiques of Traditional Measures of Self-Sufficiency and Dependency

Several scholars have criticized the traditional definitions of self-sufficiency and dependency, especially their limited focus on economic outcomes. These scholars argue that the traditional definitions overemphasize the role of people's choices and personal traits and fail to account for factors at the interpersonal, community, and institutional levels that shape people's lives.¹³ Similarly, other scholars dispute the emphasis on the duration of welfare receipt, arguing that it disregards the complex factors contributing to poverty.¹⁴ Outside of literature on welfare, literature on well-being reiterates the idea that multiple factors at both individual and structural levels shape people's quality of life. The well-being literature identifies various dimensions of well-being, including mental, physical, social, economic, and environmental.¹⁵

In rethinking the traditional, one-dimensional definitions of self-sufficiency, experts have offered more expansive perspectives on self-sufficiency that look beyond economic outcomes. Several scholars emphasize a psychological component of self-sufficiency among individuals with low incomes, identifying factors such as hope, autonomy, and self-determination.¹⁶ One definition of self-sufficiency, developed through a focus group with low-income job seekers, emphasizes that it is a goal-oriented process involving the continual growth of psychological skills rather than a static state.¹⁷ In addition to psychological capacity, education and work-related skills are also key aspects of self-sufficiency at the individual level. Through interviews with women receiving welfare, Scott et al. found that women identified their lack of education and skills as a barrier to self-sufficiency, given that many jobs that require minimal educational attainment are low-paying and do not provide the benefits that are necessary to get ahead.¹⁸

¹³ Bratt, R. G., & Keyes, L. C. (1998). Challenges confronting nonprofit housing organizations' self-sufficiency programs. *Housing Policy Debate*, 9(4), 795–824., Daugherty, R. H., & Barber, G. M. (2001). Self-Sufficiency, Ecology of Work, and Welfare Reform. *Social Service Review*, 75(4), 662–675., Robbins, S. M., & Barcus, H. R. (2004). Welfare Reform and Economic and Housing Capacity for Low-Income Households, 1997–1999. *Policy Studies Journal*, 32(3), 439–460.

¹⁴ Contini, D., & Negri, N. (2007). Would Declining Exit Rates from Welfare Provide Evidence of Welfare Dependence in Homogeneous Environments? *European Sociological Review*, 23(1), 21–33.

¹⁵ Linton, M.-J., Dieppe, P., & Medina-Lara, A. (2016). Review of 99 self-report measures for assessing well-being in adults: exploring dimensions of well-being and developments over time. *BMJ Open*, 6(7), e010641., Schulte, P. A., Guerin, R. J., Schill, A. L., Bhattacharya, A., Cunningham, T. R., Pandalai, S. P., Eggerth, D., & Stephenson, C. M. (2015). Considerations for Incorporating “Well-Being” in Public Policy for Workers and Workplaces. *American Journal of Public Health*, 105(8), e31–e44., Trudel-Fitzgerald, C., Millstein, R. A., von Hippel, C., Howe, C. J., Tomasso, L. P., Wagner, G. R., & VanderWeele, T. J. (2019). Psychological well-being as part of the public health debate? Insight into dimensions, interventions, and policy. *BMC Public Health*, 19(1), 1712., Voukelatou, V., Gabrielli, L., Miliou, I., Cresci, S., Sharma, R., Tesconi, M., & Pappalardo, L. (2021). Measuring objective and subjective well-being: dimensions and data sources. *International Journal of Data Science and Analytics*, 11(4), 279–309. <https://doi.org/10.1007/s41060-020-00224-2>

¹⁶ Gowdy, E. A., & Pearlmuter, S. (1993). Economic Self-Sufficiency: It's Not Just Money. *Affilia*, 8(4), 368–387., Hong, P. Y. P., Polanin, J. R., & Pigott, T. D. (2012). Validation of the Employment Hope Scale: Measuring Psychological Self-Sufficiency Among Low-Income Jobseekers. *Research on Social Work Practice*, 22(3), 323–332., Hong, P. Y. P., Sheriff, V. A., & Naeger, S. R. (2009). A Bottom-up Definition of Self-sufficiency: Voices from Low-income Jobseekers. *Qualitative Social Work*, 8(3), 357–376.

¹⁷ Hong, P. Y. P., Sheriff, V. A., & Naeger, S. R. (2009). A Bottom-up Definition of Self-sufficiency: Voices from Low-income Jobseekers. *Qualitative Social Work*, 8(3), 357–376, pp. 357–358.

¹⁸ Scott, E. K., London, A. S., & Gross, G. (2007). “I Try Not to Depend on Anyone but Me”: Welfare-Reliant Women's Perspectives on Self-Sufficiency, Work, and Marriage. *Sociological Inquiry*, 77(4), 601–625.

The literature on poverty more broadly has also shifted away from using purely economic-based measures toward incorporating individual capacity into measurement. Economist and philosopher Amartya Sen has argued for a measure of poverty that takes into account people’s capability to achieve well-being, as they define it, with the resources available to them.¹⁹ Likewise, Haveman proposes “earnings capacity” as a poverty measure, which he defines as the degree to which a family “[has] the skills and capabilities to earn its way out of poverty were it fully to use them.”²⁰ He argues that this is a superior measure of poverty compared to traditional measures, such as poverty thresholds and income levels, as it captures the dynamic characteristics of families rather than static snapshots of families in time. Garfinkel and Haveman further explore this measure of earnings capacity by using a human capital framework to assess earnings capacity in a way that accounts for wages, education, and experience.²¹ They found that, for the most part, people experience poverty because the skills and opportunities available to them limit their ability to leave poverty, rather than because they are not working enough to increase their income. These scholars support the idea that economic and educational factors, each of which contribute to an individual’s ability to earn income, are critical aspects of self-sufficiency in the welfare context.

Beyond individual-level capacity, scholars argue that structural factors also play an important role in self-sufficiency. Daugherty and Barber propose redefining self-sufficiency using the “ecology-of-work perspective” which recognizes the importance of interpersonal and community factors, such as childcare, parenting, domestic violence, the availability of affordable housing, transportation, neighborhood safety, and racial and gender discrimination. Moreover, they identify workplace factors that influence self-sufficiency, such as wages, work environment, and employee benefits.²²

Similarly, Robbins and Barcus redefine self-sufficiency as “increased ability to meaningfully participate in the social and economic system . . . [it] includes the ability to obtain quality housing, health care, childcare, and food security in addition to sufficient income.”²³ They propose two new measures of self-sufficiency: (1) employment and housing capacity; and (2) quality of life, which includes food scarcity and medical care. Likewise, Gowdy and Pearlmuter identified the importance of interpersonal and community influences on self-sufficiency through a factor analysis of data gathered from women with low incomes.²⁴ They identified family and individual well-being (i.e., the ability to meet basic needs like adequate food quantity and quality) and basic assets for living in the community (i.e., the ability to afford reliable transportation and decent housing) as key factors underlying economic self-sufficiency.

In addition to critiquing the limited focus of traditional definitions of self-sufficiency on economic outcomes, critics argue that viewing self-sufficiency in a binary way (i.e., either an individual is self-sufficient, or they are not) does not reflect the reality of poverty or the role of public assistance in people’s lives. These critics note that all Americans receive some form of government assistance regardless of income, such as tax deductions for mortgages and interest payments and Social Security and Medicare benefits to older adults. Moreover, they note that leaving welfare is not the

¹⁹ Albelda, R. (1999). Women and poverty: Beyond earnings and welfare. *The Quarterly Review of Economics and Finance*, 39(5), 723–742.
Sen, A. (1993). Capability and Well-Being. In M. Nussbaum & A. Sen (Eds.), *The Quality of Life*. Clarendon Press.

²⁰ Haveman, R. (1993). Changing the poverty measure: pitfalls and potential gains. *Brookings Rev*, 11, 24–27.

²¹ Garfinkel, I., & Haveman, R. H. (2013). *Earnings Capacity, Poverty, and Inequality*. Elsevier.

²² Daugherty, R. H., & Barber, G. M. (2001). Self-Sufficiency, Ecology of Work, and Welfare Reform. *Social Service Review*, 75(4), 662–675.

²³ Robbins, S. M., & Barcus, H. R. (2004). Welfare Reform and Economic and Housing Capacity for Low-Income Households, 1997–1999. *Policy Studies Journal*, 32(3), 439–460, p. 445.

²⁴ Gowdy, E. A., & Pearlmuter, S. (1993). Economic Self-Sufficiency: It’s Not Just Money. *Affilia*, 8(4), 368–387.

same as leaving poverty: people who exit the welfare system may still need significant support despite no longer accessing safety net programs.²⁵

In an effort to reflect people’s experiences related to poverty and public assistance more accurately, many new conceptualizations of self-sufficiency take a strengths-based approach and strive to reflect the lived experiences of individuals with low incomes. In redefining self-sufficiency, several studies incorporated the voices of individuals with low incomes through interviews and focus groups.²⁶ This literature emphasizes a long, nonlinear process to achieving self-sufficiency that involves incremental movement through stages of personal and economic growth.²⁷ For instance, one study that conducted interviews with women who rely on state assistance found that they viewed their use of public benefits as a “necessary part of a long-term strategy to bettering themselves . . . not a way of life but a step on the road to self-sufficiency through self-improvement.”²⁸

In more recent years, several measures of self-sufficiency have emerged which strive to assess self-sufficiency more holistically by looking at multiple dimensions of people’s lives. For example, the Arizona Self-Sufficiency Matrix measures individuals’ level of independence and quality of life across 18 different domains (such as food, housing, mental health, family relations, and children’s education) using a 5-point scale ranging from “In Crisis” to “Empowered” to assess each domain.²⁹ A product of the Arizona Homeless Evaluation Project, the measure has been used or adapted by numerous organizations across the country.³⁰ Similarly, the anti-poverty non-profit Economic Mobility Pathways (EMPath) created the Bridge to Self-Sufficiency, a tool to measure self-sufficiency and help families achieve economic mobility.³¹ The tool includes five pillars (family stability; well-being; financial management; education and training; and employment and career management), across which individuals can receive a maximum score of 100 points.³² This literature review builds on the new, more expansive conceptualizations and measures of self-sufficiency and dependency to identify a framework and measurement approach for these concepts that more accurately reflect the complexity of poverty and the experiences of individuals with low incomes.

Methods

The project team took a multi-step approach to completing the outcomes review, outlined in this section.

²⁵ Leibson Hawkins, R. (2005). From Self-Sufficiency to Personal and Family Sustainability: A New Paradigm for Social Policy. *Journal of Sociology and Social Welfare*, 32, 77., Robbins, S. M., & Barcus, H. R. (2004). Welfare Reform and Economic and Housing Capacity for Low-Income Households, 1997–1999. *Policy Studies Journal*, 32(3), 439–460.

²⁶ Gowdy, E. A., & Pearlmuter, S. (1993). Economic Self-Sufficiency: It’s Not Just Money. *Affilia*, 8(4), 368–387., Hong, P. Y. P., Sheriff, V. A., & Naeger, S. R. (2009). A Bottom-up Definition of Self-sufficiency: Voices from Low-income Jobseekers. *Qualitative Social Work*, 8(3), 357–376., Nelson, M. (2002). The Challenge of Self-sufficiency: Women on Welfare Redefining Independence. *Journal of Contemporary Ethnography*, 31, 582–614., Scott, E. K., London, A. S., & Gross, G. (2007). “I Try Not to Depend on Anyone but Me”: Welfare-Reliant Women’s Perspectives on Self-Sufficiency, Work, and Marriage. *Sociological Inquiry*, 77(4), 601–625.

²⁷ Daugherty, R. H., & Barber, G. M. (2001). Self-Sufficiency, Ecology of Work, and Welfare Reform. *Social Service Review*, 75(4), 662–675., Hong, P. Y. P., Polanin, J. R., & Pigott, T. D. (2012). Validation of the Employment Hope Scale: Measuring Psychological Self-Sufficiency Among Low-Income Jobseekers. *Research on Social Work Practice*, 22(3), 323–332., Hong, P. Y. P., Sheriff, V. A., & Naeger, S. R. (2009). A Bottom-up Definition of Self-sufficiency: Voices from Low-income Jobseekers. *Qualitative Social Work*, 8(3), 357–376., Leibson Hawkins, R. (2005). From Self-Sufficiency to Personal and Family Sustainability: A New Paradigm for Social Policy. *Journal of Sociology and Social Welfare*, 32, 77.

²⁸ Nelson, M. (2002). The Challenge of Self-sufficiency: Women on Welfare Redefining Independence. *Journal of Contemporary Ethnography*, 31, 582–614.

²⁹ Culhane, D., Parker, W., Poppe, B., Gross, K., & Sykes, E. (2008). Toward Understanding Homelessness: The 2007 National Symposium on Homelessness Research. Accountability, Cost-Effectiveness, and Program Performance: Progress Since 1998. Case Study: The Arizona Evaluation Project on Homelessness. U.S. Department of Health and Human Services, Office of Human Services Policy, Office of the Assistant Secretary for Planning and Evaluation. <https://aspe.hhs.gov/reports/toward-understanding-homelessness-2007-national-symposium-homelessness-research-accountability-cost-0>

³⁰ Schoenfeld, E. (2017). LifeWorks Self-Sufficiency Matrix User Manual. LifeWorks.

³¹ The Bridge to Self-Sufficiency. (n.d.). EMPATH Economic Mobility Pathways. Retrieved March 16, 2023, from <https://empathways.org/approach/bridge-to-self-sufficiency>

³² Some of the TOA pilots have identified the Arizona Self-Sufficiency Matrix and the Bridge to Self-Sufficiency for use in their programs.

Development of the Conceptual Framework

The team began by developing a conceptual framework to serve as a roadmap for the review, specifically related to capacity for self-sufficiency. Recognizing that unidimensional conceptualizations of capacity and self-sufficiency are inadequate, the team sought to identify the range of constructs that contribute to capacity for self-sufficiency. Finding that the term “capacity” is not widely used in the literature on welfare, poverty, and self-sufficiency, the team compiled articles and reports that defined related concepts like “self-sufficiency” and “well-being” (See Appendix A) and broke these constructs down into component parts. The team mapped the components of these constructs, as reported in the literature, and consolidated them into a single conceptual framework, presented in this report. The framework shows how elements of capacity and related influences exist across levels, residing in individuals, families, and communities and at a structural level. Within the individual level, the framework articulates five dimensions of capacity for self-sufficiency: (1) skills, knowledge, and experience; (2) economic well-being; (3) mental well-being; (4) physical well-being; and (5) social capital. More detail on the conceptual framework is provided below.

Key Informant Interviews

The team invited key informants to provide input on the conceptual framework and suggest approaches to searching and prioritizing the literature. From December 2022 to February 2023, the team conducted four one-hour, loosely-structured interviews with five experts, listed in Table 3.

Table 3. Key Informants

Name, Title, Affiliation
James Booth , Patricia and Rodes Hart Professor of Psychology and Human Development, Vanderbilt University
David Grusky , Edward Ames Edmond Professor, School of Humanities and Sciences; Director, Center on Poverty and Inequality; Senior Fellow, Institute for Economic Policy Research, Stanford University
Carolyn Heinrich , Patricia and Rodes Hart Professor of Public Policy, Vanderbilt University
Cynthia Osborne , Professor of Early Childhood Education and Policy; Executive Director, Prenatal-to-3 Policy Impact Center, Vanderbilt University
Timothy Smeeding , Lee Rainwater Distinguished Professor of Public Affairs and Economics; Former Director, Institute for Research on Poverty, University of Wisconsin-Madison

Administrative Data

The team aimed to recommend measures derived from administrative data that TDHS can access through existing data sources. Reviewing studies of public benefit usage and evaluation of programs designed promote self-sufficiency the team sought to identify outcome measures captured through administrative data that most closely linked to domains in the individual level of the conceptual framework. The team then reviewed the codebook for the P20 Connect TN (P20) database, Tennessee’s integrated, longitudinal database, to determine which of these measures were already being collected.

Identifying Surveys, Scales, and Indices

To identify measurement instruments that could be incorporated into a participant survey, the team began by identifying combinations of search terms in each of the five individual-level domains and

the family well-being domain of the conceptual framework (See Appendix B). Team members entered these terms into two databases of research literature, Google Scholar and ProQuest. They prescreened the first 50 results of each search, skipping any results that had already come up in other searches. Sources describing measurement of relevant indicators were screened into the review if they were published in or after 2000 and, if they described measurement instruments themselves, if they included psychometric information about the reliability and validity of the instruments. When a source met these criteria, team members scanned it to identify surveys, scales, and indices that were used to measure outcomes. If studies identified measurement tools had not already been identified in the search and prescreening steps, team members would search for sources describing their reliability and validity to be screened in to the review. Appendix D includes a graphical depiction of the process the team undertook to complete the literature review.

Once sources were screened into the review, team members systematically reviewed them, entering as much information as possible into the tables presented in Appendix E. The team met to discuss which instruments in each domain appeared strongest and most suited to the aims of the TOA pilots and evaluations. These instruments are highlighted in the tables in Appendix E and described in the Recommendations section of this report.

Findings

This section begins with a presentation of the conceptual framework, then shares insights from the key informant interviews. Next, there is an overview of findings regarding use of administrative data and information about the literature used to identify surveys, scales, and indices. The final section presents recommendations for specific measures TDHS might consider for the TOA pilot evaluations and other efforts to track outcomes for program participants.

Conceptual Framework

In searching for existing frameworks illustrating capacity for self-sufficiency, the team identified 60 relevant sources (See Appendix A). Each of these sources described theorized components of constructs like individual and family use of safety net programs, self-sufficiency, and well-being. Scanning each of these sources, team members mapped these components, noting significant overlap in authors' conceptualizations. The team categorized components by level, as well as by domain within the individual level. Next, the team generated a conceptual framework for capacity for self-sufficiency, shown in Figure 1.

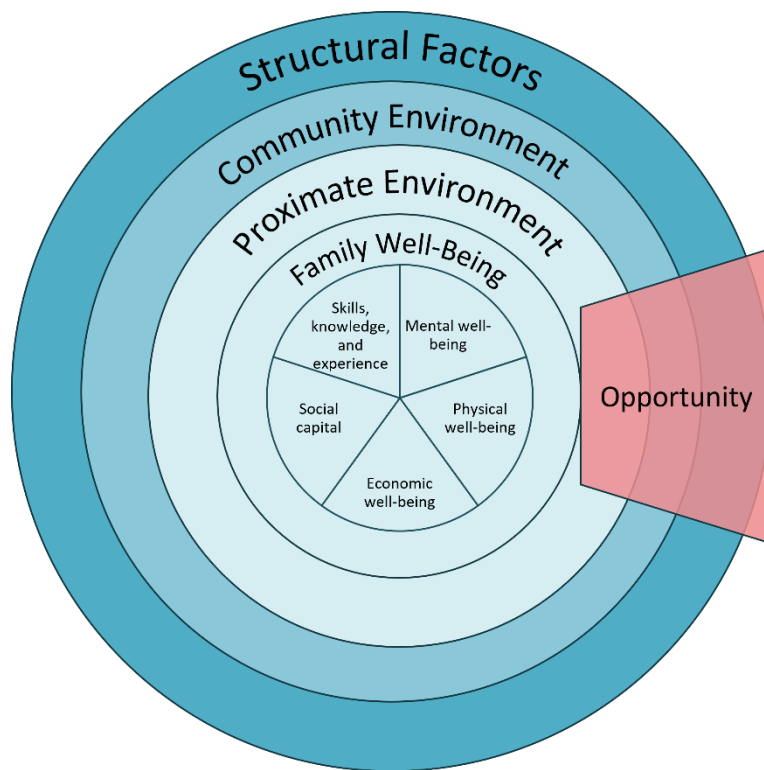


Figure 1. Conceptual Framework of Capacity for Self-Sufficiency

This conceptual framework conveys that multiple factors at multiple levels influence one’s capacity for self-sufficiency. Beginning with the innermost ring and working outward: the inner ring shows **individual states and traits**. The individual exists within the family unit, which is characterized by relationships among family members, shown in the **family well-being** ring. For the purposes of the TOA evaluations, child outcomes, which influence parental functioning, also fall into the family well-being ring of the conceptual framework. Individuals and families are most affected by their **proximate environment**, which includes their home, schools, jobs, and service providers, as well as availability and access to basic necessities like clean water, healthy food, and safe and appropriate housing. The broader community, shown in the **community environment** ring, also exerts influence, defined by characteristics like neighborhood safety, housing stock, child care availability, and transportation infrastructure, among others. On the outer ring of the framework is the superstructure in which individuals, families, and communities exist: **structural factors** such as systemic racism, classism, and sexism. Table 4 lists examples of constructs in each ring of the conceptual framework.

Table 4. Example Components of Each of the Conceptual Framework’s Levels or Domains

Level or Domain	Examples of Component Constructs ^a
Skills, Knowledge, and Experience	Money management knowledge, parenting skills, work experience, education level
Social Capital	Membership in and access to social networks that expand one’s access to resources, sense of community
Economic Well-Being	Income, debt, benefits receipt, ability to make ends meet, employment, employment benefits
Physical Well-Being	Health quality, healthcare access, substance use

Mental Well-Being	Internal subjective state (mental health, depression, stress levels, sense of purpose, sense of belonging), psychological capability, executive functioning (impulse control, working memory, mental flexibility), experience of trauma
Family Well-Being	Family relationships (e.g., domestic violence, parent-child relationships), child outcomes (behavior, school readiness, educational attainment)
Proximate Environment	Characteristics of jobs, schools, programs, etc. in which people and families participate; access to clean water and healthy food; safe and appropriate housing
Community Environment	Community resources (including housing stock, child care availability, transportation infrastructure, etc.), neighborhood safety
Structural Factors	Systemic racism, classism, and sexism

^a Examples listed are not exhaustive

Importantly, though not explicitly mapped in Figure 1, there is an interactive relationship between and among each element in the conceptual framework. Within the individual, characteristics in one domain may influence characteristics in another. For example, there is a well-established link between physical and mental health. Somatic complaints are common among those living with depression. Someone who is faring poorly in the mental health domain may also demonstrate poor physical health. Furthermore, a person with poor mental and physical health may struggle to maintain a job, threatening their economic well-being. Likewise, what is happening at one level of this framework influences what happens in another. For example, individual stress affects parenting quality. A person whose mental, physical, and economic well-being are poor may not be able to adequately respond to emerging behavioral problems in their child, which in turn affects short- and long-term outcomes for the child. The complex, multidimensional nature of these influences on capacity for self-sufficiency points to the need for measurement of outcomes in multiple domains to understand the varied impacts of interventions aimed at increasing capacity for self-sufficiency in multiple domains.

The conceptual framework also highlights the critical role of opportunity in increasing capacity for self-sufficiency. Capacity to pursue education, get a job, or move up the income ladder is of little importance if there is no opportunity to do so. Furthermore, there is variability in opportunity, causing the return on education and employment to vary depending on one's context. Opportunity can be understood broadly as the conditions or circumstances that make it possible to achieve a goal. Within the context of the TOA pilots, opportunity can refer to accessible education and training options or available, good jobs. In the conceptual framework, opportunity exists at, influences, and is influenced by the proximate environment, the community environment, and structural factors. Importantly, the conceptual framework illustrates that people will be better equipped to take advantage of the opportunities in their environment if they have achieved stability and sufficiency in the individual-level domains.

Insights from Key Informant Interviews

Most of the team's conversations with key informants centered on the conceptual framework as a tool to organize our review. Key informants agreed that the framework suitably reflected factors affecting the capacity of individuals and families to achieve self-sufficiency. Other themes from their feedback included the following:

- It is important to consider how the different domains of well-being impact development in others. For example, growth in social capital can lead to future increases in financial well-being.
- Visualized as the outermost ring in the conceptual framework, structural factors like systemic racism are embedded and experienced within the institutions making up the proximate environment, such as schools.
- While components of the framework that describe contextual factors outside of the family (i.e., proximate environment, community environment, and structural factors) are difficult to measure precisely, it is worthwhile to attempt to capture them as accurately as possible. While TOA pilots will not directly influence contextual factors, they are still important to track, especially as they may limit or facilitate participants' success.

Measures Derived from Administrative Data

Administrative data always reflects historical understanding of both important outcomes and relevant predictors. In the case of administrative data collected by state and federal agencies, widespread, longstanding views of public assistance programs and their recipients have skewed what is captured in these systems heavily toward individual and household-level measures of economic well-being: benefit receipt, employment, and earnings. As educational attainment and physical disability have direct influence for labor market outcomes, data describing these are also typically present in state and federal administrative data systems. Given the relatively narrow range of information captured in administrative records, data from these systems alone are insufficient for evaluating capacity for self-sufficiency as conceptualized in this report. That said, while such data may not provide the complete picture, what is already collected is quite useful, especially when augmented with survey data.

The P20 is a valuable resource that TDHS contributes to but does not yet fully utilize. Many of the measures included in the P20 have been used effectively both as predictors and outcome measures in previous research, including in experimental impact evaluations. Linkages between individuals and outcomes present in data contributed by various agencies can be particularly important. For example, student assessment data collected by the Tennessee Department of Education could be connected to data on parents participating in the TOA pilots to facilitate research into the impacts of those programs on child educational outcomes, a focus of “two generation” interventions.³³ Other states like Connecticut, Iowa, Minnesota, and New Jersey have successfully utilized analogous data in their own integrated databases to evaluate the effectiveness of their post-secondary and workforce policies and to support their partner agencies.³⁴ As scholars and practitioners reassess their models in efforts such as this one, the breadth and depth of data collected and tracked in databases such as the P20 Connect will continue to expand, increasing their value to participating agencies and their partners.

Surveys, Scales, and Indices

To identify instruments for subjective measurement of capacity for self-sufficiency (that is, respondents' own perception of their capacity for self-sufficiency), the team's systematic database

³³ It should be noted that, for the TOA pilot evaluations, consent for use of children's data is not required for an adult to participate in a pilot program. As a result, there may be little child data available for the pilot evaluations. However, there may be other applications for which TDHS could consider using this linked data.

³⁴ U.S. Department of Education, Institute of Educational Sciences, National Center for Education Statistics. (March, 2022). State approaches to engaging and sustaining postsecondary and workforce partners. Washington, DC: Author.

search yielded 203 sources in total; 163 of these were screened in. From these articles and reports, the team identified 68 unique measurement instruments across the five individual-level domains of the conceptual framework and the family well-being level.

Within each domain, the team assessed several facets of each instrument to arrive at recommendations for measurement. These facets included the following:

- **Reliability and Validity:** While authors reported a range of statistics related to reliability and validity, the team sought to identify instruments that demonstrated a high level of reliability as well as content and construct validity. For example, when authors reported a Cronbach's alpha for an instrument, the team considered a value around or above .70 to be acceptable, as is standard practice.
- **Population:** Some instruments identified in the search were designed for use with specific populations (e.g., disabled adults) or were normed with very particular groups that could not reasonably be compared with the adult population of Tennessee. The team prioritized instruments that had been tested with diverse populations that were not dissimilar to participants in TOA pilots in any obvious or extreme way.
- **Length:** Recognizing that any instrument will likely be incorporated into a longer survey, the team aimed to identify brief scales and indices.
- **Cost:** Some instruments are copyrighted and available to users at a cost. The team gave greater consideration to tools that are freely available, where possible. For the most part, this did not limit the review, as only a handful of the instruments that were identified in the search were proprietary.

Appendix E details each unique instrument included in this review, organized by domain and level. Further description of the tools that the team recommends for use in the TOA evaluations or for other TDHS applications is included below in the Recommendations section.

Recommendations

Based on the findings from the literature review and insights from key informants, the team makes the following recommendations to TDHS.

Recommendation 1: Use administrative data sources to track characteristics of benefits receipt, educational attainment, and employment.

TDHS collects rich administrative data describing the use of safety net programs among Tennesseans. This data feeds into the P20 database, which includes linked administrative data from seven state agencies. Using P20 data, TDHS can track a range of outcomes closely related to two domains of the conceptual framework describing capacity for self-sufficiency: individuals' economic well-being and their skills, knowledge, and experience. Information on employment, earnings, education, and training are available. The data can also be used to construct measures of use of safety net programs. Depending on each pilot's intervention and logic model, different outcome measures may be most appropriate.

Researchers studying poverty, public benefit use, and economic well-being regularly use several of the measures included in the P20 database to track gains in income, income stability, and reductions in receipt and dollar value of public benefits. For example, the Department of Labor and Workforce Development has contributed data on quarterly employment and earnings in Unemployment Insurance (UI)-covered jobs. Although state UI records do not cover all employment (out-of-state

employment, contract work, informal employment and federal jobs are not included) coverage is sufficient to allow measures derived from such data to be used as primary outcomes in a variety of analyses and impact evaluations. These include, for example, “ever employed in UI-covered job,” and “average quarterly earnings,” which are typically used in job training and placement interventions, including those targeted at current or former TANF or SNAP recipients.

Similarly, TDHS already collects data on Families First (TANF) and SNAP receipt (monthly, yes/no), benefit value (\$), current spell length (# of months receiving the benefit in current cycle), lifetime benefit value (\$) and benefit receipt over lifetime (# of months having ever received the benefit). Similar measures have been used in research for decades, including, for example, during the evaluation of the National Welfare-to-Work Strategies in the run-up to the Personal Responsibility and Work Opportunity Act of 1996. Since these measures are so widely used in the literature, the data can be used to provide apples-to-apples comparisons of programmatic impact and to identify promising intervention approaches (and avoid ineffective ones).

TDHS could also use P20 data to construct individual measures describing use of safety net programs that are adapted from the aggregate indicators included in the U.S. Department of Health and Human Services’ (HHS) annual report to Congress (Table 1). Table 5 describes how these indicators could be adapted to track individual benefits receipt. Although HHS’s definitions account for SSI receipt in addition to TANF and SNAP, the P20 does not include SSI information. Therefore, SSI is not reflected in the individual adaptations of the HHS definitions presented below. Depending on the availability of data, these definitions could be expanded to include additional benefits.

Table 5. Adaptation of HHS Welfare Indicators to Individual Measurement

HHS Indicator	Individual-Level Definition
Degree of Dependence	Proportion of income a family receives from TANF and SNAP. Alternatively, this could be a binary indicator (y/n) of whether more than half of a family’s income comes from SNAP and TANF.
Program Receipt and Family Labor Force Attachment	Binary indicator (y/n) of whether the family receives TANF or SNAP and includes a family member participating in the labor force, which includes people working, or unemployed and looking for work.
Program Reciprocity	Binary indicator (y/n) of whether an individual received or lived with a family member who received cash benefits during the year from TANF or benefits from SNAP.
Program Participation Among Those Eligible	Binary indicator (y/n) of whether, within a given period of time, an individual was eligible for and received either TANF or SNAP.
Multiple Program Receipt	Binary indicator (y/n) of whether, within a given period of time, an individual received both TANF and SNAP benefits.
Program Receipt Duration	Number of months within a given period that an individual or family received either TANF or SNAP benefits. Duration of TANF receipt and SNAP receipt could be measured independently or pooled together to capture any benefit receipt.

The P20 is not the only administrative database DHS should consider drawing indicators from. There are others that the department can access to support its work and the work of its partners. For example, the Administrative Data Research Facility (ADRF) is an integrated database developed by the Coleridge Initiative and available to staff from state and federal agencies that complete their

Applied Data Analytics training; Tennessee joined the ADRF in 2021.³⁵ Since multiple states contribute their own UI and higher education data to the ADRF, including at least two of Tennessee's neighbors (Kentucky and Missouri), the database can be used to track employment, earnings, and educational outcomes for individuals who commute out of state for work or school. As more states sign on to the initiative, the database will become more useful for this purpose. Similarly, the National Student Clearinghouse is an educational non-profit that aggregates data on enrollment, degree completion, student loans and more for most post-secondary institutions in the United States (including those in Tennessee). This data can and has been used to evaluate programs designed to improve the educational and labor force outcomes of adults with low incomes, such as Project QUEST.³⁶

Recommendation 2: Capture additional information about family finances and job quality through surveys.

While administrative data can paint a rough picture of a family's financial situation or an individual's employment circumstances, it is incomplete. To develop a more detailed picture, it is necessary to augment administrative data with objective information that program participants report. This includes factual information like whether a person has a bank account or the numbers of hours they work at their primary job in a typical week. Many large-sample surveys have questions to understand characteristics of individual and household finances, as well as features of respondents' jobs. These questions can be used or adapted in a survey of pilot participants.

Family Finances

A household's financial situation has four main components: income, expenditures, assets, and debts.³⁷ These aspects each contain different information about the household's financial resources and needs and provide necessary context for interpreting the other components; for example, information about income alone cannot answer whether a household is able to meet its needs.

One way to gather information about these components is to be comprehensive: to ask, for every potential type of each component, whether respondents have and/or receive it and the amount they have, receive, and/or owe. This approach is frequently taken by large surveys intended to track demographic trends, such as the Survey of Income and Program Participation or the Panel Study of Income Dynamics. Another approach is to ask fewer questions that are either more generalized or specifically targeted. When deciding how detailed and comprehensive questions should be, it is important to consider how each question is relevant to the goals of the research and what information it adds. Below we discuss the types of each component that may be included on a comprehensive survey as well as how researchers have commonly chosen to ask about it. There are several large-sample surveys that include batteries of questions about family finances in the dimensions listed below. As these are quite lengthy, we have not reproduced them in this document.

Income

There are several types of income: wage and salary income; income from self-employment; investment earning; unemployment insurance; worker's compensation; child support; alimony; pensions; Social Security retirement; Social Security Insurance (SSI)/Social Security Disability

³⁵ Kuehn, D. (2022). Better data for better policy: The Coleridge Initiative in Tennessee: Leveraging the Applied Data Analytics training for expansion and policy impact. Washington, D.C.: Urban Institute.

³⁶ Order, A. & Elliott, M. (2021). Eleven year gains: Project QUEST's investment continues to pay dividends. New York, NY: Economic Mobility Corporation, Inc.

³⁷ Ratcliffe, C., Burke, J., Gardner, J., & Knoll, M. (2020). Evidence-Based Strategies to Build Emergency Savings. Consumer Financial Protection Bureau Office of Research.

Insurance (SSDI); TANF; SNAP; and gifts. Researchers evaluating programs are usually primarily interested in a program's effect on whether and how much income participants make in two categories: earnings (money earned from employment or self-employment) and benefits. Administrative data is an adequate source for information about use of safety net programs, but the information it captures related to individual and family income is lacking. While it is useful to use administrative data to track quarterly earnings in UI-covered jobs, this measure does not account for many types of income, as noted in the previous section. Asking participants to report earnings from all sources provides a more complete snapshot of their household income.

In addition to asking about the types of inflows participants have and the amount of each type, researchers frequently ask questions to try to understand how stable respondents' income is. For example, a questionnaire might ask respondents if their income stays about the same from month to month, varies a little, or varies a lot. Another useful question is whether respondents feel they usually have a good sense of what their future income will be.

Expenditures

Types of expenditures include: outlays for transportation, rent, utilities, medical costs, and debt payments, among others. Unless they want to understand the cost of specific essentials, researchers are more often interested whether respondents can afford those essentials than in the specific amount spent on each category. This can be achieved by asking about respondents' cash flow: whether they usually have enough money left over at the end of the month, have just enough to make ends meet, or do not have enough to make ends meet. Another approach, which yields more discrete data points, is asking whether participants have had trouble affording each type of essential expenditure in the past year.

Assets

Assets include checking accounts, IRA accounts, CDs, savings accounts, bonds, stocks, investment accounts, life insurance, real estate, and businesses. Assets can be divided into two categories based on whether they are liquid (that is, easily accessible) or illiquid (not easily accessible). Researchers commonly ask about whether respondents have any savings, and, if so, the total amount, as well as the total of specifically liquid assets. Alternatively, researchers may ask about the existence and amount of non-retirement savings to assess available assets.

An important aspect of financial well-being is *how* one saves: whether respondents have bank accounts or if they use high-cost financial services like check cashing services. Researchers may ask whether respondents have a bank account or inquire more specifically about how often and in what ways they use it. They may also ask whether and how frequently respondents use services like check-cashers.

Debt

Debt can be incurred through credit cards, revolving store accounts, student loans, auto loans, personal loans, medical debt, and payday loans. Researchers will frequently ask simply about total debt or about the types of debt most likely to be relevant to the population or intervention they are studying. They may also distinguish between debt from lower-interest, lower-cost forms of credit like bank loans and credit cards, and high-interest, high-cost forms of credit like payday loans. Researchers may also ask about the frequency of use of different types of credit.

Job Quality

The quality of a job is inherently subjective. However, most conceptualizations of job quality include dimensions that can be objectively measured. For example, a recent literature review conducted by

the Urban Institute identified the following dimensions of job quality: wages and earnings; hours and scheduling; benefits and leave; working conditions; job design (e.g., autonomy, task composition); nonmonetary value (e.g., meaningfulness); and forward prospects (e.g., training opportunities).³⁸ Likewise, an analysis of General Social Survey data identified five job quality dimensions: individual task discretion, monetary compensation, job security, low work intensity, and work conditions.³⁹ Some of these dimensions are more salient to the TOA project than others. Specifically, these include dimensions of job quality that relate to compensation, benefits, and flexibility—features of a job that are most closely linked to self-sufficiency. Table 6 lists sample questions from large-sample surveys that have been used to track indicators of job quality, as recommended in the Urban Institute’s report. These items, or others from the surveys listed in this table, could be adapted or used as written in a survey of pilot participants.

Table 6. Job Quality Survey Questions by Dimension

Job Quality Dimension	Survey Source	Sample Questions
Wages and earning	Current Population Survey Outgoing Rotation Group	<ul style="list-style-type: none"> • Are you paid by the hour on this job? • How much do you earn per hour? • How much do you usually earn per week at this job before deductions?
Hours	Current Population Survey Outgoing Rotation Group	<ul style="list-style-type: none"> • How many hours did you work last week at all jobs? • Do you usually work 35 hours or more a week at this job?
Scheduling and predictability	National Longitudinal Survey of Youth 1997	<ul style="list-style-type: none"> • How far in advance do you usually know what days and hours you will need to work? <i>Response options include: 3 days or less, 4 to 7 days, between 1 and 2 weeks, 3 weeks or more, always works same schedule.</i>
Schedule flexibility	National Longitudinal Survey of Youth 1997	<ul style="list-style-type: none"> • Which of the following best describes how your working hours are decided? By working hours we mean the time you start and finish work, and not the total hours you work per week or month. <ul style="list-style-type: none"> ○ Starting and finishing times are decided by my employer and I cannot change them on my own. ○ Starting and finishing times are decided by my employer but with my input. ○ I can decide the time I start and finish work, within certain limits. ○ I am entirely free to decide when I start and finish work. ○ When I start and finish work depends on things outside of my control and outside of my employer’s control.
Benefits	Panel Study of Income Dynamics (1994 Wave)	<ul style="list-style-type: none"> • Does your employer pay for life insurance that would cover you if your death were not job related?
Leave	American Time Use Survey 2018 Leave Survey	<ul style="list-style-type: none"> • Can you take paid leave for: <ul style="list-style-type: none"> ○ Your own illness or medical care? ○ The illness or medical care of another family member? ○ Childcare, other than for illness? ○ Eldercare? ○ Vacation? ○ Errands or personal reasons? ○ The birth or adoption of a child?

³⁸ Katz, B., Congdon, W., & Shakesprere, J. (2022). Measuring Job Quality: Current Measures, Gaps, and New Approaches. Washington, D.C.: Urban Institute.

³⁹ Horowitz, J. (2016). Dimensions of Job Quality, Mechanisms, and Subjective Well-Being in the United States. *Sociological Forum*, 31(2), 419–440. <https://doi.org/10.1111/socf.12251>

		<ul style="list-style-type: none"> • Would your employer approve of you taking unpaid leave for the following reasons: (same list as above)
Job security	General Social Survey	<ul style="list-style-type: none"> • Respondents are asked to indicate whether the statement is very true, somewhat true, not too true, or not at all true with respect to the work they do: The job security is good. • Thinking about the next 12 months, how likely do you think it is that you will lose your job or be laid off? Response options include: very likely, fairly likely, not too likely, or not at all likely. • Were you laid off your main job at any time in the last year?

Recommendation 3: Use well-tested scales with demonstrated reliability and validity among similar populations to capture participants’ subjective assessments of well-being in multiple individual and family domains.

Participants’ subjective assessments of their own well-being add to a fuller picture of how they are faring. Many components of the domains of the conceptual framework, such as parental stress, cannot be easily directly observed. There are, however, scales and indices that researchers and practitioners have developed to measure them. This document includes recommendations for scales and indices that are best suited to capturing constructs in domains of the conceptual framework; taking into consideration reliability and validity; populations with which instruments were tested, length; and cost. In selecting instruments for a survey, TDHS and evaluators should consider which domains of the conceptual framework are most important to tap based on pilots’ targets and logic models. Surveys for pilots’ impact evaluations should be tailored for each program’s design, incorporating measures from those domains that are most likely to be affected by the intervention. Recommended instruments, by domain, follow in Table 7.

Table 7. Recommended Survey Instruments

Economic Well-Being
Consumer Financial Protection Bureau (CFPB) Financial Well-Being Scale
<p>Description: The CFPB Financial Well-Being Scale assesses financial well-being, defined as “a state wherein a person can fully meet current and ongoing financial obligations, can feel secure in their financial future, and is able to make choices that allow them to enjoy life.” It includes 10 statements⁴⁰ with Likert-type response scales, such as “I am concerned that the money I have or will save won’t last” and “I have money left over at the end of the month.” Respondents score between 0 and 100, with 0 representing the lowest level of financial security and 100 representing the highest level of financial security.⁴¹</p> <p>Testing, Reliability, and Validity: The scale underwent systematic development and extensive testing outlined in a thorough Technical Report. Data from nearly 15,000 respondents ages 18 and up participated in scale development and testing, representing diversity in gender, race, income, and education. The instrument’s developers applied Item Response Theory⁴² for item selection and calculated marginal reliability scores. For both the 10-item and a 5-item scale, marginal reliability is above .80, indicating that the instrument is highly reliable (that is, it consistently taps the same construct in the same way, such that variations in respondents’ scores are due to variation in subjective financial</p>

⁴⁰ A five-item version is also available.

⁴¹ Consumer Financial Protection Bureau (2015). Measuring financial well-being: A guide to using the CFPB financial well-being scale. Washington, DC: Author.

⁴² Item Response Theory includes a family of statistical models that can be applied to optimize the set of items included in a scale measuring latent traits to maximize reliability and validity.

well-being, not measurement error). Items correlated in the expected direction with related constructs and indicators, demonstrating content and construct validity.⁴³

Length:	10 items, 4 minutes
Cost:	Free
Publication Year:	2015

Additional Information Supporting the Recommendation: This scale is best used as a complement to administrative indicators of financial well-being, including income and benefits receipt. It is subjective, but it provides a fuller picture of how a family is faring financially than what can be gleaned from administrative data. Neither this instrument nor administrative data in the P20 get at aspects of financial well-being like amounts of savings or debt, which are difficult or laborious to capture accurately. However, individuals responding to the CFPB Financial Well-Being Scale will be taking that information into account when furnishing responses to the scale’s statements.

USDA Household Food Security Module

Description: Widely used in government-sponsored and independent research studies, the USDA Household Food Security Module measures the extent to which families have access to “nutritionally adequate and safe foods” and “an assured ability to acquire acceptable foods in socially acceptable ways (e.g., without resorting to emergency food supplies, scavenging, stealing, and other coping strategies).” The USDA points out that food insecurity, the inverse of food security, is a result of financial limitations. The scale includes 18 items in total, beginning with three screening questions. Of the remaining items, respondents are shown those that are relevant, per the screening items.⁴⁴ Scale scoring ranges from 0 to 18 for households with children present, with 0 representing the highest food security and 18 representing the lowest food security.⁴⁵

Testing, Reliability, and Validity: The food security module was developed for the Current Population Survey (CPS) of the Census in 1994 with extensive testing. With minor revisions, it has remained relatively unchanged. It has remained in the CPS and appeared in multiple federally-sponsored surveys, including the Panel Survey of Income Dynamics, the Survey of Income and Program Participation, and the Early Childhood Longitudinal Study. It remains a consensus measure for food security, with literature supporting its adaptation for use with many subgroups of people.⁴⁶

Length:	18 items for households with children; 2-4 minutes
Cost:	Free
Publication Year:	2012 (most recent update)

Additional Information Supporting the Recommendation: The team’s inclusion of the USDA Household Food Security Module reflects its assessment of the importance of capturing material hardship among pilot participants. While participants’ assessments of their financial situation, along with administrative indicators of income and benefits receipt, provide important information about the flow of resources into a household, these measures do not offer insight into the day-to-day experiences of families. Inclusion of the Food Security Module expands the understanding of financial well-being to incorporate the ways families are able to use the resources they have to make ends meet.

Mental Well-Being

Short Form Health Survey (SF-36)

Description: The SF-36 is a widely used tool for assessing self-reported health and functioning in physical, mental, and social domains. Specifically, it includes subscales for the following constructs: “1) limitations in physical activities

⁴³ Ibid.

⁴⁴ For example, if a respondent answers, “Often true” or “Sometimes true,” to the statement, “We worried about whether our food would run out before we got money to buy more,” the respondent would be taken to the next section of the instrument. If the respondent answers, “Never true” to this statement and the other two screening items, their survey is complete.

⁴⁵ Bickel, G., Nord, M., Price, C., Hamilton, W., & Cook, J. (2000). Guide to measuring household food security, revised 2000. Washington, DC: U.S. Department of Agriculture, Food and Nutrition Service.

⁴⁶ Wunderlich, G. S., & Norwood, J. K. (2006). History of the development of food insecurity and hunger measures. In *Food insecurity and hunger in the United States: An assessment of the measure*, (Eds., Wunderlich, G. S. & Norwood, J.L.), 23-40. Washington, DC: National Academies Press.

because of health problems; 2) limitations in social activities because of physical or emotional problems; 3) limitations in usual role activities because of health problems; 4) bodily pain; 5) general mental health (psychological distress and well-being); 6) limitations in usual role activities because of emotional problems; 7) vitality (energy and fatigue); and 8) general health perceptions.”⁴⁷ Scores range from 0 to 100 for each subscale, with higher scores reflecting more positive health. The subscales can also be aggregated into two summary measures: physical health and mental health.

Testing, Reliability, and Validity: The SF-36 has been extensively tested in general and specific populations, demonstrating excellent reliability and validity. Consistently, studies have found Cronbach’s alphas for each of the eight subscales exceeding .80, with summary scores for physical and mental health above .90. Likewise, because the SF-36 is so widely used in so many contexts, researchers have thoroughly examined its validity, demonstrating construct, content, and criterion validity.⁴⁸

Length:	36 items, ⁴⁹ 5 minutes
Cost:	Free
Publication Year:	1992

Additional Information Supporting the Recommendation: The SF-36 is not a diagnostic tool, nor does it capture specific symptoms of common mental health challenges, like depression and anxiety. However, in the context of the TOA evaluation, the most salient information is not whether one meets diagnostic criteria for given conditions but rather what is the effect of their symptoms on functioning. The SF-36 accurately and reliably captures functioning as it is influenced by physical and mental health. Also, in addition to being widely used and widely accepted, the SF-36 captures both physical and mental health, covering two of the domains at the individual level in the conceptual framework. Using a single instrument to tap two domains would reduce respondent burden.

Physical Well-Being

Short Form Health Survey (SF-36)

Description: See above. The SF-36 covers both mental and physical health domains.

Social Capital

Social Provisions Scale (SPS)

Description: The Social Provisions Scale taps respondents’ sense of perceived support. It includes 24 statements and a 4-point agree/disagree response scale. Example statements include, “There are people I can depend on to help me if I really need it,” and “There is no one I feel comfortable talking about problems with.” The items map to six subscales: reliable alliance, attachment, nurturance, social integration, reassurance of worth, and guidance. In addition to subscale scores, it is also possible to calculate a total score. In all cases, higher scores are indicative of greater perceived social support.⁵⁰

Testing, Reliability, and Validity: The SPS was validated with nearly 1,800 respondents across three studies, most of whom were college students and public school teachers. It has since been used with general population samples and samples drawn from specific groups, including people with low incomes.⁵¹ Testing indicates the instrument has high reliability, with a combined Cronbach’s alpha of .92. Subscale alphas are lower, and the author of the SPS cautions against their use, as some subscales appear to be correlated. The authors have demonstrated construct validity in a college student sample.⁵²

⁴⁷ Ware Jr, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical care*, 473-483.

⁴⁸ Ware Jr, J. E. (2000). SF-36 health survey update. *Spine*, 25(24), 3130-3139.

⁴⁹ A 12-item version of the survey is available for a cost. It also has high reliability and validity.

⁵⁰ Gottlieb, B. H., & Bergen, A. E. (2010). Social support concepts and measures. *Journal of Psychosomatic Research*, 69(5), 511–520.

⁵¹ Orpana, H. M., Lang, J. J., & Yurkowski, K. (2019). Original quantitative research-validation of a brief version of the Social Provisions Scale using Canadian national survey data. *Health Promotion and Chronic Disease Prevention in Canada: Research, Policy and Practice*, 39(12), 323.

⁵² Gottlieb, B. H., & Bergen, A. E. (2010). Social support concepts and measures. *Journal of Psychosomatic Research*, 69(5), 511–520.

Length:	24 items, ⁵³ 5 minutes
Cost:	Free
Publication Year:	1990

Additional Information Supporting the Recommendation: N/A

Bridging Social Capital Questionnaire, v2 (BSC)

Description: The BSC captures the extent to which individuals’ social networks include people who are demographically different from them. This element of social capital provides individuals with conduits to resources outside of their homogenous social networks. The BSC includes seven subscales: (1) socialization in the job place; (2) membership in community activities; (3) participation in community activities; (4) contact with similar/different people; (5) assistance; (6) trust of institutions, corporations, and other people; and (7) trust of immediate people. Subscales can be administered independently, increasing the flexibility of the instrument. Subscales six and seven can also be combined into a single “Trust” subscale.⁵⁴

Testing, Reliability, and Validity: The instrument’s developers conducted focus groups and consultations with experts to confirm content validity. The original version of the BSC was developed and tested with 138 Latino immigrants living in the U.S. All subscales⁵⁵ achieved a Cronbach’s alpha equal to or greater than .80. Confirmatory factor analysis demonstrated discriminant validity for the subscales, indicating each subscale is tapping a discrete component of bridging social capital.⁵⁶ The BSC’s developers refined the instrument using a Rasch model to assess the instrument’s performance in a new sample of Latino immigrants living in the U.S. and identified several items to delete from the original BSC, yielding version 2 of the instrument.⁵⁷

Length:	61 items, approximately 25 minutes for the full questionnaire, though subscales may be used independently (subscales range from 3 to 17 items)
Cost:	Free
Publication Year:	2021

Additional Information Supporting the Recommendation: Scholars identify two types of social capital: bonding social capital and bridging social capital. The SPS taps bonding social capital while the BSC taps bridging social capital. Given the length of the BSC, the team recommends using select subscales that are most relevant to the TOA pilots (for instance, subscales four through seven). The team recommends the BSC with some cautions. First, it should be noted that the instrument was tested with a particular population: Latino immigrants living in the U.S. It has not been validated in the general population or low-income populations. Second, the BSC has not been widely used to date. However, according to our review, this instrument appears to be the most rigorously tested measure of bridging social capital, which may be essential to understanding the impact of efforts to build program participants’ social capital. It has a strong conceptual foundation, backed by trustworthy qualitative methods. If this instrument is used in the TOA evaluations, the team recommends examining the psychometric performance of the instrument and/or its subscales with participants before drawing conclusions. This could be as simple as calculating a Cronbach’s alpha for the subscales to ensure that the internal consistency of the scales tracks into the population served by the pilots.

Brief Sense of Community Scale (BSCS)

Description: The BSCS measures the degree to which individuals feel belonging in and valued by a community, and the extent to which they feel their and others’ needs will be met. The instrument provides a total sense of community score and scores for four subscales: (1) needs fulfillment, (2) group membership, (3) influence, and (4) emotional connection. Respondents are presented with eight statements and a five-point Likert response scale (“Strongly Agree”

⁵³ Also available are 10-item and 5-item versions of the SPS, each demonstrating satisfactory reliability and validity.

⁵⁴ Villalonga-Olives, E., Adams, I., & Kawachi, I. (2016). The development of a bridging social capital questionnaire for use in population health research. *SJM-population Health*, 2, 613-622.

⁵⁵ The developers combined the trust subscales for their analyses.

⁵⁶ Villalonga-Olives, E., Adams, I., & Kawachi, I. (2016). The development of a bridging social capital questionnaire for use in population health research. *SJM-population Health*, 2, 613-622.

⁵⁷ Villalonga-Olives, E., Kawachi, I., & Rodriguez, A. M. (2021). Rasch model of the bridging social capital questionnaire. *SJM-Population Health*, 14, 100791.

to “Strongly Disagree”). Example statements include: “I can get what I need in this neighborhood,” and “I have a say about what goes on in my neighborhood.”⁵⁸

Testing, Reliability, and Validity: The instrument was validated with a random sample of nearly 300 people in the Midwest participating in a community health promotion initiative. The sample is diverse in terms of gender, age, Hispanic origin, income, and educational attainment. Cronbach’s alphas for the measure include: .92 for the full scale, .86 for needs fulfillment, .94 for group membership, .77 for influence, and .87 for emotional connection. Supporting the instrument’s construct validity, scores on the total scale and its subscales correlate in the expected directions with related measures.⁵⁹ The instrument has also been tested with many other populations, including non-Hispanic Black and Hispanic college students, youth of color in urban areas, and LGBT people in the U.S south, among others.

Length:	8 items, 3 minutes
Cost:	Free
Publication Year:	2008

Additional Information Supporting the Recommendation: The inclusion of the BSCS among the recommended measures reflects the importance of neighborhood context in individual outcomes. While it is a subjective measure, it may, in some models, be even more predictive of outcomes than indicators of neighborhood quality derived from administrative data. Put differently, whether someone thinks their neighborhood increases their capacity may matter more than the amount or type of resources present. Furthermore, the sense of belonging among people living in the same neighborhood may vary depending on their race, cultural practices, income levels, and more.

Family Well-Being

Parenting Stress Index – 4 Short Form (PSI-4 SF)

Description: The PSI-4 SF captures stress in the parent-child relationship that may contribute to problematic behaviors, including child maltreatment. The instrument contains three subscales: (1) parental distress, (2) parent-child dysfunctional interaction, and (3) difficult child. These relate to three domains of stress: child characteristics, parent characteristics, and situational life stress. Parents of children 0-12 years complete the survey, reporting on themselves and their child(ren). Total scores and subscale scores are computed and translated into percentiles relative to the general population.⁶⁰

Testing, Reliability, and Validity: Studies of the PSI-SF’s reliability have yielded an average Cronbach’s alpha of .85, demonstrating high internal consistency. In a normative sample of 800 parents, subscale alphas ranged from .80 to .87.⁶¹ Similar results were found in a sample of 100 parents with children enrolled in Head Start.⁶² Items for the PSI-SF were pulled verbatim from the long-form PSI, and the factor structure of the instrument was confirmed (that is, the items still grouped into the three subscales), reflecting content validity and demonstrating construct validity. Additionally, PSI-SF scores have been found to correlate in the expected direction with related constructs.

Length:	36 items, 10 minutes
Cost:	\$223 for a PSI-4 Short Form Kit containing the Professional Manual and 25 Short Form Record/Profile Forms
Publication Year:	Initially developed in 1983, the PSI is now in its fourth edition

Additional Information Supporting the Recommendation: Acknowledging the need for any participant survey to minimize respondent burden, the team is recommending a single measure of family well-being, despite the complexity of this area. The PSI captures elements of both child behavior and parenting quality, providing a high-level snapshot of the family dynamic. TOA interventions targeting either or both parent and child well-being could be expected to affect PSI-4 SF scores. There are a few reservations to the team’s recommendation of this instrument. First, the

⁵⁸ Peterson, N. A., Speer, P. W., & McMillan, D. W. (2008). Validation of a brief sense of community scale: Confirmation of the principal theory of sense of community. *Journal of community psychology*, 36(1), 61-73.

⁵⁹ Ibid.

⁶⁰ PAR, Inc. (n.d.). Parenting Stress Index, Fourth Edition Short Form.

⁶¹ Abidin, R. R. (1995). Parenting Stress Index, Third Edition: Professional Manual. Psychological Assessment Resources, Inc.

⁶² Reitman, D., Currier, R. O., & Stickle, T. R. (2002). A critical evaluation of the parenting stress index-short form (PSI-SF) in a Head Start population. *Journal of Clinical Child & Adolescent Psychology*, 31(3), 384-392.

instrument is proprietary and has costs associated with its use. Copyright restrictions limited the team from fully exploring the PSI-4 SF. Second, the instrument is designed for parents with children ages 0-12, though it has been used in studies of parents of adolescents. The PSI-4 SF's developers have a similar tool for use with parents of adolescents, the Stress Index for Parents of Adolescents (SIPA) that could be considered.

Cross-Domain

Everyday Discrimination Scale (EDS)

Description: The EDS measures the degree to which respondents perceive being discriminated against for a range of reasons, including but not limited to race, gender, religion, sexual orientation, and education or income level. Respondents are presented with the question, “In your day-to-day life, how often do any of the following things happen to you?” followed by nine statements, such as, “You are treated with less respect than other people are,” and “You are threatened or harassed.” Response options range from “Never” to “Almost every day.” If respondents answer, “More than a few times a year” or more to any statement, they see the follow-up question, “What do you think is the main reason for these experiences?” Response options list characteristics, such as those listed earlier, that could be seen as the reason for perceived discrimination.

Testing, Reliability, and Validity: In a study exploring the validity and reliability of a measure of experiences of discrimination, researchers found that the EDS achieved a Cronbach’s alpha score of .88 in both African American and Latino respondents drawn from a sample of working-class adults.⁶³ A smaller validation study involving African American men calculated an alpha of .80 and statistically significant correlations in the expected direction with theoretically linked concepts, demonstrating construct validity.⁶⁴

Length:	10 items, with one potential follow-up question to each, ⁶⁵ 5 minutes
Cost:	Free
Publication Year:	1997

Additional Information Supporting the Recommendation: The conceptual framework depicts the role of structural influences on individual and family well-being. While it is not possible to measure the extent of influences like racism, sexism, and classism, among others, the EDS allows researchers to capture the individual experience of them. The perceptions of participants about the degree to which their race, gender, and class affect their ability to avail themselves of opportunities describe important context for the TOA pilots.

Recommendation 4: Examine variability in opportunity in the communities where participants live and work and, if necessary, include indicators of opportunity as covariates in outcome or impact analyses.

Scholars theorize well-being as a multidimensional construct, made up of multiple factors. As such, the conceptual framework shows how individuals and families are inextricably situated in communities and a broader social structure. They interact both directly and indirectly with formal institutions and policies, as well as informal systems, that exert their influence on all aspects of daily life. Of particular relevance to TDHS and the TOA pilots, the availability of opportunities to support educational advancement, skill-building, and employment (which in turn contribute to skills, knowledge, and experience as well as economic well-being) are dependent on myriad factors, including the policy climate, public funding, community values, and much more.

In evaluating the TOA pilots, it is essential to consider how individual and program contexts that affect the availability of opportunities contribute to outcomes and impacts. Although the pilots are

⁶³ Krieger N., Smith K., Naishadham D., Hartman C., Barbeau E.M. “Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health.” *Social Science & Medicine*. 2005; 61(7):1576-1596.

⁶⁴ Taylor T.R., Kamarck T.W., Shiffman S. “Validation of the Detroit area study discrimination scale in a community sample of older African American adults: the Pittsburgh healthy heart project.” *International Journal of Behavioral Medicine*. 2004; 11:88–94.

⁶⁵ A five-item version is also available (Cronbach’s alpha = .77).

not targeting community or structural outcomes, neighborhood variation in factors like the transportation infrastructure, income inequality, and the availability of high-quality child care may exert far more influence on individual and family outcomes than program services. The same is true for participants' experiences of structural and systemic racism. While rigorous randomization procedures assure that program participants in both intervention and comparison groups are embedded in comparable contexts, TDHS and its evaluators should confirm equivalence on a predetermined set of contextual indicators linked to opportunity. If equivalence on community-level measures of opportunity is not achieved, these indicators may be included in statistical analyses as covariates to adjust for variability.

As part of the TOA evaluation, MEF's subcontractor, the Urban Institute, is preparing a data dashboard to track macro-level indicators in Tennessee counties, grand divisions, and across the state. This dashboard includes several indicators that can be used at multiple levels to characterize opportunity. Since these data are already being compiled and analyzed by the project team, it makes sense to use them in the outcome and impact analyses as described above.

Recommendation 5: Develop logic models to explore the relationship between program services and participant outcomes that account for the complex, dynamic nature of capacity for self-sufficiency.

At a high level, the goal of the TOA pilots is to move participants toward self-sufficiency, characterized by decreased use of safety net programs. However, the comprehensive array of services that pilots offer target a range of individual and family outcomes not limited to economic well-being. Further, as described several times in this report, how a person is faring in one domain affects their functioning in others. An individual's success in a program targeting economic well-being is influenced by their mental and physical health, which the program may not address. The complex nature of capacity for self-sufficiency and the comprehensive approach to support services that pilot programs are taking to support it necessitate clear explication of a logic model that articulates inputs (services), outputs (dosages), intermediate outcomes, and long-term outcomes.

The indicators listed in this report are reliable, valid, and appropriate for use in program evaluation. Depending on a program's logic model, indicators listed in this report could be used as outcomes or covariates in statistical analyses, or as descriptive contextual information provided alongside findings.

Conclusion and Next Steps

As of the writing of this report, MEF has worked with the seven pilot grantees to develop evaluation plans and is in the process of mapping outcomes aligned with each program's intervention and evaluation designs. Having identified each pilot's targeted outcomes, MEF will use the recommendations from this report to clarify which measures to track using administrative data and which questions, scales, and indices to include in follow-up surveys. MEF researchers will also use the recommendations from this report to inform their multidimensional approach to assessing poverty (as a way of measuring use of safety net programs) among pilot participants before, during, and after participation in the pilots.

Appendix A: References Informing the Conceptual Framework

The following sources were gathered and used to inform the development of the conceptual frame. Many of these sources were also identified during the systematic search of literature for surveys, scales, and indices. References for sources that were screened in to the systematic search, as well as sources that informed the rest of this report, are included either in the tables in Appendix E or in footnotes throughout the report text.

- Albelda, R. (1999). Women and poverty: Beyond earnings and welfare. *The Quarterly Review of Economics and Finance*, 39(5), 723–742.
- Alfred, M. V. (2007). Welfare Reform and Black Women’s Economic Development. *Adult Education Quarterly*, 57(4), 293–311.
- Andrews, F. M., & McKennell, A. C. (1980). Measures of self-reported well-being: Their affective, cognitive, and other components. *Social Indicators Research*, 8(2), 127–155.
- Babcock, E. D. (2014). Using Brain Science to Design New Pathways out of Poverty (p. 37). Crittenton Women’s Union.
- Bratt, R. G., & Keyes, L. C. (1998). Challenges confronting nonprofit housing organizations’ self-sufficiency programs. *Housing Policy Debate*, 9(4), 795–824.
- Brüggen, E. C., Hogreve, J., Holmlund, M., Kabadayi, S., & Löfgren, M. (2017). Financial well-being: A conceptualization and research agenda. *Journal of Business Research*, 79, 228–237.
- Bush, I. R., & Kraft, M. K. (2001). Self-Sufficiency and Sobriety: Substance-Abusing Women and Welfare Reform. *Journal of Social Work Practice in the Addictions*, 1(1), 41–64.
- Cherlin, A., & Reeder, L. G. (1975). The Dimensions of Psychological Well-Being: A Critical Review. *Sociological Methods & Research*, 4(2), 189–214.
- Crouse, G. (2022). Welfare Indicators and Risk Factors, 21st Report to Congress. U.S. Department of Health and Human Services, Office of Human Services Policy, Office of the Assistant Secretary for Planning and Evaluation.
- Daugherty, R. H., & Barber, G. M. (2001). Self-Sufficiency, Ecology of Work, and Welfare Reform. *Social Service Review*, 75(4), 662–675.
- Dugan, J., Booshehri, L. G., Phojanakong, P., Patel, F., Brown, E., Bloom, S., & Chilton, M. (2020). Effects of a trauma-informed curriculum on depression, self-efficacy, economic security, and substance use among TANF participants: Evidence from the Building Health and Wealth Network Phase II. *Social Science & Medicine* (1982), 258, 113136.
- Finn, M., Gilmore, B., Sheaf, G., & Vallières, F. (2021). What do we mean by individual capacity strengthening for primary health care in low- and middle-income countries? A systematic scoping review to improve conceptual clarity. *Human Resources for Health*, 19(1), 5.
- Gardiner, J. (2000). Rethinking self-sufficiency: Employment, families and welfare. *Cambridge Journal of Economics*, 24(6), 671–689.
- Gowdy, E. A., & Pearlmutter, S. (1993). Economic Self-Sufficiency: It’s Not Just Money. *Affilia*, 8(4), 368–387.

- Headey, B., Holmström, E., & Wearing, A. (1984). Well-Being and Ill-Being: Different Dimensions? *Social Indicators Research*, 14(2), 115–139.
- Hong, P. Y. P. (2013). Toward A Client-Centered Benchmark for Self-Sufficiency: Evaluating the ‘Process’ of Becoming Job Ready. *Journal of Community Practice*, 21(4), 356–378.
- Hong, P. Y. P., Polanin, J. R., & Pigott, T. D. (2012). Validation of the Employment Hope Scale: Measuring Psychological Self-Sufficiency Among Low-Income Jobseekers. *Research on Social Work Practice*, 22(3), 323–332. <https://doi.org/10.1177/1049731511435952>
- Hong, P. Y. P., Sheriff, V. A., & Naeger, S. R. (2009). A Bottom-up Definition of Self-sufficiency: Voices from Low-income Jobseekers. *Qualitative Social Work*, 8(3), 357–376.
- Jayakody, R., Danziger, S., & Pollack, H. (2000). Welfare Reform, Substance Use, and Mental Health. *Journal of Health Politics, Policy and Law*, 25(4), 623–652.
- Joseph, R. (2018). The welfare/self-sufficiency gap among single mothers through theoretical lenses. *Journal of Human Behavior in the Social Environment*, 28(6), 731–745.
- Joseph, R. (2019). Poverty, Welfare, and Self-Sufficiency: Implications for the Social Work Profession. *Journal of Poverty*, 23(6), 505–520.
- Julnes, G., Fan, X., & Hayashi, K. (2001). Understanding self-sufficiency of welfare leavers in Illinois: Elaborating models with psychosocial factors. *New Directions for Evaluation*, 2001(91), 33–44.
- Kabeer, N. (1996). Agency, Well-being & Inequality: Reflections on the Gender Dimensions of Poverty. *IDS Bulletin*, 27(1), 11–21.
- Kapteyn, A., Lee, J., Tassot, C., Vonkova, H., & Zamarro, G. (2015). Dimensions of Subjective Well-Being. *Social Indicators Research*, 123(3), 625–660.
- King, M. F., Renó, V. F., & Novo, E. M. L. M. (2014). The Concept, Dimensions and Methods of Assessment of Human Well-Being within a Socioecological Context: A Literature Review. *Social Indicators Research*, 116(3), 681–698.
- Kneipp, S. M. (2000). Economic Self-Sufficiency: An Insufficient Indicator of How Women Fare after Welfare Reform. *Policy, Politics, & Nursing Practice*, 1(4), 256–266.
- Krzczkowska, A., Spalding, D. M., McGeown, W. J., Gow, A. J., Carlson, M. C., & Nicholls, L. A. B. (2021). A systematic review of the impacts of intergenerational engagement on older adults’ cognitive, social, and health outcomes. *Ageing Research Reviews*, 71, 101400.
- Leibson Hawkins, R. (2005). From Self-Sufficiency to Personal and Family Sustainability: A New Paradigm for Social Policy. *Journal of Sociology and Social Welfare*, 32, 77.
- Lie, G.-Y., & Moroney, R. M. (1992). A Controlled Evaluation of Comprehensive Social Services Provided to Teenage Mothers Receiving AFDC. *Research on Social Work Practice*, 2(4), 429–447.
- Lindeblad-Fry, M. (2019). An Analysis of Government Dependency and Tribal Economic Performance. *Berkeley Public Policy Journal*.
- Linton, M.-J., Dieppe, P., & Medina-Lara, A. (2016). Review of 99 self-report measures for assessing well-being in adults: Exploring dimensions of well-being and developments over time. *BMJ Open*, 6(7), e010641.

- London, R. A. (2006). The Role of Postsecondary Education in Welfare Recipients' Paths to Self-Sufficiency. *The Journal of Higher Education*, 77(3), 472–496.
- Luthans, F., Avey, J. B., Avolio, B. J., & Peterson, S. J. (2010). The development and resulting performance impact of positive psychological capital. *Human Resource Development Quarterly*, 21(1), 41–67.
- Mayne, J. (2016). The Capabilities, Opportunities and Motivation Behaviour-Based Theory of Change Model. (Working paper).
- McDowell, I. (2010). Measures of self-perceived well-being. *Journal of Psychosomatic Research*, 69(1), 69–79.
- Metsch, L. R., & Pollack, H. A. (2005). Welfare Reform and Substance Abuse. *The Milbank Quarterly*, 83(1), 65–99.
- Nelson, M. (2002). The Challenge of Self-sufficiency: Women on Welfare Redefining Independence. *Journal of Contemporary Ethnography*, 31, 582–614.
- Netemeyer, R. G., Warmath, D., Fernandes, D., & Lynch, J. G., Jr. (2018). How Am I Doing? Perceived Financial Well-Being, Its Potential Antecedents, and Its Relation to Overall Well-Being. *Journal of Consumer Research*, 45(1), 68–89.
- New Findings on Children, Families, and Economic Self-Sufficiency: Summary of a Research Briefing. (1995). National Academies Press.
- Nieboer, A., Lindenberg, S., Boomsma, A., & Bruggen, A. C. Van. (2005). Dimensions Of Well-Being And Their Measurement: The Spf-II Scale. *Social Indicators Research*, 73(3), 313–353.
- Panel on Measuring Subjective Well-Being in a Policy-Relevant Framework, Committee on National Statistics, Division on Behavioral and Social Sciences and Education, & National Research Council. (2013). Subjective Well-Being: Measuring Happiness, Suffering, and Other Dimensions of Experience (A. A. Stone & C. Mackie, Eds.). National Academies Press.
- Parker, L. (1994). The Role of Workplace Support in Facilitating Self-Sufficiency among Single Mothers on Welfare. *Family Relations*, 43(2), 168–173.
- Partners Ending Homelessness. (n.d.). Vulnerability Index—Service Prioritization Decision Assistance Tool (VI-SPDAT): Prescreen Triage Tool for Families.
- Pew Charitable Trust. (2018) Philadelphia's Poor: Experiences From Below the Poverty Line. <https://www.pewtrusts.org/en/research-and-analysis/reports/2018/09/26/philadelphias-poor-experiences-from-below-the-poverty-line>
- Raymond, C., & Cleary, J. (2013). A Tool and Process that Facilitate Community Capacity Building and Social Learning for Natural Resource Management. *Ecology and Society*, 18(1).
- Ruggeri, K., Garcia-Garzon, E., Maguire, Á., Matz, S., & Huppert, F. A. (2020). Well-being is more than happiness and life satisfaction: A multidimensional analysis of 21 countries. *Health and Quality of Life Outcomes*, 18(1), 192.
- Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57, 1069–1081.
- Ryff, C. D., & Keyes, C. L. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, 69(4), 719–727.

- Ryff, C. D., & Singer, B. (1996). Psychological well-being: Meaning, measurement, and implications for psychotherapy research. *Psychotherapy and Psychosomatics*, 65(1), 14–23.
- Schochet, O. N. (2022). Case Management or Child Care: Which Has the Greater Impact on Parental Human Capital and Self-Sufficiency in Two-Generation Programs? *Social Service Review*, 96(4), 703–743.
- Schulte, P. A., Guerin, R. J., Schill, A. L., Bhattacharya, A., Cunningham, T. R., Pandalai, S. P., Eggerth, D., & Stephenson, C. M. (2015). Considerations for Incorporating “Well-Being” in Public Policy for Workers and Workplaces. *American Journal of Public Health*, 105(8), e31–e44.
- Scott, E. K., London, A. S., & Gross, G. (2007). “I Try Not to Depend on Anyone but Me”: Welfare-Reliant Women’s Perspectives on Self-Sufficiency, Work, and Marriage*. *Sociological Inquiry*, 77(4), 601–625.
- Scott, E. K., London, A. S., & Myers, N. A. (2002). Dangerous Dependencies: The Intersection of Welfare Reform and Domestic Violence. *Gender & Society*, 16(6), 878–897.
- Smith, S., & Others, A. (1992). Pathways to Self-Sufficiency for Two Generations: Designing Welfare-to-Work Programs That Benefit Children Strengthen Families. *Foundation for Child Development*, 345 E.
- Tosun, J., Arco-Tirado, J. L., Caserta, M., Cemalcilar, Z., Freitag, M., Hörisch, F., Jensen, C., Kittel, B., Littvay, L., Lukeš, M., Maloney, W. A., Mühlböck, M., Rainsford, E., Rapp, C., Schuck, B., Shore, J., Steiber, N., Sümer, N., Tsakoglou, P., ... Vegetti, F. (2019). Perceived economic self-sufficiency: A country- and generation-comparative approach. *European Political Science*, 18(3), 510–531.
- Trudel-Fitzgerald, C., Millstein, R. A., von Hippel, C., Howe, C. J., Tomasso, L. P., Wagner, G. R., & VanderWeele, T. J. (2019). Psychological well-being as part of the public health debate? Insight into dimensions, interventions, and policy. *BMC Public Health*, 19(1), 1712.
- Vlaev, I., & Elliott, A. (2014). Financial Well-Being Components. *Social Indicators Research*, 118(3), 1103–1123.
- Voukelatou, V., Gabrielli, L., Miliou, I., Cresci, S., Sharma, R., Tesconi, M., & Pappalardo, L. (2021). Measuring objective and subjective well-being: Dimensions and data sources. *International Journal of Data Science and Analytics*, 11(4), 279–309.
- Xiao, J. J., Huang, J., Goyal, K., & Kumar, S. (2022). Financial capability: A systematic conceptual review, extension and synthesis. *International Journal of Bank Marketing*, 40(7), 1680–1717.
- Zaslow, M., Moore, K., Coiro, M. J., & Morrison, D. R. (1994). Programs To Enhance the Self-Sufficiency of Welfare Families: Working towards a Model of Effects on Young Children.

Appendix B: Glossary of Key Terms

Capacity	For the purposes of this review, the term “capacity” refers to the abilities, skills, traits, and contextual facilitators, including opportunities, that make up an individual’s or a family’s potential to achieve self-sufficiency.
Construct	A grouping of concepts into a single complex model. In this report, “construct” refers to human characteristics, behaviors, or perceptions that are not directly observable but may be assessed with measurement tools like surveys. For example, “well-being” is a construct.
Construct validity	The degree to which a measurement tool (such as a survey or scale) accurately describes what it is intended to describe. When a measurement tool has strong construct validity, users can make inferences based on it. For example, a doctor might administer a 9-question depression screen with demonstrated construct validity to patients. The doctor can be reasonably sure that patients’ responses are indicative of depressive symptoms.
Content validity	The degree to which a measurement tool captures all aspects or dimensions of the construct it is assessing.
Convergent validity	The degree to which a measurement tool operates similarly to other measures of the same construct. For example, a tool assessing children’s math skills should yield student scores that positively correlate with assessments of numeracy and quantitative reasoning.
Cronbach’s alpha	A measure of reliability, Cronbach’s alpha is a statistic that describes the internal consistency of a scale or subscale, or the degree to which items in a scale or subscale are tapping the same construct. High internal consistency is associated with strong measurement tools. Alpha ranges from 0 to 1, with higher values indicating greater internal consistency. Values between .70 and .80 are considered acceptable; above .80 is good; and above .90 is excellent.
Discriminant validity	The degree to which a measurement tool operates differently from measures to which it should not be similar. For example, scores from a tool assessing children’s math skills should not be correlated with score from tools assessing social skills or motor development.
Reliability	The degree to which a measurement tool operates consistently. For example, a tool with high reliability will produce similar scores for people with similar levels of the trait being measured.
Self-sufficiency	The ability to make ends meet for oneself and one’s family without relying on safety net programs like TANF and SNAP. For a thorough discussion of the definition of self-sufficiency and critiques of the concept as it has been defined in policy, see the report section, Review of Traditional Measures of Self-Sufficiency and Use of Safety Net Programs.
Well-being	A multi-dimensional construct describing a positive state of being that includes happiness, fulfillment, wellness, and health. While there is no consensus definition of well-being, most operationalizations include physical health, mental health, economic well-being, and positive social networks.

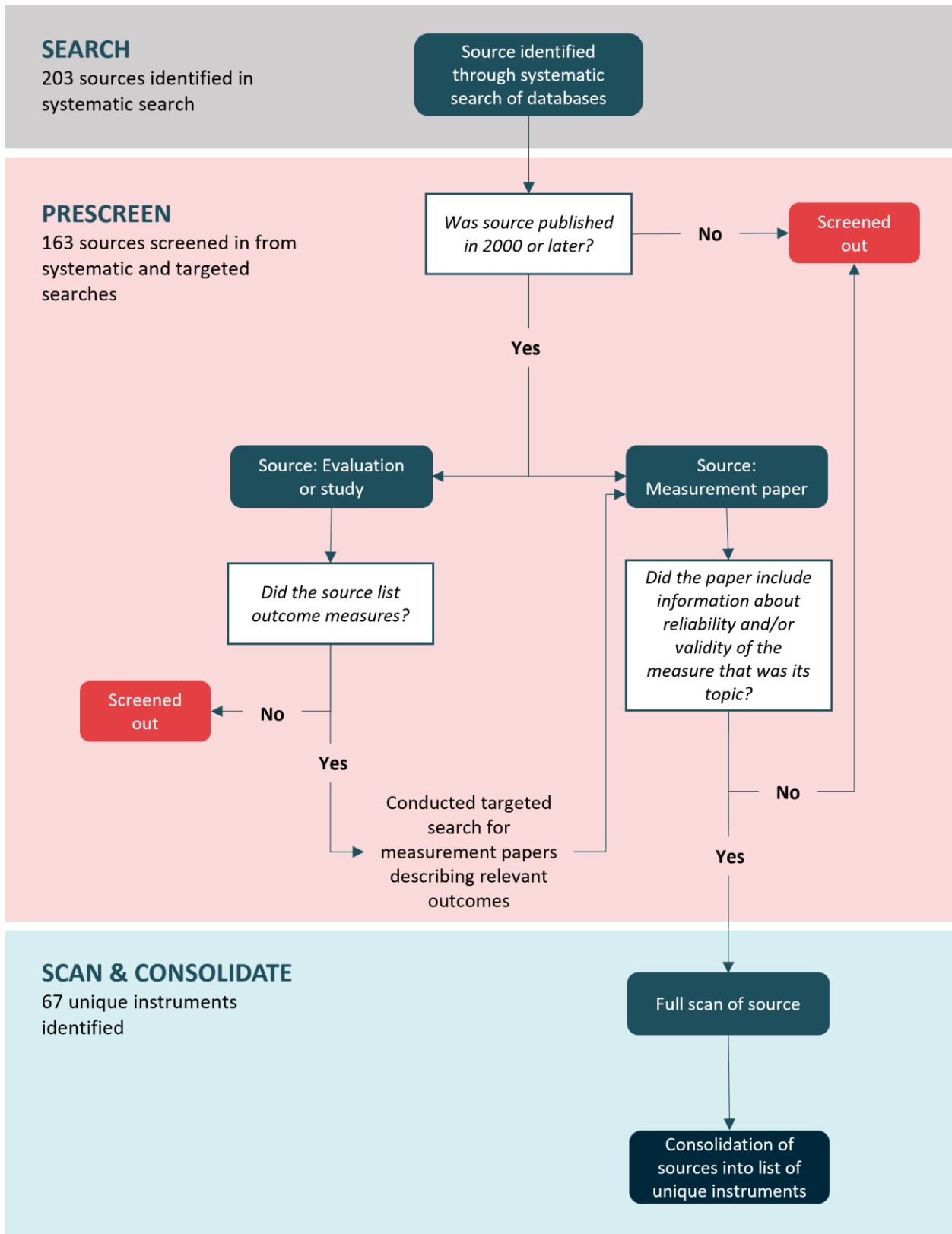
Appendix C: List of Search Terms for Surveys, Scales, and Indices

Table 8. Literature Review Search Terms

Domain	Search Terms
Economic Well-Being	economic wellbeing indicator economic wellbeing measurement financial resilienc* indicator financial reilienc* measurement
Mental Well-Being	mental health measurement psychological capability measurement executive functioning indicator executive functioning measurement executive skills indicator executive skills measurement
Social Capital	social capital indicator social capital measurement social support indicator social support measurement social network indicator social network measurement
Physical Well-Being	health indicator health measurement physical health indicator physical health measurement
Skills, Knowledge and Experience	human capital indicator human capital measurement education indicator employment indicator employment quality measurement parenting measurement
Family Well-Being	family wellbeing indicator family wellbeing measurement child wellbeing indicator child wellbeing measurement parent-child relationship indicator parent-child relationship measurement
Cross-Domain	quality of life indicator quality of life measurement

Appendix D: Literature Review Process Graphic

The following image pertains to the literature search for surveys, scales, and indices.



Appendix E: Tables of Surveys, Scales, and Indices

The following tables present the surveys, scales, and indices that the team identified during the literature search. Information is presented here as it was compiled by multiple team members for analysis. In places, information about psychometrics and the populations with which measures have been normed on or used with is taken verbatim from the sources listed in the Citation(s) column

Economic Well-Being Measures

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation(s)
Financial wellbeing	InCharge Financial Distress/Financial Well-Being Scale	InCharge Education Foundation	- Cronbach's alpha of 0.956 - Development included modified Delphi study of experts	Tested for validity/normed using data from general population (n = 1,097) and financially distressed credit counseling clients (n = 590)	8	Pravitz, A. D., Garman, E. T., Sorhaindo, B., O'Neill, B., Kim, J., & Drentea, P. InCharge Financial Distress/Financial Well-Being Scale. <i>European Journal of Psychological Assessment</i> . https://doi.org/10.1037/t60365-000 .
Material hardship	Survey of Income and Program Participation (SIPP) Adult Well-Being Topical Module	Census Bureau	Carle et al. 2009 developed model of selected SIPP questions using confirmatory factor analyses for ordered categorical measures: RMSEA = 0.05, CFI > 0.88, TLI > .95, gamma hat = .95.	SIPP given to large scale panel of American households	Carle et al. 2009's model covers different constructs related to material hardship: consumer durables (14 questions), resources to meet needs (12 questions), housing conditions (13 questions),	Carle, A. C., Bauman, K. J., & Short, K. (2009). Assessing the measurement and structure of material hardship in the United States. <i>Social Indicators Research</i> , 92, 35-51. https://doi.org/10.1007/s11205-008-9287-7 . Nelson, G. (2011). Measuring poverty: The official US measure and material hardship. <i>Poverty & Public Policy</i> , 3(3), 1-35. https://doi.org/10.2202/1944-2858.1077 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation(s)
					neighborhood problems and crime (17 questions), community services (4 questions)	
Food security	U.S. Household Food Security Survey Module	USDA			18 on standard; 6 on abbreviated	Bickel, G., Nord, M., Price, C., Maloney, W., & Cook, J. (2000). Guide to Measuring Household Food Security: Measuring Food Security in the United States: Reports of the Federal Interagency Food Security Measurement Project. Washington, DC: USDA Food and Nutrition Service Office of Analysis, Nutrition, and Evaluation. https://www.fns.usda.gov/guide-measuring-household-food-security-revised-2000 . Council, N. R., & Council, N. R. (2006). Food insecurity and hunger in the United States: An assessment of the measure. Washington, DC: The National Academies Press. doi, 10, 11578. https://doi.org/10.17226/11578 .
Financial strain	Financial Strain Index	Levine and Chase-Lansdale 2000	Cronbach's alpha of .81	Data from Families in Communities Study, study of mother-adolescent daughter pairs in three high poverty	6 questions coded on 4 or 5 level scales	Levine, R., & Chase-Lansdale, P. L. (2000). Welfare receipt, financial strain, and African-American adolescent functioning. <i>Social Service Review</i> , 74(3), 380-404. https://doi.org/10.1086/516410 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation(s)
				neighborhoods in the south side of Chicago.		
Financial hardship (broken into domains of food, housing, health care, and child care)	n/a	Economic Policy Institute	No discussion of reliability but discusses validity for each individual domain. Questions are drawn from two national surveys: the Survey of Income and Program Participation (SIPP) and the National Survey of American Families (NSAF)	Data from NSAF, cross-sectional survey of families' experiences that focuses on a number of hardships and family-stress indicators	15 total; 4 for food, 5 for housing, 2 for healthcare, and 3 for child care	Boushey, H. & Gundersen, B. (2001). "When Work Just Isn't Enough: Measuring Hardships Faced by Families after Moving from Welfare to Work." Briefing Paper. Washington, DC: Economic Policy Institute. https://files.eric.ed.gov/fulltext/E_D457249.pdf .
Financial wellbeing	CFPB Financial Wellbeing Scale	Consumer Financial Protection Bureau	- Validated using comparisons to common existing measures or empirical measures. Validation statistics in Appendix D of CFPB report on development. - For phone survey with those 18-61, Cronbach's alpha of .90 for standard scale and .84 for 5 item scale	10k sample of general population; demographics reported in Appendix A of CFPB report on development.	12 on standard, 7 on abbreviated	Bureau, C. F. P. (2015). Measuring financial well-being: A guide to using the CFPB financial well-being scale. <i>Washington, DC: Consumer Financial Protection Bureau.</i> https://files.consumerfinance.gov/f/documents/201705_cfpb_financial-well-being-scale-technical-report.pdf .

Social Capital Measures

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
Individual-level social capital	Resource Generator	van der Gaag & Snijders 2005	Robust reliability based on results from exploratory factor analyses (α^2 (29) $\frac{1}{4}$ 30.93, p $\frac{1}{4}$ 0.37; CFI $\frac{1}{4}$ 0.994; TLI $\frac{1}{4}$ 0.996; RMSEA	Primarily African American urban residents in Midwestern city	26	Foster, K. A., & Maas, C. D. (2014). An exploratory factor analysis of the resource generator-United States: A social capital measure. <i>The British Journal of Social</i>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			¼ 0.024; AIC ¼ 1917.26; BIC ¼ 1863.71)			<i>Work</i> , 46(1), 8-26. https://doi.org/10.1093/bjsw/bcu111 .
Social network	Social Network Instrument	Jason et al. 2014	Cronbach's alpha of 0.85	Oxford House recovery home participants across the US	6	Jason, L. A., & Stevens, E. (2017). The reliability and reciprocity of a social network measure. <i>Alcoholism treatment quarterly</i> , 35(4), 317-327. https://doi.org/10.1080/07347324.2017.1355220 .
Frequency of receipt of verbal and behavioral expressions of social support	Inventory of Socially Supportive Behaviors (ISSB)	Barrera 1981	-Excellent internal consistency: Cronbach's alpha between 0.93 and 0.94 for first and second testing sessions -Evidence for validity of ISSB comes from its moderate direct correlation with social support network size (r=0.32 to r=0.42 as assessed by the Arizona Social Support Interview Schedule	71 undergraduate students	40	Gottlieb, B. H., & Bergen, A. E. (2010). Social support concepts and measures. <i>Journal of psychosomatic research</i> , 69(5), 511-520. https://doi.org/10.1016/j.jpsychores.2009.10.001 .
Social support	Social Provisions Scale (SPS)	Cutrona & Russell 1990	-Excellent overall internal consistency with combined Cronbach's alpha of 0.92 -Adequate internal consistency of subscales (range of 0.65-0.76)	1792 respondents across several studies (including students, nurses, and public school teachers)	24	Gottlieb, B. H., & Bergen, A. E. (2010). Social support concepts and measures. <i>Journal of psychosomatic research</i> , 69(5), 511-520. https://doi.org/10.1016/j.jpsychores.2009.10.001 .
Availability of support from any network member	ENRICHD Social Support Inventory (ESSI)	Mitchell et al. 2003	-Good internal consistency with Cronbach's alpha of 0.86 -Evidence for convergent validity comes from moderate correlation of ESSI scores with Perceived Social Support Scale	196 patients recovering from myocardial infarction across the US	7	Gottlieb, B. H., & Bergen, A. E. (2010). Social support concepts and measures. <i>Journal of psychosomatic research</i> , 69(5), 511-520. https://doi.org/10.1016/j.jpsychores.2009.10.001 .
Social support	Personal Resource	Weinert 1981	-Cronbach's alpha of 0.92 for PRQ85	899 participants from the Family Health Study, a nation-wide study of	25 (10 in part 1/PRQ85	Weinert, C. (2000). Measuring Social Support: PRQ2000. In Strickland, O.L. & Dilorio

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
	Questionnaire (PRQ)		-Alpha ranging from 0.87 to 0.93 for PRQ2000	families managing multiple sclerosis	, 15 in part 2/PRQ2000)	(Eds.) <i>Measurement of Nursing Outcomes: Self Care and Coping</i> (pp. 161–73). Springer.
Loneliness levels	Revised UCLA Loneliness Scale (R-ULS-6)	UCLA	-Rasch person reliability of 0.8 -Cronbach's alpha of 0.82 -Item reliability of 0.98	Thai medical students	6	Wongpakaran, N., Wongpakaran, T., Pinyopornpanish, M., Simcharoen, S., Suradom, C., Varnado, P., & Kuntawong, P. (2020). Development and validation of a 6-item Revised UCLA Loneliness Scale (RULS-6) using Rasch analysis. <i>British Journal of Health Psychology, 25</i> (2), 233-256. https://doi.org/10.1111/bjhp.12404 .
Emotional and social loneliness	6-Item De Jong Gierveld Loneliness Scale	De Jong Gierveld	-Alpha coefficients varied between 0.7 and 0.76 -All correlations above 0.3	8154 Dutch people between the ages of 18 to 79; 4659 Dutch people between the ages of 21 to 99	6	De Jong Gierveld, J., & Van Tilburg, T. (2010). The De Jong Gierveld short scales for emotional and social loneliness: tested on data from 7 countries in the UN generations and gender surveys. <i>European journal of ageing, 7</i> , 121-130. https://doi.org/10.1007/s10433-010-0144-6 .
Social capital	University of Minnesota Scale/Our Community	University of Minnesota	-Strong goodness-of-fit model -Convergent validity: statistically significant with mostly moderate correlations (strongest among three trust scales, strong correlation between bonding trust and bonding engagement) -Internal reliability coefficients for all seven scales were high, with all but one above 0.7 (bonding trust scale alpha was 0.669)	1293 adults in four Minnesota communities (three rural, one located close to major metro area). Disproportionately female, well educated, and wealthy	34	Chazdon, S., Scheffert, D. R., Allen, R., & Horntvedt, J. (2013). Developing and validating University of Minnesota Extension's social capital model and survey. http://conservancy.umn.edu/handle/11299/171657 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
Bridging social capital	Bridging Social Capital (BSC) questionnaire	Villalonga-Olives et al. 2016	-Good internal consistency in all subscales - Cronbach alphas ranged from 0.8 to 0.92 -Good factor analysis results, CFI and TLI higher than 0.9 in almost all scales	Latinx immigrant population in US (Maryland)	61 (v2)	Villalonga-Olives, E., Adams, I., & Kawachi, I. (2016). The development of a bridging social capital questionnaire for use in population health research. <i>S.S.M-population Health</i> , 2, 613-622. https://doi.org/10.1016/j.ssmph.2016.08.008 Villalonga-Olives, E., Kawachi, I., & Rodríguez, A. M. (2021). Rasch model of the bridging social capital questionnaire. <i>S.S.M-Population Health</i> , 14, 100791. https://doi.org/10.1016/j.ssmph.2021.100791 .
Social capital	Personal Social Capital Scale (PSCS)	Chen, Stanton & Gong 2009	-PSCS-16: Cronbach's alpha of 0.9; PSCS-8: Cronbach's alpha of 0.83 -Highly correlated with original PSCS	Rural-to-urban migrants, non-migrant rural residents, and urban residents ages 18-50 in Wuhan	42 (original); 8-item and 16-item version	Wang, P., Chen, X., Gong, J., & Jacques-Tiura, A. J. (2014). Reliability and Validity of the Personal Social Capital Scale 16 and Personal Social Capital Scale 8: Two Short Instruments for Survey Studies. <i>Social Indicators Research</i> , 119(2), 1133–1148. https://www.jstor.org/stable/24721474 .
Sense of community	Brief Sense of Community Scale (BSCS)	Peterson, N.A., Speer, P.W., and McMillan, D. 2008	-Cronbach's alpha of .92 for overall BSBC. -Cronbach's alpha of .86 for needs fulfillment subscale. -Cronbach's alpha of .94 for group membership subscale. -Cronbach's alpha of .77 for influence subscale. -Cronbach's alpha of .87 for emotional connection subscale	293 randomly selected individuals in survey evaluating a community health promotion initiative in the midwestern U.S.	8 items; made up of 4 sets of 2 items that make up their own subscale.	Peterson, N.A., Speer, P.W. & McMillan, D. (2008). Validation of a brief sense of community scale: Confirmation of the principal theory of sense of community. <i>Journal of Community Psychology</i> , 36, 61-73.
Frequency of receipt of verbal and behavioral expression	Social Capital Questionnaire	Onyx & Bullen 2000	-Low to modest factor correlations	496 randomly selected adults in Midwestern community of approx 350,000 residents. Nationally representative	36	O'Brien, M. S., Burdsal, C. A., & Molgaard, C. A. (2004). Further development of an Australian-based measure of social capital in a US sample. <i>Social Science & Medicine (1982)</i> , 59(6), 1207–1217.

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
s of social support				based on 2000 Census data		https://doi.org/10.1016/j.socscimed.2004.01.007 .

Skills, Knowledge, and Experience Measures

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
Work-life balance	Job-Nonjob Conflicts	Sternthal, Slopen, & Williams 2011	Cronbach's alpha of .91	Data from the Chicago Community Adult Health Study (CCAHS), a household probability sample of 3,105 adults aged 18 and over living in Chicago, Illinois stratified into 343 neighborhood clusters.	2 item mean index, range 1-4	Sternthal, M. J., Slopen, N., & Williams, D. R. (2011). Racial disparities in health: how much does stress really matter? 1. <i>Du Bois review: social science research on race</i> , 8(1), 95-113.
Job hazards	Job Hazards	Sternthal, Slopen, & Williams 2011	Cronbach's alpha of .71	Data from the Chicago Community Adult Health Study (CCAHS), a household probability sample of 3,105 adults aged 18 and over living in Chicago, Illinois stratified into 343 neighborhood clusters.	3 item mean index, range 1-3	Sternthal, M. J., Slopen, N., & Williams, D. R. (2011). Racial disparities in health: how much does stress really matter? 1. <i>Du Bois review: social science research on race</i> , 8(1), 95-113.
Parenting	Alabama Parenting Measure	Frick 1991	Acceptable psychometric ratings: internal reliability, test-retest reliability, cross informant reliability, face validity, convergent validity, discriminant validity, exploratory factor analysis, confirmatory factor analysis	Diverse populations as reported by manuals and articles; range of ages for children of the parent varied from 1 to 18 years of age	42	Hurley, K. D., Huscroft-D'Angelo, J., Trout, A., Griffith, A., & Epstein, M. (2014). Assessing parenting skills and attitudes: A review of the psychometrics of parenting measures. <i>Journal of Child and Family Studies</i> , 23, 812-823. https://doi.org/10.1007/s10826-013-9733-2 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
Parenting	Parenting Alliance Measure	Abidin and Konold 1999	Acceptable psychometric ratings: internal reliability, test-retest reliability, cross informant reliability, face validity, predictive validity, convergent validity, discriminant validity, exploratory factor analysis	Diverse populations as reported by manuals and articles; range of ages for children of the parent varied from 1 to 18 years of age	20	Hurley, K. D., Huscroft-D'Angelo, J., Trout, A., Griffith, A., & Epstein, M. (2014). Assessing parenting skills and attitudes: A review of the psychometrics of parenting measures. <i>Journal of Child and Family Studies, 23</i> , 812-823. https://doi.org/10.1007/s10826-013-9733-2 .
Parenting	Parenting Scale	Arnold et al. 1993	Acceptable psychometric ratings: internal reliability, test-retest reliability, cross informant reliability, face validity, convergent validity, discriminant validity, exploratory factor analysis, confirmatory factor analysis	Diverse populations as reported by manuals and articles; range of ages for children of the parent varied from 1 to 18 years of age	30	Hurley, K. D., Huscroft-D'Angelo, J., Trout, A., Griffith, A., & Epstein, M. (2014). Assessing parenting skills and attitudes: A review of the psychometrics of parenting measures. <i>Journal of Child and Family Studies, 23</i> , 812-823. https://doi.org/10.1007/s10826-013-9733-2 .
Parenting behaviors of caregivers of young children			-Construct validity: all items were adequately correlated, and all items under the same theoretical heading were significantly correlated with each other	579 primary caregivers (mostly biological mothers) initially recruited from WIC programs in Pittsburgh, PA, Eugene, OR, and Charlottesville, VA. Racially diverse sample.	21	McEachern, A. D., Dishion, T. J., Weaver, C. M., Shaw, D. S., Wilson, M. N., & Gardner, F. (2012). Parenting Young Children (PARYC): Validation of a self-report parenting measure. <i>Journal of child and family studies, 21</i> , 498-511. https://doi.org/10.1007/s10826-011-9503-y .

Mental Well-Being Measures

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
Depression, anxiety, tension/stress	Depression Anxiety Stress Scale-21 (DASS-21)	Psychology Foundation of Australia	-A systemic review of mental health measurement scales found DASS-21 to be a widely used measurement of a range of mental health symptoms and is appropriate for use with a general, non-clinical population. Additionally, the	-DASS was tested on undergraduate psychology students (van Ballegooijen et al., 2016) -DASS and DASS-21 were tested on nonclinical volunteers and outpatients with psychiatric diagnoses	42- and 21-item versions	Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. <i>Psychological Assessment, 10</i> , 176-181. https://doi.org/10.1037/1040-3590.10.2.176 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>review found evidence to support its validity and internal consistency. The review also found that DASS-21 has been successfully used in a variety of cultures and countries and is available in multiple languages (Breedvelt et al., 2020)</p> <p>-A systemic review of online instruments for common mental health disorders found evidence for strong internal consistency (Cronbach's alpha ranged from 0.93 to 0.95 across subscales) (van Ballegooijen et al., 2016)</p> <p>-Strong internal consistency of DASS-42 (Cronbach's alpha ranging from 0.92 to 0.97 across subscales) and DASS-21 (Cronbach's alpha ranging from 0.87 to 0.94 across subscales) (Antony et al., 1998)</p> <p>-Concurrent validity of DASS-42 and DASS-21 in the acceptable to excellent ranges (Antony et al., 1998)</p> <p>-DASS-21 has a cleaner factor structure and smaller interfactor correlations relative to DASS-42 (Antony et al., 1998)</p> <p>-DASS-42 and DASS-21 distinguish well between features of depression, physical arousal, and</p>	<p>between 18 and 65 years (Antony et al., 1998)</p> <p>-DASS-21 was tested on a large, diverse sample matched to distributions of gender, age, race, ethnicity, education, and income based on US 2004 Census Bureau Data (Sinclair et al., 2012)</p> <p>-DASS-21 was tested on a sample of undergraduates in Houston of four racial groups (African American/Black (non-Hispanic), Caucasian/White (non-Hispanic), Hispanic/Latino(a), and Asian ages 17-51 (mean = 20.67 years) (Norton, 2007)</p>		<p>Breedvelt, J. J. F., Zamperoni, V., South, E., Uphoff, E. P., Gilbody, S., Bockting, C. L. H., Churchill, R., & Kousoulis, A. A. (2020). A systematic review of mental health measurement scales for evaluating the effects of mental health prevention interventions. <i>European Journal of Public Health</i>, 30(3), 510–516. https://doi.org/10.1093/eurpub/ckz233.</p> <p>Norton, P. J. (2007). Depression Anxiety and Stress Scales (DASS-21): Psychometric analysis across four racial groups. <i>Anxiety, Stress, & Coping</i>, 20(3), 253–265. https://doi.org/10.1080/10615800701309279.</p> <p>Sinclair, S. J., Siefert, C. J., Slavin-Mulford, J. M., Stein, M. B., Rennan, M., & Blais, M. A. (2012). Psychometric Evaluation and Normative Data for the Depression, Anxiety, and Stress Scales-21 (DASS-21) in a Nonclinical Sample of U.S. Adults. <i>Evaluation & the Health Professions</i>, 35(3), 259–279. https://doi.org/10.1177/0163278711424282.</p> <p>van Ballegooijen, W., Riper, H., Cuijpers, P., van Oppen, P., & Smit, J. H. (2016). Validation of online psychometric instruments for common mental health disorders: A systematic review. <i>BMC Psychiatry</i>, 16(1), 45. https://doi.org/10.1186/s12888-016-0735-7.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>psychological tension and agitation (Antony et al., 1998)</p> <p>-Acceptable internal consistency reliability of DASS-21 with Cronbach's alpha ranging from 0.80 to 0.91 across subscales (Sinclair et al., 2012)</p> <p>-Correlations between DASS-21 scales and other concurrent measures were higher than reported elsewhere (r ranged from 0.68 to 0.73) (Sinclair et al., 2012)</p> <p>-DASS-21 demonstrated good convergent validity with Mental Health Summary score ($r = -0.58$ to -0.69) and divergent validity with the Physical Component Summary score ($r = -0.16$ to -0.34) from the SF-8 Health Survey (Sinclair et al., 2012)</p> <p>-Principal components analysis supported the extraction of only one component accounting for 47% of the item-level variance, though confirmatory factor analysis (CFA) favored a three-factor structure when compared to a one-factor model (Sinclair et al., 2012)</p> <p>-DASS-21 showed strong internal consistency across four racial groups, with Cronbach's alpha ranging from 0.778 to 0.871 across the</p>			<p>Zlomke, K. R. (2009). Psychometric properties of internet administered versions of Penn State Worry Questionnaire (PSWQ) and Depression, Anxiety, and Stress Scale (DASS). <i>Computers in Human Behavior</i>, 25(4), 841–843. https://doi.org/10.1016/j.chb.2008.06.003.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>three subscales for all subgroups, and similar internal consistency for each subgroup (Norton, 2007)</p> <p>-Multi-group confirmatory factor analysis of DASS-21 with a sample of four racial groups indicated that item loadings were invariant while scale covariances were not invariant, suggesting that the depression, stress, and anxiety maybe be differentially inter-related across groups (Norton, 2007)</p> <p>-Convergent and divergent validity of the DASS-21 were found to be similar across four racial groups (Norton, 2007)</p>			
Positive and negative affect	Positive and Negative Affect Schedule (PANAS)	University of Minnesota and Southern Methodist University	<p>-A systemic review of mental health measurement scales found evidence to support the construct validity, criterion validity and internal consistency of PANAS as a measurement of general mental health (Breedvelt et al., 2020)</p> <p>-High internal consistency reliabilities of the PANAS PA and NA scales with Cronbach's alpha values of 0.89 and 0.85, respectively (Crawford & Henry, 2004)</p> <ul style="list-style-type: none"> Demographic variables had only very modest influences on PANAS scores and the PANAS exhibited 	<p>-PANAS was tested with a non-clinical sample broadly representative of the general adult UK population (Crawford & Henry, 2004)</p> <p>-I-PANAS-SF was tested with a large, culturally diverse sample, although the sample population may represent a wealthier and more educated subset (Thompson, 2007)</p> <p>-PANAS-X was tested with a variety of populations, including undergraduates, adults, and psychiatric patients (Clark & Watson, 1994)</p> <p>-PANAS was tested with undergraduate students and employees from a private</p>	Original scale has 20 items, short form version has 10 items	<p>Bagozzi, R. P. (1993). An examination of the psychometric properties of measures of negative affect in the PANAS-X scales. <i>Journal of Personality and Social Psychology</i>, 65, 836–851. https://doi.org/10.1037/0022-3514.65.4.836.</p> <p>Breedvelt, J. J. F., Zamperoni, V., South, E., Uphoff, E. P., Gilbody, S., Bockting, C. L. H., Churchill, R., & Kousoulis, A. A. (2020). A systematic review of mental health measurement scales for evaluating the effects of mental health prevention interventions. <i>European Journal of Public Health</i>, 30(3), 510–516. https://doi.org/10.1093/eurpub/ckz233.</p> <p>Clark, L. A., & Watson, D. (1994). The PANAS-X: Manual for the Positive and</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>measurement invariance across demographic subgroups (Crawford & Henry, 2004)</p> <ul style="list-style-type: none"> • Cross-sample stability, internal reliability, temporary stability, cross-cultural factorial invariance, and convergent and criterion validity of I-PANAS-SF were found to be psychometrically acceptable (Thompson, 2007) • High internal consistency reliability of the PANAS-X subscales (Cronbach's alpha ranging from 0.83 to 0.90 for PA and 0.85 to 0.90 for NA) (Clark & Watson, 1994) • High convergent correlations and low divergent correlations of the PANAS-X subscales (Clark & Watson, 1994) • High internal consistency reliability of the PANAS scale ranging from 0.86 to 0.90 on the PANAS PA subscale and 0.84 to 0.87 on the PANAS NA subscale (Watson et al., 1988) • The PANAS NA and PA subscales had low intercorrelations, were stable over a 2-month time period, had strong convergent and discriminant correlations with longer measures of underlying mood factors, and correlated at predicted levels with 	southwestern university in the US, as well as adults not affiliated with the university (Watson et al., 1988)		<p>Negative Affect Schedule - Expanded Form [Data set]. University of Iowa. https://doi.org/10.17077/48vt-m4t2.</p> <p>Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. <i>British Journal of Clinical Psychology</i>, 43(3), 245–265. https://doi.org/10.1348/0144665031752934.</p> <p>Thompson, E. R. (2007). Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). <i>Journal of Cross-Cultural Psychology</i>, 38(2), 227–242. https://doi.org/10.1177/0022022106297301.</p> <p>Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. <i>Journal of Personality and Social Psychology</i>, 54, 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			measures of related constructs (Watson et al., 1988)			
General psychological distress	General Health Questionnaire (GHQ-12)	Developed by British Scholar D.P. Goldberg	<p>-A systemic review of mental health measurement scales found evidence to support the construct and criterion validity, internal consistency, and test-retest reliability of GHQ-12 as a measurement of general mental health (Breedvelt et al., 2020)</p> <p>-A bifactor model showed that GHQ-12 and WEMWBS items assess mainly the same construct (Böhnke & Croudace, 2016)</p> <p>-GHQ-12 had high internal consistency (with Cronbach's alpha ranging from 0.82 to 0.90) and unidimensional factor structure across three distinct samples (Banks et al., 1980)</p> <p>-Gender, age, and educational level were shown to have no significant effect on the validity of the GHQ (Goldberg et al., 1997)</p> <p>-One study found that traditional psychometric assessments using factor analysis and reliability estimates have obscured substantial measurement error in the GHQ-12 due to response bias on the negative items (Hankins, 2008)</p> <p>-Substantial factor variation</p>	<p>-A large sample of respondents of the Health Survey for England (Böhnke & Croudace, 2016; Hankins, 2008)</p> <p>-Employees in an engineering firm, recent school-leavers, and unemployed men (Banks et al., 1980)</p> <p>-Large samples of patients at health clinics around the world (Goldberg et al., 1997; Werneke et al., 2000)</p>	28-item and 12-item versions	<p>Banks, M. H., Clegg, C. W., Jackson, P. R., Kemp, N. J., Stafford, E. M., & Wall, T. D. (1980). The use of the General Health Questionnaire as an indicator of mental health in occupational studies. <i>Journal of Occupational Psychology</i>, 53(3), 187–194. https://doi.org/10.1111/j.2044-8325.1980.tb00024.x.</p> <p>Böhnke, J. R., & Croudace, T. J. (2016). Calibrating well-being, quality of life and common mental disorder items: Psychometric epidemiology in public mental health research. <i>The British Journal of Psychiatry</i>, 209(2), 162–168. https://doi.org/10.1192/bjp.bp.115.165530.</p> <p>Breedvelt, J. J. F., Zamperoni, V., South, E., Uphoff, E. P., Gilbody, S., Bockting, C. L. H., Churchill, R., & Kousoulis, A. A. (2020). A systematic review of mental health measurement scales for evaluating the effects of mental health prevention interventions. <i>European Journal of Public Health</i>, 30(3), 510–516. https://doi.org/10.1093/eurpub/ckz233.</p> <p>Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. <i>Psychological Medicine</i>, 27(1), 191–197. https://doi.org/10.1017/S0033291796004242.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			for the GHQ-12 was found between participating health centers (less factor variation was found for the GHQ-28). Validity as a case detector was not affected by variance in factor structure (Werneke et al., 2000)			<p>Hankins, M. (2008). The reliability of the twelve-item general health questionnaire (GHQ-12) under realistic assumptions. <i>BMC Public Health</i>, 8(1), 355. https://doi.org/10.1186/1471-2458-8-355.</p> <p>Hardy, G. E., Shapiro, D. A., Haynes, C. E., & Rick, J. E. (1999). Validation of the General Health Questionnaire-12: Using a sample of employees from England's health care services. <i>Psychological Assessment</i>, 11, 159–165. https://doi.org/10.1037/1040-3590.11.2.159.</p> <p>McCabe, C. J., Thomas, K. J., Brazier, J. E., & Coleman, P. (1996). Measuring the Mental Health Status of a Population: A Comparison of the GHQ-12 and the SF-36 (MHI-5). <i>The British Journal of Psychiatry</i>, 169(4), 517–521. https://doi.org/10.1192/bjp.169.4.516.</p> <p>Werneke, U., Goldberg, D. P., Yalcin, I., & Üstün, B. T. (2000). The stability of the factor structure of the General Health Questionnaire. <i>Psychological Medicine</i>, 30(4), 823–829. https://doi.org/10.1017/S0033291799002287.</p>
Positive mental health	Warwick-Edinburgh Mental Well-being Scale (WEMWBS)	NHS Health Scotland, University of Warwick and University of Edinburgh	<p>-A bifactor model showed that GHQ-12 and WEMWBS items assess mainly the same construct (Böhnke & Croudace, 2016)</p> <p>-One study found that WEMWBS had good content validity, and confirmatory factor analysis supported that</p>	<p>-A large sample of respondents from three waves of the Health Survey for England (Böhnke & Croudace, 2016)</p> <p>-A large sample of respondents from the Scottish Health Education Population Survey in Autumn 2006 in the</p>	14-item and 7-item versions	<p>Böhnke, J. R., & Croudace, T. J. (2016). Calibrating well-being, quality of life and common mental disorder items: Psychometric epidemiology in public mental health research. <i>The British Journal of Psychiatry</i>, 209(2), 162–168. https://doi.org/10.1192/bjp.bp.115.165530.</p> <p>Stewart-Brown, S. L., Platt, S., Tennant, A.,</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>the scale measured a single construct (Tennant et al., 2007)</p> <p>-Internal consistency of WEMWBS indicated some item redundancy in the scale (with a Cronbach's alpha score of 0.89 and 0.91 for the student and population samples, respectively) (Tennant et al., 2007)</p> <p>-WEMWBS showed high correlations with other mental health and well-being scales and lower correlations with scales measuring overall health, had a near normal distribution and no ceiling effects in a population sample, and discriminated between population groups in a way consistent with the results of other population surveys (Tennant et al., 2007)</p> <p>-Test-retest reliability of WEMWBS at one week was high (Tennant et al., 2007)</p> <p>-One study found that a short 7-item version (SEMWBS) provides an interval scale estimate of mental well-being, as it satisfied strict one-dimensionality expectations of the Rasch model and was largely free of bias (S. Stewart-Brown et al., 2009)</p> <p>-WEMWBS did not meet the strict criteria for measurement</p>	<p>age range of 16-74 years (average age = 41.9 years) (S. Stewart-Brown et al., 2009)</p> <p>-Undergraduate and postgraduate student sample from universities in the UK and a representative Scottish population sample from 2006 datasets (Tennant et al., 2007)</p>		<p>Maheswaran, H., Parkinson, J., Weich, S., Tennant, R., Taggart, F., & Clarke, A. (2011). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): A valid and reliable tool for measuring mental well-being in diverse populations and projects. <i>J Epidemiol Community Health</i>, 65(Suppl 2), A38–A39. https://doi.org/10.1136/jech.2011.143586.86.</p> <p>Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J., & Weich, S. (2009). Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): A Rasch analysis using data from the Scottish Health Education Population Survey. <i>Health and Quality of Life Outcomes</i>, 7(1), 15. https://doi.org/10.1186/1477-7525-7-15.</p> <p>Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. <i>Health and Quality of Life Outcomes</i>, 5(1), 63. https://doi.org/10.1186/1477-7525-5-63.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			as it demonstrated Differential Item Functioning and multidimensionality (S. Stewart-Brown et al., 2009)			
Positive mental health	Positive Mental Health Scale (PMH-scale)	Lukat et al., 2016	-9-item PMH-scale was confirmed to be a unidimensional instrument with high internal consistency (Cronbach's alpha was .93 for all groups together), good test-retest reliability, scalar invariance across samples and over time, good convergent and discriminant validity, and sensitivity to therapeutic change (Lukat et al., 2016) -8-item PMH-scale proved to be unidimensional and had strong reliability (Person Separation Index = 0.89) (Vaganian et al., 2022)	-Student, patient, and representative German adult populations (Lukat et al., 2016) -Adult cancer patients, with a greater female portion of the sample (80.7%) relative to male (19.0%) (Vaganian et al., 2022)	9-item and 8-item versions	Lukat, J., Margraf, J., Lutz, R., van der Veld, W. M., & Becker, E. S. (2016). Psychometric properties of the Positive Mental Health Scale (PMH-scale). <i>BMC Psychology</i> , 4(1), 8. https://doi.org/10.1186/s40359-016-0111-x . Vaganian, L., Boecker, M., Bussmann, S., Kusch, M., Labouvie, H., Margraf, J., Gerlach, A. L., & Cwik, J. C. (2022). Psychometric evaluation of the Positive Mental Health (PMH) scale using item response theory. <i>BMC Psychiatry</i> , 22(1), 512. https://doi.org/10.1186/s12888-022-04162-0 .
Executive functioning	Behavior Rating Inventory of Executive Function-Adult Version (BRIEF-A)	Western Psychological Services	-Moderate to high alphas obtained for both BRIEF-A normative and mixed clinical samples: ranging from 0.9 to 0.93 in normative samples and 0.85 to 0.95 in mixed clinical samples -Inter-rater reliability: 0.57 reliability coefficient mean with range of 0.44-0.68 (Roth & Isquith 2014)	-BRIEF-A standardization sample collected with goal of approximating population of US; even split for gender; 72.6% white; age range of 18-90; spread out across geographic regions of Northeast, Midwest, South, and West; N = 1050 (Roth & Isquith 2014)	75 items	Baggetta, P., & Alexander, P. A. (2016). Conceptualization and Operationalization of Executive Function. <i>Mind, Brain, and Education</i> , 10(1), 10–33. https://doi.org/10.1111/mbe.12100 . Gioia, G. A., Isquith, P. K., & Kenealy, L. E. (2008). Assessment of behavioral aspects of executive function. In Anderson, V., Jacobs, R., & Anderson, P.J. (Eds.), <i>Executive Functions and the Frontal Lobes</i> . Psychology Press. Isquith, P. K., Roth, R. M., & Gioia, G. (2013). Contribution of Rating Scales to the Assessment of Executive Functions. <i>Applied</i>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
						<p><i>Neuropsychology: Child</i>, 2(2), 125–132. https://doi.org/10.1080/21622965.2013.748389.</p> <p>Roth, R. M., Isquith, P. K., & Gioia, G. A. (2014). Assessment of Executive Functioning Using the Behavior Rating Inventory of Executive Function (BRIEF). In S. Goldstein & J. A. Naglieri (Eds.), <i>Handbook of Executive Functioning</i> (pp. 301–331). Springer. https://doi.org/10.1007/978-1-4614-8106-5_18.</p>
Executive functioning	Dysexecutive Questionnaire (DEX)	Wilson et al. 1996	<p>-Good internal consistency: Cronbach's alpha was 0.9 for community sample, 0.91 for psychiatric sample, 0.91 for neurologically impaired sample (Shaw & Sawang 2015)</p> <p>-Responses did not show fit to the Rasch model, suggesting it measures more than one subconstruct (Wakely 2022)</p> <p>-Excellent internal consistency with Cronbach's alpha of 0.93 (Wakely 2022)</p> <p>-Good concurrent validity when compared to responses given to FrSBe (other validated measure of dysexecutive problems), correlation between DEX-R and FrSBe was $r = 0.98$, $p < 0.01$ (Wakely 2022)</p>	<p>-997 individuals across three settings (51% male); community sample (n=663), psychiatric sample (depressed patients, n=92; anxiety patients, n=122), and neurologically impaired sample (n=120). Psychiatric outpatients recruited from CBT unit at psychiatric hospital, neurological patients recruited from head injuries unit from large public general hospital, community sample was mix of university students and other individuals. Study conducted in Australia, age range of 15 to 72, median age of 23 (Shaw & Sawang 2015)</p> <p>-125 participants between 19 and 69 years, 80% female, 77% White British (Wakely 2022)</p>	20 items	<p>Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. <i>Archives of Clinical Neuropsychology</i>, 21(3), 217–227. https://doi.org/10.1016/j.acn.2005.12.002.</p> <p>Gioia, G. A., Isquith, P. K., & Kenealy, L. E. (2008). Assessment of behavioral aspects of executive function. In Anderson, V., Jacobs, R., & Anderson, P.J. (Eds.), <i>Executive Functions and the Frontal Lobes</i>. Psychology Press.</p> <p>Mooney, B., Walmsley, C., & McFarland, K. (2006). Factor Analysis of the Self-Report Dysexecutive (DEX-S) Questionnaire. <i>Applied Neuropsychology</i>, 13(1), 12–18. https://doi.org/10.1207/s15324826an1301_2.</p> <p>Shaw, S., Oei, T. P. S., & Sawang, S. (2015). Psychometric validation of the Dysexecutive Questionnaire (DEX). <i>Psychological Assessment</i>, 27, 138–147.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
						https://doi.org/10.1037/a0038195 . Wakely, H., Radakovic, R., Bateman, A., Simblett, S., Fish, J., & Gracey, F. (2022). Psychometric Properties of the Revised Dysexecutive Questionnaire in a Non-clinical Population. <i>Frontiers in Human Neuroscience</i> , 16, 767367. https://doi.org/10.3389/fnhum.2022.767367 .
Executive functioning	Brock Adaptive Functioning Questionnaire (BAFQ)	Dywan & Segalowitz 1996	-Adequate internal consistency reliability (alpha = 0.92)	-46 adults recruited from outpatients referred for neuropsychological assessment (epilepsy, traumatic brain injury, tumor, MS, etc.). Age: M = 40.87, range = 19-75 years; 56% male; avg level of education = 13.48 years (range = 9-21 years) (Chaytor 2006)	68 items	Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. <i>Archives of Clinical Neuropsychology</i> , 21(3), 217–227. https://doi.org/10.1016/j.acn.2005.12.002 .
Executive functioning	Executive Function Index (EFI)	Spinella 2005	-Pearson correlations coefficients were obtained between scales of the EFI, FrSBe, BIS, and IRI: inverse correlations between the all scales of the EFI with the all scales of the FrSBe; inverse correlations occurred between the all BIS scales and MD, IC, ORG, and SP whereas only BISnp correlated with EM; both IRI scales correlated positively with MD, IC, and EM whereas only IRIpt correlated positively with SP (Spinella 2005)	-188 adults (81 male, 107 female), ages 17 to 60 years (M = 26.6), completed between 11 and 18 years of formal education (M = 14.5) (Spinella 2005)	27 items	Baggetta, P., & Alexander, P. A. (2016). Conceptualization and Operationalization of Executive Function. <i>Mind, Brain, and Education</i> , 10(1), 10–33. https://doi.org/10.1111/mbe.12100 . Gioia, G. A., Isquith, P. K., & Kenealy, L. E. (2008). Assessment of behavioral aspects of executive function. In Anderson, V., Jacobs, R., & Anderson, P.J. (Eds.), <i>Executive Functions and the Frontal Lobes</i> . Psychology Press. Spinella, M. (2005). Self-Rated Executive Function: Development of the Executive Function Index. <i>International Journal of Neuroscience</i> , 115(5), 649–667.

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
						https://doi.org/10.1080/00207450590524304 .
Executive functioning	Frontal Behavior Inventory (FBI)	Kertesz, Davidson, & Fox 2000	-High interrater reliability (Cohen's Kappa = 0.89), high internal consistency (Cronbach's alpha = 0.89) -High validity: able to successfully distinguish frontal lobe dementia from several other dementing disorders (Kertesz 2000)	-108 geriatric patients with Alzheimer's disease, frontal lobe dementia, primary progressive aphasia, vascular dementia, and depressive disorders (Kertesz 2000)	24 items	Gioia, G. A., Isquith, P. K., & Kenealy, L. E. (2008). Assessment of behavioral aspects of executive function. In Anderson, V., Jacobs, R., & Anderson, P.J. (Eds.), <i>Executive Functions and the Frontal Lobes</i> . Psychology Press.
Executive functioning	Frontal Systems Behavior Scale	Grace & Malloy 2001	-Reliability: High overall internal consistency (Cronbach's alpha for Total Score: 0.94 on family form, 0.92 on self-report form), similar coefficients reported in additional studies. Test-retest reliability was reported as 0.78 for Total scale (Gioia 2010) -Validity: Construct validity demonstrated in several ways: exploratory factor analysis of FrSBe-Family Version with 324 neurological patients showed three-factor solution accounting for 41% of the variance, five-factor solution was found by Azzara (2005), with 85% of items loading saliently. Convergent and discriminant validity have been documented for the FrSBe (Gioia 2010)	-436 individuals between the ages of 18-95 years of age with T-scores stratified for age, gender, and education (Gioia 2010)	46 items	Gioia, G. A., Isquith, P. K., & Kenealy, L. E. (2008). Assessment of behavioral aspects of executive function. In Anderson, V., Jacobs, R., & Anderson, P.J. (Eds.), <i>Executive Functions and the Frontal Lobes</i> . Psychology Press.
Resilience	Connor-Davidson Resilience	Kathryn M. Connor and Jonathan	-Factor analysis of the 25-item CD-RISC yielded five factors (Connor & Davidson, 2003)	-A community sample, primary care outpatients, general psychiatric outpatients,	2-, 10- and 25-	Burns, R., & Anstey, K. (2010). The Connor-Davidson Resilience Scale (CD-RISC): Testing the invariance of a uni-

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
	Scale (CD-RISC)	R.T. Davidson	<ul style="list-style-type: none"> -The 25-item CD-RISC had good internal consistency with a Cronbach's alpha value of 0.89 and strong test-retest reliability (Connor & Davidson, 2003) -The 25-item CD-RISC demonstrated convergent validity relative to other measures of stress and hardiness (Connor & Davidson, 2003) -Exploratory factor analysis of the 25-item CD-RISC found an unstable factor structure across two demographically equivalent samples (Campbell-Sills & Stein, 2007) -The 10-item CD-RISC demonstrated good internal consistency (Cronbach's alpha was 0.85) and construct validity (Campbell-Sills & Stein, 2007) -The 2-item version of the CD-RISC showed good test-retest reliability, convergent validity, and divergent validity, and showed significant correlation with the overall CD-RISC score as well as with each item of the CD-RISC (Vaishnavi et al., 2007) -The 25-item CD-RISC was found to be a unidimensional measure that was independent of affect at an item level, but there were strong associations 	<ul style="list-style-type: none"> clinical trial of generalized anxiety disorder, and two clinical trials of PTSD; across five of the groups, 65% were female, 35% were male, 77% were white, 23% were non-white, the mean age was 43.8 (SD = 15.3) (Connor & Davidson, 2003) -Undergraduates from San Diego State University; 74.4% were women, the mean age was 18.8 years (SD = 2.2), and the majority of participants were Caucasian (53.1%) (Campbell-Sills & Stein, 2007) -Individuals with PTSD, major depressive disorder, and generalized anxiety disorder; the general population, family medicine clinic patients, and psychiatric outpatients (Vaishnavi et al., 2007) -A large sample of Australian individuals ages 20-24 which was 45.9% male, had a mean education level of 15 years, and 88.3% rated their health as good, very good, or excellent (Burns & Anstey, 2010) 	item versions	dimensional measure that is independent of positive and negative affect. <i>Personality and Individual Differences</i> , 48, 527–533.

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>between resilience and affect factors (Burns & Anstey, 2010)</p> <p>-Comparable Goodness of Fit Indices supported strict invariance between genders on an oblique 3-factor model of resilience and affect for the 25-item CD-RISC (Burns & Anstey, 2010)</p>			
Psychological capital	Psychological Capital Questionnaire (PCQ)	Fred Luthans, Bruce J. Avolio, & James B. Avey	<p>-Strong internal consistency of the PCQ with Cronbach alphas ranging from 0.88-0.89 across four samples (Luthans et al., 2007)</p> <p>-Strong psychometrics supported by confirmatory factor analysis with heterogeneous samples (Luthans et al., 2007)</p> <p>-One analysis identified issues around the construct and structural validity and raised the possibility of problems around response distortion on the self-report PCQ (Harms et al., 2018)</p>	-PCQ was tested with college students, engineers and technicians from a large firm (the average age was 44.83 years and 80% were men), and employees from a mid-sized insurance services firm (the average age was 33.79 years and 65% were women) (Luthans et al., 2007)	12- and 24-item versions	Luthans, F., Avolio, B. J., Avey, J. B., & Norman, S. M. (2007). Positive Psychological Capital: Measurement and Relationship with Performance and Satisfaction. <i>Personnel Psychology</i> , 60(3), 541–572. https://doi.org/10.1111/j.1744-6570.2007.00083.x .
Psychological capital	Implicit Psychological Capital Questionnaire (I-PCQ)	P. D. Harms, Dina V. Krasikova, and Fred Luthans	-I-PCQ demonstrated acceptable structural validity, was resistant to response distortion, and predicted work outcomes above and beyond the widely used self-report PCQ and Big Five personality traits (Harms et al., 2018)	-I-PCQ was tested with students in business school (Harms et al., 2018)	24 items (includes three story prompts with each followed by eight questions)	Harms, P. D., Krasikova, D. V., & Luthans, F. (2018). Not Me, But Reflects Me: Validating a Simple Implicit Measure of Psychological Capital. <i>PLoS ONE</i> , 551–562. https://doi.org/10.1080/00223891.2018.1480489 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
Emotional awareness	Levels of Emotional Awareness Scale (LEAS)	Lane, R. D., Quinlan, D. M., Schwartz, G. E., Walker, P. A., & Zeitlin, S. B.	<ul style="list-style-type: none"> -Various studies have found good internal consistency as measured by Cronbach's alpha for the 20- and 10-item versions (Lane & Smith, 2021) -Various studies have found excellent inter-rater reliability (Lane & Smith, 2021) -One study found good test-retest reliability at 4 weeks (Lane & Smith, 2021) -Various studies support the construct validity of LEAS in its status as an individual difference variable (Lane & Smith, 2021) 	-Healthy adult populations, undergraduate and graduate student populations, clinical populations including those with mental health or systemic medical disorders (Lane & Smith, 2021)	20- and 10-item versions	Lane, R. D., & Smith, R. (2021). Levels of Emotional Awareness: Theory and Measurement of a Socio-Emotional Skill. <i>Journal of Intelligence</i> , 9(3), 42. https://doi.org/10.3390/jintelligence9030042 .
Psychological capital	Compound PsyCap Scale (CPC-12)	Timo Lorenz, Clemens Beer, Jan Pütz, Kathrin Heinitz	<ul style="list-style-type: none"> -Strong external validity demonstrated by moderate to high correlations to other work-related (meaning of work, job satisfaction and engagement; $r = .28-.40$) and more general constructs of positive psychology (i.e., subjective well-being, proactive attitude, and gratitude; $r = .22-.58$) (Lorenz et al., 2016) -The four subscales of hope, optimism, resilience, and self-efficacy were found to be identifiable as subcomponents of the overall measure which the higher-order factor can explain additional variance in the data (Lorenz et al., 2016) 	-German individuals (Lorenz et al., 2016)	12 items	Lorenz, T., Beer, C., Putz, J., & Heinitz, K. (2016). Measuring Psychological Capital: Construction and Validation of the Compound PsyCap Scale (CPC-12). <i>PLoS ONE</i> . https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0152892 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
Psychological need	Balanced Measure of Psychological Needs (BMPN)	Sheldon & Hilpert, 2012	<ul style="list-style-type: none"> -Reliability analyses found alpha coefficients ranging from 0.78 to 0.79 across subscales (Sheldon & Hilpert, 2012) -BMPN demonstrated sound evidence of convergent and discriminant construct validity at the matrix level, and findings suggested that the need satisfaction construct as measured by the BMPN is comprised of three separate need factors measured via satisfaction and dissatisfaction items (Sheldon & Hilpert, 2012) -Responses to the BMPN exceeded rigorous cutoff criteria for model fit (Sheldon & Hilpert, 2012) 	<ul style="list-style-type: none"> -Undergraduate students from the University of Missouri; the sample was 88% Caucasian, 6% African American, 3% Asian American, and 3% “other” (Sheldon & Hilpert, 2012) 	18 items	Sheldon, K. M., & Hilpert, J. C. (2012). The balanced measure of psychological needs (BMPN) scale: An alternative domain general measure of need satisfaction. <i>Motivation and Emotion</i> , 36, 439–451.
Psychological need	Basic Psychological Needs Scale (BPNS)	Gagné, M., 2003	<ul style="list-style-type: none"> -Reliability analyses found alpha coefficients ranging from 0.68 to .84 across subscales (Sheldon & Hilpert, 2012) -Responses to the BPNS did not differentiate adequately between one general need factor or three related but distinguishable need factors and findings suggested that the methods of measurement may bias the measurement of the need factors (Sheldon & Hilpert, 2012) -Responses to the BPNS did not meet rigorous cutoff 	<ul style="list-style-type: none"> -Undergraduate students from the University of Missouri; the sample was 88% Caucasian, 6% African American, 3% Asian American, and 3% “other” (Sheldon & Hilpert, 2012) 	21 items	Sheldon, K. M., & Hilpert, J. C. (2012). The balanced measure of psychological needs (BMPN) scale: An alternative domain general measure of need satisfaction. <i>Motivation and Emotion</i> , 36, 439–451.

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			criteria for model fit (Sheldon & Hilpert, 2012)			
Impulsiveness	Barratt Impulsivity Scale (BIS)	E. S. Barratt, 1985	<p>-Full information item bifactor analysis of the 30-item BIS-11 found no evidence supporting the 3-factor model in which BIS-11 is held to measure the theoretical subtraits of attentional, motor, and non-planning impulsiveness (Steinberg et al., 2013)</p> <p>-Scores obtained on the 8-item unidimensional BIS-BRIEF across three different samples and diverse age ranges had indices of construct validity to those found with BIS-11 total scores (Steinberg et al., 2013)</p>	<p>-BIS-11 was tested with a large sample of undergraduates at a private mid-sized Southern university (77.4% were female, 22.6% were male; the mean age was 19.4 years; 61.2% were Caucasian, 12.8% were Asian/Pacific Islander, 8.8% were African American, and 4.4% self-identified as “other” or were multiracial) (Steinberg et al., 2013)</p> <p>• BIS-BRIEF was tested on a community sample of borderline personality patients and healthy controls who were women from an urban Southwestern city; 46.6% were Caucasian and 43.2% reported low annual income (Steinberg et al., 2013)</p> <p>-BIS-BRIEF was tested with men recently convicted of domestic violence in two southeastern Louisiana parishes (ages 18-71, mean age = 34.56, SD = 10.87) (Steinberg et al., 2013)</p> <p>-BIS-BRIEF was tested on an inpatient sample of young adults and adolescents ages 13-22 (mean age = 16.72, SD = 2.36) (Steinberg et al., 2013)</p>	30- and 8-item versions	Steinberg, L., Sharp, C., Stanford, M. S., & Tharp, A. T. (2013). New tricks for an old measure: The development of the Barratt Impulsiveness Scale-Brief (BIS-Brief). <i>Psychological Assessment</i> , 25(1), 216–226. https://doi.org/10.1037/a0030550 .

Physical Well-Being Measures

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
General health (physical, mental, social subscales)	Medical Outcomes Survey (MOS) 36-Item Short-Form Health Survey (SF-36)	RAND	<p>-Relative validity of the physical functioning, role limitations due to physical health problems, and bodily pain subscales with the physical health component were 1.00, 0.79, and 0.77, respectively (Mchorney et al., 1993)</p> <p>-Relative validity of the mental health, role limitations due to mental health problems, and social functioning subscales with the mental health component were 1.00, 0.81, and 0.62, respectively (Mchorney et al., 1993)</p> <p>-Social functioning, vitality, and general health perceptions subscales showed moderate to strong associations with both mental and physical health components, though there was variation in relative validity (Mchorney et al., 1993)</p> <p>-Internal consistency reliability ranging from 0.83 to 0.93 for the eight subscales (Gandek et al., 2004)</p> <p>-Internal consistency</p>	<p>-Elderly and disabled; respondents had a mean age of 73.1 years, 56.9% of respondents were female, and 88.0% of respondents were white (Gandek et al., 2004)</p> <p>-London-based civil servants aged 35-55; 66.9% of respondents were men and 33.1% were female (Stansfeld et al., 1997)</p> <p>-Patients and physicians from health examinations who were English-speaking adults 18-years-old and older (Mchorney et al., 1993)</p> <p>-A large UK population sample between the ages of 18-64 years (Jenkinson et al., 1994)</p> <p>-A subset of the general US population participating in the National Survey of Functional Health Status (NSFHS) and a sample from the UK population (Ware et al., 1996)</p>	36-item version, 12-item version, and 8-item version	<p>Gandek, B., Sinclair, S. J., Kosinski, M., & Ware, J. E. (2004). Psychometric Evaluation of the SF-36® Health Survey in Medicare Managed Care. <i>Health Care Financing Review</i>, 25(4), 5–25.</p> <p>McCabe, C. J., Thomas, K. J., Brazier, J. E., & Coleman, P. (1996). Measuring the Mental Health Status of a Population: A Comparison of the GHQ-12 and the SF-36 (MHI-5). <i>The British Journal of Psychiatry</i>, 169(4), 517–521. https://doi.org/10.1192/bjp.169.4.516.</p> <p>Jenkinson, C., Wright, L., & Coulter, A. (1994). Criterion validity and reliability of the SF-36 in a population sample. <i>Quality of Life Research</i>, 3(1), 7–12. https://doi.org/10.1007/BF00647843.</p> <p>Mchorney, C. A., John, W. J., & Anastasiae, R. (1993). The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and Clinical Tests of Validity in Measuring Physical and Mental Health Constructs. <i>Medical Care</i>, 31(3), 247.</p> <p>Mchorney, C. A., & Ware, J. E. J. (1995). Construction and Validation of an Alternate Form General Mental Health Scale for the Medical Outcomes Study Short-Form 36-Item Health Survey. <i>Medical Care</i>, 33(1), 15.</p> <p>Stansfeld, S. A., Roberts, R., & Foot, S. P. (1997). Assessing the validity of the SF-36 General Health Survey. <i>Quality of Life Research</i>, 6(3), 0–0.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			reliability for physical functioning subscale of 0.93, role limitations due to physical problems subscale of 0.91, and the bodily pain subscale of 0.88 (Gandek et al., 2004) -Internal consistency reliability for mental health subscale of 0.83 and role limitations due to mental problems subscale of 0.88 (Gandek et al., 2004) -Evidence for discriminant validity of the general mental health and physical functioning subscales (Stansfeld et al., 1997) -Moderate construct and criterion validity for the social functioning subscale (Stansfeld et al., 1997) -Considerable overlap between the general mental health and social functioning scales (Stansfeld et al., 1997) -High internal consistency of the SF-36 for the whole sample (with Cronbach's alpha ranging from 0.76 to 0.90 across subscales) and also when broken down by subgroups (Cronbach's alpha was above 0.7 for all subgroups except for			https://doi.org/10.1023/A:1026406620756 . Ware, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. <i>Medical Care</i> , 34(3), 220–233. Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. <i>Medical Care</i> , 30(6), 473–483.

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>those on the social functioning subscale) (Jenkinson et al., 1994)</p> <ul style="list-style-type: none"> -SF-36 dimensions were strongly associated with patient reports of overall general health, and Kruskal-Wallis tests indicated clear linear trends for decreasing SF-36 scores across all subscales and worsening self-rated general health (Jenkinson et al., 1994) -Among the general US population, the 12-item Physical Component Summary had test-retest reliability of 0.890, while the 12-item Mental Component Summary had test-retest reliability of 0.760, respectively (Ware et al., 1996) -In 14 validity tests involving physical criteria of the 12-item Physical Component Summary, relative validity ranged from 0.43 to 0.93 (median = 0.67) in comparison with the best 36-item short-form scale (Ware et al., 1996) -In 6 validity tests involving mental criteria of the 12-item Mental Component Summary, 			

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>relative validity ranged from 0.60 to 1.07 (median = 0.97) in relation to the best 36-item short-form scale (Ware et al., 1996)</p> <p>-Average scores for the 12-item physical and mental summary measures, and those for most scales in the 8-scale profile based on the 12-item short-form, were very similar to those of the 36-item short-form, although standard errors were nearly always larger for the 12-item short-form (Ware et al., 1996)</p>			
General health, general mental health, and general physical health	PROMIS Global Health, PROMIS Global Mental Health, and PROMIS Global Physical Health	Patient-Reported Outcomes Measurement Systems (PROMIS®)	<p>-4-item global mental health scale (GMH-4) had internal consistency reliability coefficient of 0.86 (Hays et al., 2009)</p> <p>-4-item global physical health scale (GPH-4) had an internal consistency reliability coefficient of 0.81 (Hays et al., 2009)</p> <p>-2-item global mental health scale (GMH-2) had internal consistency reliability of 0.81 (Hays et al., 2017)</p> <p>-2-item global physical health scale (GPH-2) had internal consistency reliability of 0.73 (Hays et al., 2017)</p>	<p>-GPH-4 and GMH-4 were tested with a sample selected to be comparable to distributions of gender, age groups, race/ethnicity, and education based on the 2000 U.S. Census; average age was 53, 52% were female; 80% were non-Hispanic white, 9% were Latino, and 9% non-Hispanic black; 19% had only a high school degree or less (Hays et al., 2009)</p> <p>-11 item banks measuring components of physical, mental, social health, and global health and their short forms were tested on the U.S. general population and multiple disease populations; a scale-setting sub-</p>	10 for Global Health scale, 2 for Global Mental Health (GPH-2) scale, and 2 for Global Physical Health (GPH-2) scale	<p>Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., Amtmann, D., Bode, R., Buysse, D., Choi, S., Cook, K., DeVellis, R., DeWalt, D., Fries, J. F., Gershon, R., Hahn, E. A., Lai, J.-S., Pilkonis, P., Revicki, D., ..., Hays, R. (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. <i>Journal of Clinical Epidemiology</i>, 63(11), 1179–1194. https://doi.org/10.1016/j.jclinepi.2010.04.011.</p> <p>Hays, R. D., Bjorner, J. B., Revicki, D. A., Spritzer, K. L., & Cella, D. (2009). Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. <i>Quality of Life Research</i>, 18(7), 873–880.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>-Product-moment correlation between the GPH-2 and GPH-4 was 0.94 and between GMH-2 and GMH-4 was 0.97 (Hays et al., 2017)</p> <p>-The 2-item and 4-item versions of the GPH and GMH scales had similar correlations with PROMIS domain scores, the EQ-5D-3L and comorbidities, but the 4-item versions were more strongly correlated with these measures (Hays et al., 2017)</p> <p>-11 item banks measuring components of physical, mental, social health, and global health and their short forms demonstrated good reliability across the majority of score distributions (Cella et al., 2010)</p> <p>-Construct validity of the 11 item banks and their short forms was supported by moderated to strong correlations with other well-validated and widely accepted measures (Cella et al., 2010)</p> <p>-Reliabilities for physical and mental health summary scores were 0.98</p>	<p>sample was created that reflected demographics proportional to the 2000 U.S. Census (Cella et al., 2010)</p> <p>-The 2-item and 4-item versions of the GPH and GMH scales were tested on a sample that was 52% female; 82% White, 9% Black, 9% Hispanic; median age of 50 years (Hays et al., 2017)</p> <p>-Physical and mental health summary scores were tested with a large, demographically mixed population of people aged 18-93 (Hays et al., 2018)</p>		<p>https://doi.org/10.1007/s11136-009-9496-9.</p> <p>Hays, R. D., Schalet, B. D., Spritzer, K. L., & Cella, D. (2017). Two-item PROMIS® global physical and mental health scales. <i>Journal of Patient-Reported Outcomes</i>, 1(1), 2. https://doi.org/10.1186/s41687-017-0003-8</p> <p>Hays, R. D., Spritzer, K. L., Schalet, B. D., & Cella, D. (2018). PROMIS®-29 v2.0 profile physical and mental health summary scores. <i>Quality of Life Research</i>, 27(7), 1885–1891. https://doi.org/10.1007/s11136-018-1842-3.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			and 0.97, respectively (Hays et al., 2018)			
General well-being	Public Health Surveillance Well-being Scale (PHS-WB)	RTI International	-Strong internal consistency (alpha = 0.87), high correlation with scores for the entire item pool (r = 0.94) (Bann et al., 2012) -Strong construct validity demonstrated through high correlation with global and domain-specific measures of similar constructs (Bann et al., 2012)	-Community-dwelling adults in a nationally representative sample; 55% were women, around 50% were aged 35–54, about two-thirds had some college or more education, 68% were white, 68% married, 44% had household incomes of \$60,000 or more, 80% owned their own homes, and were geographically diverse (Bann et al., 2012)	10	Bann, C. M., Kobau, R., Lewis, M. A., Zack, M. M., Luncheon, C., & Thompson, W. W. (2012). Development and psychometric evaluation of the public health surveillance well-being scale. <i>Quality of Life Research</i> , 21(6), 1031–1043. https://doi.org/10.1007/s11136-011-0002-9 .

Family Well-Being Measures

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
Child well-being; Health-related quality of life (HRQOL)	KIDSCREE N-52	Ravens-Sieberer et al., 2005	Ravens-Sieberer et al., 2014 found that KIDSCREEN is a reliable, valid, sensitive, and conceptually/linguistically appropriate QoL measure in 38 countries/languages. The three KIDSCREEN versions showed excellent scale characteristics in terms of missing responses, floor and ceiling effects, and internal consistency. Cronbach's alphas are ranging from 0.77 to 0.89 for the dimensions of the 52-item version. Test-retest reliability was also generally satisfactory to excellent with ICCs ranging from 0.56 to 0.77 for the 52-item	Designed to be used in populations aged 8 to 18 years; tested in 13 countries (Austria (AT), Czech Republic (CZ), France (FR), Germany (DE), Greece (EL), Hungary (HU), Ireland (IE), the Netherlands (NL), Poland (PL), United Kingdom (UK), Spain (ES), Sweden (SE), and Switzerland	52 items	Ravens-Sieberer, U., Herdman, M., Devine, J. et al. (2014). The European KIDSCREEN approach to measure quality of life and well-being in children: development, current application, and future advances. <i>Qual Life Res</i> 23, 791–803. https://doi.org/10.1007/s11136-013-0428-3 . Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Duer, W., ... & KIDSCREEN

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>version. The three KIDSCREEN versions generally showed excellent scale characteristics in terms of missing responses, floor and ceiling effects, and internal consistency. All three KIDSCREEN instruments showed good results in terms of convergent, known groups', and criterion validity.</p> <p>Ravens-Sieberer et al. (2005) found that Cronbach α for the overall sample ranged from 0.77 to 0.89. Classical psychometric analysis confirmed the instrument's ability for sound measurement with sufficient psychometric properties. Instrument reliability was good. A comparison of the KIDSCREEN-52 dimensions and the KINDLR scales showed the highest correlations for all similar concepts/dimensions, showing good convergent validity. The KIDSCREEN-52 questionnaire has acceptable levels of reliability and validity.</p> <p>Ravens-Sieberer et al. (2008) found Cronbach's alpha values ranged from 0.77 to 0.89. Scaling success (Multitrait Analysis Program) was >97.8% for all dimensions and Rasch analysis item fit (INFITmsq) ranged from 0.80 to 1.27. The intraclass correlation coefficients ranged from 0.56 to 0.77.</p> <p>Essler, Christner, and Paulus (2021) found that Cronbach's Alphas for the</p>	<p>(CH)); applicable for healthy and chronically ill children and adolescents (Ravens-Sieberer et al., 2014).</p> <p>The mean age for overall child sample was 9.6 years and 14.3 years for adolescents. There were slightly more females than males in both. Age and gender were similar among participating countries. 46.2% Medium Family Affluence Scale (Ravens-Sieberer et al., 2005).</p> <p>Response rates by country ranged from 18.9% to 91.2%. Response rates were highest in the school-based surveys (69.0%–91.2%). Sample proportions by age and gender were similar to the reference Eurostat population in most countries, although</p>		<p>Group, E. (2005). KIDSCREEN-52 quality-of-life measure for children and adolescents. <i>Expert review of pharmacoeconomics & outcomes research</i>, 5(3), 353-364. https://doi.org/10.1586/14737167.5.3.353.</p> <p>Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Power, M., ... & Kidscreen Group. (2008). The KIDSCREEN-52 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. <i>Value in health</i>, 11(4), 645-658. https://doi.org/10.1111/j.1524-4733.2007.00291.x.</p> <p>Ramil, L., Berra, S., Erhart, M., et al. (2007). Survey methods and representativeness of European national surveys of the KIDSCREEN study. <i>BMC Public Health</i>, 7: 182.</p> <p>Europe, T. K. G. (2006). The KIDSCREEN Questionnaires. Quality of life questionnaires for children and adolescents. Lengerich: Pabst Science Publishers.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>dimensions range between 0.76 and 0.89. The KIDSCREEN evidences good convergent and discriminant validity.</p>	<p>boys and adolescents were slightly underrepresented (PFR <1). Parents in lower educational categories were less likely to participate (PFR <1 in 5 countries). Parents in higher educational categories were overrepresented when the school and household sampling strategies were used (PFR = 1.78–2.97) (Ravens-Sieberer et al., 2008; Ramil et al., 2007).</p> <p>Essler, Christner, and Paulus (2021) sample included most having a university degree (49%) or vocational training (24%) for T1 (first strict lockdown). 44% of parents were in a home office and 18% were at a job outside of the home for T1, while 31% of parents were in the home and 35% of parents had a job outside of the home</p>		<p>Essler, S., Christner, N. & Paulus, M. Longitudinal Relations Between Parental Strain, Parent–Child Relationship Quality, and Child Well-Being During the Unfolding COVID-19 Pandemic. <i>Child Psychiatry Hum Dev</i> 52, 995–1011 (2021).</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
				for T2 (loosened restrictions). This study doesn't provide many demographic variables.		
Child well-being; Health-related quality of life (HRQoL)	KIDSCREEN-10	Ravens-Sieberer et al., 2005; Ravens-Sieberer et al., 2010.	<p>Cronbach alpha was 0.82 (0.78), test-retest reliability was ICC = 0.70 (0.67) for the self- (proxy-)report version. KIDSCREEN-10 provides a valid measure of a general HRQoL factor in children and adolescents, but the instrument does not represent well most of the single dimensions of the original KIDSCREEN-52 (Ravens-Sieberer et al., 2010).</p> <p>Cronbach's alpha was 0.82 for the 10-item dimensions. Test-retest reliability was also generally satisfactory to excellent with an ICC of 0.70 for the 10-item version. Statistically significant correlations between the 10- and 27-item KIDSCREEN scores and the majority of the KIDSCREEN-52 scales indicated satisfactory criterion validity (Ravens-Sieberer et al., 2014).</p>	<p>13 countries participated in the KIDSCREEN-57 study. The target population for the KIDSCREEN study was children and adolescents aged 8–18 and their parents. The following 13 countries participated in the KIDSCREEN study: Austria (AT), Czech Republic (CZ), France (FR), Germany (DE), Greece (EL), Hungary (HU), Ireland (IE), Poland (PL), Spain (ES), Sweden (SE), Switzerland (CH), the Netherlands (NL), and the United Kingdom (UK). The target population for the KIDSCREEN study was children and adolescents aged 8–18 and their parents. Parent reports were not collected in Sweden and Ireland.</p>	10 items	<p>Ravens-Sieberer, U., Erhart, M., Rajmil, L. et al. (2010). Reliability, construct and criterion validity of the KIDSCREEN-10 score: a short measure for children and adolescents' well-being and health-related quality of life. <i>Qual Life Res</i> 19, 1487–1500. https://doi.org/10.1007/s11136-010-9706-5.</p> <p>Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Power, M., et al. (2008). The KIDSCREEN-52 Quality of Life measure for children and adolescents: Psychometric results from a cross-cultural survey in 13 European Countries. <i>Value in Health</i>, 11, 645–658.</p> <p>Ravens-Sieberer, U., Herdman, M., Devine, J. et al. (2014). The European KIDSCREEN approach to measure quality of life and well-being in children: development, current application, and future advances. <i>Qual Life Res</i> 23, 791–803.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
				<p>Data on the target child's and the parents' perceived health, together with data on parents' marital and educational status, and place of residence were collected from parents who refused to participate. These data were compared with similar data from participants. Statistically significant differences were observed between the two groups on some variables. However, the magnitudes of the differences were small—e.g., 90.6% of the responders and 86.2% of the refuses rated their child's health as good, very good, or excellent; 23.5% of the responders and 37.5% of the refuses had a low educational level. The most notable differences between countries occurred in socioeconomic status (FAS) with, for</p>		<p>https://doi.org/10.1007/s11136-013-0428-3.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
				example, 45.5% of the Czech Republic child sample reporting low FAS compared to only 7.5% of the French sample. Notable differences between countries in socioeconomic status occurred. Besides the lack of parents data from IE and SE, and a generally lower response rate the distribution of age, gender, and SES was quite similar in both the parents and the children samples. (Ravens-Sieberer et al., 2010).		
Child well-being; Health-related quality of life (HRQOL)	KIDSCREEN N-27	Ravens-Sieberer et al., 2005; Ravens-Sieberer et al., 2007.	KIDSCREEN-27 is a valid measure of HRQoL in children and adolescents. Correlation with corresponding scales of the KIDSCREEN-52 ranged from $r = 0.63$ to $r = 0.96$, and r^2 ranged from 0.39 to 0.92. Correlations between other HRQoL questionnaires and KIDSCREEN-27 dimensions were moderate to high for those assessing similar constructs ($r = 0.36$ to 0.63). Statistically significant and sizeable differences between physically and mentally healthy and ill children were found in all KIDSCREEN-27	13 countries participated in the KIDSCREEN-57 study. The following countries participated in the KIDSCREEN study: Austria (AT), Czech Republic (CZ), France (FR), Germany (DE), Greece (EL), Hungary (HU), Ireland (IE), Poland (PL), Spain (ES), Sweden (SE), Switzerland (CH), the	27 items	Ravens-Sieberer, U., Auquier, P., Erhart, M. et al. (2007). The KIDSCREEN-27 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. <i>Qual Life Res</i> 16, 1347–1356. https://doi.org/10.1007/s11136-007-9240-2 . Ravens-Sieberer, U., Herdman, M., Devine, J. et al. (2014). The European KIDSCREEN approach to measure quality of

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>dimensions together with strong associations with psychosomatic complaints ($r = -0.52$). Most of the KIDSCREEN-27 dimensions showed a gradient according to socio-economic status, age and gender. The KIDSCREEN-27 seems to be a valid measure of HRQoL in children and adolescents. Further research is needed to assess longitudinal validity and sensitivity to change (Ravens-Sieberer et al., 2007).</p> <p>Cronbach's alphas ranged from from 0.80 to 0.84 for the 27-item dimensions. Test-retest reliability was also generally satisfactory to excellent with ICCs ranging from 0.61 to 0.74 for the 27-item version. Statistically significant correlations between the 10- and 27-item KIDSCREEN scores and the majority of the KIDSCREEN-52 scales indicated satisfactory criterion validity, and KIDSCREEN-27 dimensions were found to explain 39–92 % of the variance in the corresponding KIDSCREEN-52 dimensions (Ravens-Sieberer et al., 2014)</p>	<p>Netherlands (NL), and the United Kingdom (UK). The target population for the KIDSCREEN study was children and adolescents aged 8–18 (Ravens-Sieberer et al., 2007).</p> <p>Response rates by country ranged from 18.9% to 91.2%. Response rates were highest in the school-based surveys (69.0%–91.2%). Sample proportions by age and gender were similar to the reference Eurostat population in most countries, although boys and adolescents were slightly underrepresented (PFR <1). Parents in lower educational categories were less likely to participate (PFR <1 in 5 countries). Parents in higher educational categories were overrepresented when the school and</p>		<p>life and well-being in children: development, current application, and future advances. <i>Qual Life Res</i> 23, 791–803. https://doi.org/10.1007/s11136-013-0428-3.</p> <p>Berra, S., Ravens-Sieberer, U., Erhart, M., Tebé, C., Bisegger, C., Duer, W., von Rueden, U., Herdman, M., Alonso, J., Rajmil, L., & European KIDSCREEN group [kidscreen] (2007). Methods and representativeness of a European survey in children and adolescents: the KIDSCREEN study. <i>BMC public health</i>, 7, 182. https://doi.org/10.1186/1471-2458-7-182.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
				household sampling strategies were used (PFR = 1.78–2.97) (Berra et al., 2007).		
Child well-being	KINDLR	Ravens-Sieberer and Bullinger, 1998	The results of psychometric testing in these populations showed that the German KINDL is a reliable, valid and practical instrument to assess the health-related quality of life of children which should be supplemented by disease-specific modules and needs to be further tested in clinical populations.	8–12 year and 13–17 years of age; five participating countries (i.e., Austria, Czech Republic, Germany, Greece and Spain (Ravens-Sieberer et al., 2005). The 45 children (23 girls and 22 boys) were 10–16 years old (mean = 13.2 years and SD = ± 2.7 years) and 56% were 13 years or younger. Approximately one-third of the children were at higher elementary level schooling, another third were at lower secondary level and the last third were at higher secondary level. Of the 45 children, two-thirds suffered from diabetes and one-third from asthma. 45 chronically ill children suffering from diabetes or asthma in comparison to 45 age- and	24 items which cover six dimensions of quality of life (e.g., physical, psychological, self-esteem, family, friends and functional aspects) for three age groups (4–7, 8–12 and 13–17 years)	Ravens-Sieberer U. and Bullinger M. (1998). Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric and content analytical results. <i>Qual Life Res.</i> 7, 399–407. Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Duer, W., ... & KIDSCREEN Group, E. (2005). KIDSCREEN-52 quality-of-life measure for children and adolescents. <i>Expert review of pharmacoeconomics & outcomes research</i> , 5(3), 353-364. https://doi.org/10.1586/14737167.5.3.353 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
				gender-matched healthy children. (Ravens-Sieberer and Bullinger, 1998)		
Child well-being; Health-related quality of life (HRQOL)	The Pediatric Quality of Life Inventory (PedsQL) 4.0	Varni, Seid, & Kurtin, 2001	Internal consistency reliability for the Total Scale Score (a 5 0.88 child, 0.90 parent report), Physical Health Summary Score (a 5 0.80 child, 0.88 parent), and Psychosocial Health Summary Score (a 5 0.83 child, 0.86 parent) were acceptable for group comparisons. Validity was demonstrated using the known-groups method, correlations with indicators of morbidity and illness burden, and factor analysis. The PedsQL distinguished between healthy children and pediatric patients with acute or chronic health conditions, was related to indicators of morbidity and illness burden, and displayed a factor-derived solution largely consistent with the a priori conceptually-derived scales. The results demonstrate the reliability and validity of the PedsQL 4.0 Generic Core Scales. The PedsQL 4.0 Generic Core Scales may be applicable in clinical trials, research, clinical practice, school health settings, and community populations (Varni, Seid, & Kurtin, 2001).	Targets children and adolescents ages 2 to 18. The average age of the 815 boys (48.6%) and 830 girls (49.5%; 32, 1.9% missing) was 9.3 years (SD 5 4.37) with a range of 2.0 to 18.8 years. For child self-report, the average age of the 643 boys (48.1%) and 664 girls (49.7%; 30, 2.2% missing) was 10.78 years (SD 5 3.61) with a range of 5.0 to 18.8. The sample was heterogeneous with respect to race/ethnicity, with 612 (36.5%) White non-Hispanic, 667 (39.8%) Hispanic, 117 (7.0%) Black non-Hispanic, 49 (2.9%) Asian/Pacific Islander, 18 (1.1%) American Indian or Alaskan Native, 99 (5.9%) Other, and 115 (6.9%) missing. With respect to	23 items	Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Power, M., ... & Kidscreen Group. (2008). The KIDSCREEN-52 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. <i>Value in health</i> , 11(4), 645-658. https://doi.org/10.1111/j.1524-4733.2007.00291.x . Varni, J. W., Seid, M., & Kurtin, P. S. (2001). PedsQL™ 4.0: Reliability and Validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Core Scales in Healthy and Patient Populations. <i>Medical Care</i> , 39(8), 800–812. doi:10.1097/00005650-200108000-00006

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
				<p>insurance type, 31% of the sample had commercial insurance, with 56% covered by Medicaid, 2.8% Other, 0.7% self-pay, and 9.5% missing. The sample was also diverse with respect to socioeconomic status, using mother's education as a proxy. In the sample, 7.1% of mothers had no more than an elementary school education, 15.8% had some secondary education but did not graduate from high school, 14.8% were high school graduates, 25.8% had some posthigh school education, 12.2% were college graduates, 4.4% had a graduate or professional degree, with 19.8% missing data. The sample included chronically ill children, acutely ill children, and healthy children (Varni, Seid, & Kurtin, 2001).</p>		

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
Child problem behaviors	Strengths and Difficulties Questionnaire (SDQ)	Goodman, 1997	<p>Reliabilities of the original SDQ subscales evidenced acceptable to good values and range between Cronbach's $\alpha = 0.58-0.76$ (Goodman, 1997).</p> <p>Scores derived from the SDQ and Rutter questionnaires were highly correlated; parent-teacher correlations for the two sets of measures were comparable or favoured the SDQ (Goodman, 1997).</p> <p>Confirmatory factor analysis supported a 5-factor measurement model. Internal reliability of subscales ranged from $\omega = 0.66$ (peer problems) to $\omega = 0.83$ (hyperactivity). Item-factor structures revealed measurement invariance over time. Strong positive correlations between ages 3 and 5 SDQ scores were not significantly different from correlations between age 5 and 7 scores. Conduct problems and hyperactivity subscales independently predicted developmental and clinical outcomes 2 years later (Croft et al., 2015).</p> <p>The 5-factor model previously found in studies using exploratory factor analysis was fit to the data for both parent and teacher questionnaires. Fit indices for both versions were marginally adequate. Model fit was comparable across gender and ethnic groups. Factor fit for the parent questionnaire was invariant across parent educational level. The examination of convergent and</p>	<p>Children aged 4-16 years attending one of two London child psychiatric clinics or the children's department of a London dental hospital (Goodman, 1997).</p> <p>784 children attending one of three school districts in southeastern Texas participating in a longitudinal study of school achievement. 42% White, 25% African American, 27% Hispanic, and 6% other. The participants were recruited across two sequential first-grade cohorts in Fall 2001 and Fall 2002 (Hill & Hughes, 2007).</p> <p>Representative sample of ten thousand 5-15-year-olds, all of whom had psychiatric assessments (Goodman, 2001).</p> <p>10,367 4- to 17-year-</p>	25 items	<p>Essler, S., Christner, N. & Paulus, M. (2021). Longitudinal Relations Between Parental Strain, Parent-Child Relationship Quality, and Child Well-Being During the Unfolding COVID-19 Pandemic. <i>Child Psychiatry Hum Dev</i> 52, 995-1011.</p> <p>Goodman R. (1997) The strengths and difficulties questionnaire: a research note. <i>J Child Psychol Psychiatry</i> 38:581-586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x.</p> <p>Croft, S., Stride, C., Maughan, B., & Rowe, R. (2015). Validity of the strengths and difficulties questionnaire in preschool-aged children. <i>Pediatrics</i>, 135(5), e1210-e1219.</p> <p>Hill, C. R., & Hughes, J. N. (2007). An examination of the convergent and discriminant validity of the Strengths and Difficulties Questionnaire. <i>School Psychology Quarterly</i>, 22(3), 380-406. https://doi.org/10.1037/1045-3830.22.3.380.</p> <p>Goodman, R. (2001). Psychometric properties of the strengths and difficulties</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>discriminant validity included peer evaluations of each SDQ construct. Thus, each of the five constructs was evaluated by three sources (parent, teacher, and child). On the basis of D. T. Campbell and D. W. Fiske's (1959) multitrait-multimethod approach as well as a confirmatory factor analysis using the correlated uniqueness model, the SDQ has good convergent validity but relatively poor discriminant validity (Hill & Hughes, 2007).</p> <p>The predicted five-factor structure (emotional, conduct, hyperactivity-inattention, peer, prosocial) was confirmed. Internalizing and externalizing scales were relatively "uncontaminated" by one another. Reliability was generally satisfactory, whether judged by internal consistency (mean Cronbach α: .73), cross-informant correlation (mean: 0.34), or retest stability after 4 to 6 months (mean: 0.62). SDQ scores above the 90th percentile predicted a substantially raised probability of independently diagnosed psychiatric disorders (Goodman, 2001).</p> <p>Results indicated good acceptability and internal consistency. Normative scoring bands were similar, though not identical, to the original British bands. Results of each scoring method had a strong association with service contact/use. This study supports the usefulness of the</p>	<p>olds in the 2001 National Health Interview Survey. Black and Hispanic populations are oversampled. a nationally representative sample of the civilian noninstitutionalized household population of the United States (Bourdon et al., 2005).</p> <p>Mothers of 5481 2 year olds (52 % male) from the Growing Up in New Zealand cohort, and investigated the psychometric properties of the preschool SDQ within this cohort (D'Souza et al., 2017).</p>		<p>questionnaire. <i>Journal of the American Academy of Child & Adolescent Psychiatry</i>, 40(11), 1337-1345. https://doi.org/10.1097/00004583-200111000-00015.</p> <p>Bourdon, K. H., Goodman, R., Rae, D. S., Simpson, G., & Koretz, D. S. (2005). The Strengths and Difficulties Questionnaire: US normative data and psychometric properties. <i>Journal of the American Academy of Child & Adolescent Psychiatry</i>, 44(6), 557-564. https://doi.org/10.1097/01.chi.000159157.57075.c8.</p> <p>D'Souza, S., Waldie, K. E., Peterson, E. R., Underwood, L., & Morton, S. M. (2017). Psychometric properties and normative data for the preschool strengths and difficulties questionnaire in two-year-old children. <i>Journal of abnormal child psychology</i>, 45(2), 345-357. https://doi.org/10.1007/s10802-016-0176-2.</p> <p>Van Roy, B., Veenstra, M., & Clench-Aas, J. (2008). Construct validity of the five-factor Strengths and Difficulties Questionnaire (SDQ) in pre-,</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>Strengths and Difficulties Questionnaire as an effective and efficient screener for child and adolescent mental health problems in the United States (Bourdon et al., 2005).</p> <p>Support for a modified five-factor model, in which the prosocial factor was extended into a positive construal factor by allowing cross-loadings of reverse-scored items. Full measurement invariance was found across gender and socioeconomic status, and partial invariance was found across mother's ethnicity. Cronbach's alpha was satisfactory for all subscales (α range = 0.71–0.84), except peer problems (α = 0.54). Overall, satisfactory psychometric properties were found for the preschool SDQ in 2 year olds (D'Souza et al., 2017).</p> <p>The results of Van Roy, Veenstra, & Clench-Aas (2008) indicated support for the proposed five-factor structure of the SDQ (Goodman, 2001) across a wide age range (10-19years), including older adolescents and different informants. All factor loadings were higher than .30, except for item 11 (good friend). The loadings differed across age groups and differed markedly between the parent/proxy and self-report measures. The MTMM showed that the source of ratings made a difference on the validity of all subscale ratings, with self-reports</p>			<p>early, and late adolescence. <i>Journal of Child Psychology and Psychiatry</i>, 49(12), 1304-1312. https://doi.org/10.1111/j.1469-7610.2008.01942.x.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			discriminating more on ratings of emotional and peer problems, and parents/proxies discriminating more on hyperactivity symptoms. A positive construal factor was identified but had a modest effect compared with the original five traits. Results suggested an unclear construct and meaning of the Prosocial behaviour subscale.			
Parental Self-efficacy	The Parenting Self-Efficacy Questionnaire—FSW	Kilem et al., 2014	The original FSW showed good psychometric properties (Cronbach's $\alpha = 0.78$). Good values of internal consistency were found for the total score of the FSW (Mothers: $\alpha = .78$; Fathers: $\alpha = .79$). Correlations with other questionnaires were mostly found to be as expected.	Sample of N = 271 mothers and N = 198 fathers of children who were 2.5 to 6.5 years old	9 items	Essler, S., Christner, N. & Paulus, M. Longitudinal Relations Between Parental Strain, Parent–Child Relationship Quality, and Child Well-Being During the Unfolding COVID-19 Pandemic. <i>Child Psychiatry Hum Dev</i> 52, 995–1011 (2021). Kliem S., Kessemeier Y., Heinrichs N., Döpfner M., Hahlweg K. (2014) The parenting self-efficacy questionnaire (FSW). <i>Diag</i> 60:35–45. https://doi.org/10.1026/0012-1924/a000107 .
Parental Self-efficacy	Karitane Parenting Confidence Scale (KPCS)	Crnec, Barnett, and Matthey, 2008	The KPCS showed acceptable internal consistency (Cronbach's alpha = .81), test-retest reliability ($r = .88$), and discriminant and convergent validity. A cut-off score was determined, and the scale's sensitivity and positive predictive power was 86% and 88%, respectively. The KPCS may prove a useful addition to tools for the assessment of parents	Demographic Data for the Total Sample (Crnec, Barnett, and Matthey, 2008): - Mean mother's age in years (SD) -32.0 (4.8) - Mean infant age in weeks (SD) - 24.7 (25.5)	15 items	Crnec, R., Barnett, B., & Matthey, S. (2008). Development of an instrument to assess perceived self-efficacy in the parents of infants. <i>Research in nursing & health</i> , 31(5), 442–453. https://doi.org/10.1002/nur.20271 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			and infants presenting to clinical services (Crnec, Barnett, and Matthey, 2008).	- Mean number of children - 1.5 (.8) - % of male index infants - 60 - % of married/defacto parents - 94 - % of mothers who have not completed a University or vocational course - 8 - Ethnic background (> 5% of total): Australian -63%		
Parent-child relationship stress	Parent-Child Relationship Inventory (PCRI)	Gerard, 1994	Results revealed acceptable internal consistency for most scales and moderate to high 1-year stability for all scales. Both parents' PCRI scores correlated with their views of family climate. Cross-informant concordance was pervasive and strong between mothers' PCRI scores and adolescents' perceptions of the parent-child relationship and family climate; however, convergence was not evident between fathers' and adolescents' reports. All scales revealed acceptable levels of internal consistency except the Communication scale (for mothers) and the Autonomy scale (for both parents). 1-year temporal stability coefficients were significant and reached moderate-to-strong magnitudes for all scales. (Coffman, Guerin, & Gottfried, 2006). For convergent validity, parent reports on the PCRI predicted their	The adolescents comprised 90% Euro-American, 52% male, and a wide range of middle SES families (M 45.6, SD 11.9 at 17 years). Children ranged in age from 3-15: 50.2% male, 49.8% female; 85.7% White, 6.9% Black, 1.9% Asian, 1.5% Hispanic, 1% Native American, and 3% Other Ethnicity. Mothers made up 55.2% of the sample, fathers (39.1%), stepparents (2.7%), and Other (3.1%). Age of parents 18-24 (2.3%), 25-34 (32.8%),	78 items	Gerard, A. B. (1994). Parent-Child Relationship Inventory (PCRI): Manual. Los Angeles: Western Psychological Services. Coffman, J. K., Guerin, D. W., & Gottfried, A. W. (2006). Reliability and validity of the Parent-Child Relationship Inventory (PCRI): Evidence from a longitudinal cross-informant investigation. <i>Psychological Assessment</i> , 18(2), 209-214. https://doi.org/10.1037/1040-3590.18.2.209 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>impressions of family climate assessed one year later. Fourth, when individual parent– child dyads were examined, cross-informant data demonstrated that PCRI scores related concurrently and predictively to adolescents’ perceptions of the quality of their relationship with their mothers, but not with their fathers.</p> <p>Gerard (1994) reported norm sample internal consistencies for the five scales investigated herein ranging from .76 to .88 with a median of .82</p>	<p>35-44 (55.4%), 45-54 (8%), and 55+ (1.4%). Respondents over age 54 were excluded from the normative sample. Education of parents: less than a high school degree (4.5%), high school graduates (4.5%), some college education (27.6%), and college degree or higher (35.4%) (Gerard, 1994).</p>		
Parent-child relationship stress	Parental Stress Index	Abidin, 1983	<p>In the original study by Abidin (1995) the norm group consisted of mothers aged 16-61 (M=30.9) and their children aged 1 month to 12 years (M=4.9, SD=3.1). The majority (41%) were recruited from well-child pediatric clinics in Central Virginia, USA. Normative data were also collected from 200 fathers aged 18-65 (M=32.1, SD=6.01).</p> <p>Reliability: The authors reported good internal consistency for the PPC, with Cronbach’s alpha values ranging from 0.70 to 0.95 (average .81). Test-retest reliability were assessed after 60 days, and the authors reported acceptable Pearson r between .63 and .96 (average .83). Test-retest in four other studies, using intervals of 3 weeks to 1 year have generally reports correlations above .60.</p>	<p>In the original study by Abidin (1995) the norm group consisted of mothers aged 16-61 (M=30.9) and their children aged 1 month to 12 years (M=4.9, SD=3.1). The majority (41%) were recruited from well-child pediatric clinics in Central Virginia, USA. Normative data were also collected from 200 fathers aged 18-65 (M=32.1, SD=6.01).</p> <p>Normative data indicated that the Montreal sample represented a</p>	101 items (optional 19-item Life Stress scale)	<p>Abidin, R. R. (1983). Parenting Stress Index: Manual, Administration Booklet,[and] Research Update. https://eric.ed.gov/?id=ED238896.</p> <p>Abidin, R.R. (1995). Parenting Stress Index, Third Edition: Professional Manual. Odessa, FL: Psychological Assessment Resources, Inc.</p> <p>Loyd, B. H., & Abidin, R. R. (1985). Revision of the parenting stress index. <i>Journal of Pediatric Psychology</i>, 10(2), 169-177. https://doi.org/10.1093/jpepsy/10.2.169.</p> <p>Haskett, M. E., Ahern, L. S.,</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>Validity: According to the NCTSN, the PSI has been used in over 50 studies, and has been validated with a broad variety of populations. Construct validity as been established in a wide variety of populations including a myriad of developmental issues, behavioural problems, disabilities, illnesses and ethnic backgrounds.</p> <p>Bigras, LaFreniere, & Dumas (1996) show that the two main factors of the PSI represent valid and relatively independent sources of information. Considering clinical and research applications, the PSI may be particularly useful in assessing combinations of stressors, which appear to be specifically harmful for the parent-child relationship. The results support the convergent and discriminant validity of the parent and child scales of the French version of the Parenting Stress Index. convergent validity was achieved with correlations between the PSI scales and three sets of criteria. Results indicate that mothers who report more stress also report</p> <p>1) more isolation, more symptoms of depression and marital maladjustment; 2) more negative attitudes and behavioral problems in their children and; 3) reported stress was also associated with independent assessment of more negative and controlling parenting practices. results also support</p>	<p>somewhat less educated, and poorer population of mothers than the normative sample of Abidin (1990). For example, forty percent of the mothers were separated or divorced (Bigras, LaFreniere, & Dumas, 1996)</p>		<p>Ward, C. S., & Allaire, J. C. (2006). Factor structure and validity of the parenting stress index-short form. <i>Journal of Clinical Child & Adolescent Psychology</i>, 35(2), 302-312. https://doi.org/10.1207/s15374424jccp3502_14.</p> <p>Bigras, M., LaFreniere, P. J., & Dumas, J. E. (1996). Discriminant validity of the parent and child scales of the parenting stress index. <i>Early Education and Development</i>, 7(2), 167-178. https://doi.org/10.1207/s15566935eed0702_5.</p> <p>Abidin, R. R. (1997). Parenting Stress Index: a measure of the parent-child system.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>the discriminant validity of the parent and child scales of the French version of the PSI. Both scales were shown to be related to measurement of similar constructs and accounted for most, if not all, of the variance in related construct</p> <p>Loyd & Abidin (1985) found that the alpha reliability coefficients are sufficiently large to indicate a high degree of internal consistency for these measures. These significant ($p < .001$) and relatively high reliability coefficients provide strong support for the stability of the scores across these time intervals, given a developing child and normal changes in the family system.</p>			
Parent-child relationship stress	Parental Stress Index - Short Form (PSI-SF)	Abidin, 1995	<p>Reitman and colleagues (2002) found that scores on the Brief Symptom Index, a derivative of the SCL-90-R, were related to PSI-SF scores;</p> <p>Results of our exploratory factor analysis indicated the presence of two reasonably discrete factors. This structure is consistent with early multidimensional models of parenting stress (Abidin, 1983; Burke & Abidin, 1980)</p> <p>parents who reported a high level of stress on the PD subscale also reported poor overall emotional health on the SCL-90-R. Further, parents who</p>	<p>Mostly White, primarily married mothers of young children (mean age under 4 years)</p> <p>Haskett et al., 2006 included 185 parents (80% mothers or mother figures, 20% fathers or father figures), 90 had documented histories of reported physical abuse and 95 had no known</p>	36 item; 3 scales with 12 items each.	<p>Haskett, M. E., Ahern, L. S., Ward, C. S., & Allaire, J. C. (2006). Factor structure and validity of the parenting stress index-short form. <i>Journal of Clinical Child & Adolescent Psychology</i>, 35(2), 302-312. https://doi.org/10.1207/s15374424jccp3502_14.</p> <p>Reitman, D., Currier, R. O., & Stickle, T. R. (2002). A critical evaluation of the Parenting Stress Index-Short Form (PSI-SF) in a head start population. <i>Journal of Clinical Child and Adolescent</i></p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>reported difficulty with their child’s self-regulatory capacity and demandingness on the CS scale also viewed their child’s adjustment as highly problematic on a separate measure of child adjustment. That finding is congruent with research showing that the original Difficult Child subscale of the PSI–SF was related to other parent-report measures of child adjustment (e.g., Reitman et al.,2002; Silovsky & Niec, 2002)</p> <p>The correlation between the scales (.58) suggested that the PD and CS scales might represent overlapping yet distinct constructs. In our direct tests of discriminant validity, results indicated reasonable support for the distinction between the two scales.</p> <p>With respect to predictive validity, PSI–SF scores were related to parents’ reports of their children’s disruptive behaviors in the home 1 year later.</p> <p>Reitman, Currier,& Stickle (2002) found that internal consistencies for the PSI–SF were very good to excellent. However, confirmatory factor analysis (CFA) indicated that a 3-factor model comprised of Parental Distress, Difficult Child, and Parent–Child Dysfunctional Interaction subscales was only marginally superior to a single-factor model. results appear to support the use of the PSI–SF</p>	<p>history of abuse; children (48% female) between the ages of 4 and 10 years; majority (68%) were African American, 34% were Caucasian, and 2% were Hispanic or biracial; . Forty-five percent of parents were married or living with a partner. Although 18% had not completed high school, 23% had a college degree. The full range of socioeconomic status (SES; Hollingshead, 1975) was represented, with 36% at the two highest levels and 43% at the two lowest levels. Mean parent age was 34.4 years</p> <p>Reitman, Currier,& Stickle (2002) examines psychometric characteristics of the 36-item Parenting Stress Index–Short</p>		<p><i>Psychology</i>, 31(3), 384-392. https://www.tandfonline.com/doi/abs/10.1207/S15374424JCCP3103_10.</p> <p>Whiteside-Mansell, L., Ayoub, C., McKelvey, L., Faldowski, R. A., Hart, A., & Shears, J. (2007). Parenting stress of low-income parents of toddlers and preschoolers: Psychometric properties of a short form of the Parenting Stress Index. <i>Parenting: Science and Practice</i>, 7(1), 26-56. https://doi.org/10.1080/15295190709336775.</p> <p>Ahern, L. S. (2005). Psychometric properties of the parenting stress index-short form. https://repository.lib.ncsu.edu/handle/1840.16/2765.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			<p>with lower socioeconomic, primarily African American mothers. The regression analyses supported the construct validity of the PSI-SF.</p> <p>Whiteside-Mansell et al. (2007) found that CFA indicated that the 5-factor scales fit the data better than the 2-factor scales. Both 2- and 5-factor scales had high internal consistency, and the pattern of relations between the new scales and validity constructs support their usefulness. This concluded that the PSI-SF can be used in clinical applications. Detailed examinations of internal consistency and validity coefficients for PD and PCDI supported the two scales in this sample. These subscales had good internal consistency and to correlate with validity constructs in patterns that suggest they provided a more focused assessment of the broader constructs. Each of the proposed scales contributed to the prediction of validity constructs after controlling for the four other dimensions of stress.</p> <p>Ahern, L. S. (2005) found that the Difficult Child and Parent Distress subscales, as well as Total PSI-SF, were internally consistent. Confirmatory factor analysis did not support a three-factor model. Results were mixed in terms of support for convergent and discriminant validity. The PSI-SF Total and subscales were related to measures of parent</p>	<p>Form (PSI-SF) in a low-income, predominantly minority population.</p> <p>Whiteside-Mansell et al. (2007) examined psychometric properties of 2 scales of the Parenting Stress Index – Short Form (PSI – SF) in a sample of preschool children (Early Head Start parents) from low-income families. Most families (98%) had biological mothers. Forty-two percent of the mothers were European American, 40% African American, 13% Latin American, and 5% other minority. Most mothers (60%) had completed 12 years of education, 39% were younger than 20 years of age (M = 22.1, SD = 5.5), and 34% received public assistance. At intake into the EHSRE study, 26% of</p>		

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			psychopathology and perceptions of child adjustment, but not to observed parent and child behavior.	mothers were married. Most target children (64%) were firstborn (the mean number of children in family at study intake was 1.83, SD = 1.2), and 50% of the children were boys.		
Parental Stress	Parental Stress Scale	Berry & Jones, 1995)	Highly reliable, both internally and over time, and related to the general measure of stress. Berry & Jones (1995) reported high correlations between the PSS and the Perceived Stress Scale in a US sample. Results are consistent across different parental characteristics, inferring the stability of scale characteristics. Validity was supported by predicted correlations with measures of relevant emotions and role satisfaction and significant discrimination between mothers of children in treatment for emotional/behavioral problems and development disability disabilities vs. those not receiving treatment. Factor analysis showed high internal reliability	First Group: - mean age was 34.4 - median number of children was 2 and mean age was 6.7 - median education level of 15.5 years and 91% white family incomes in excess of \$40,00 reported for half of participants Second Group: - mean age was 36.8 - median number of children was 2 and mean age was 7.8 - well educated with 50% having college degrees and 95% white family incomes in excess of \$40,00 reported for half of participants	18-item self-report measure	Berry, J. O., & Jones, W. H. (1995). The parental stress scale: Initial psychometric evidence. <i>Journal of social and personal relationships</i> , 12(3), 463-472. https://doi.org/10.1177/0265407595123009 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
Maternal Depression	Center for Epidemiological Studies Depression Scale (CESD)	Radloff (1977)	<p>Cronbach's αs were computed as .89 (15 months) and .88 (37 months) (Whiteside-Mansell et al., 2007).</p> <p>It was found to have very high internal consistency and adequate test-retest repeatability. Validity was established by patterns of correlations with other self-report measures, by correlations with clinical ratings of depression, and by relationships with other variables which support its construct validity. Reliability, validity, and factor structure were similar across a wide variety of demographic characteristics in the general population samples tested. The scale should be a useful tool for epidemiologic studies of depression (Radloff, 1977).</p> <p>Cronbach's alpha of .90 was observed. The 4-factor model had the best model fit. Conclusions: High internal consistency was demonstrated alongside a replication of the original 4-factor structure (Cosco et al., 2017).</p>	<p>The new scale was tested in household interview surveys and in psychiatric settings. The author states that it was tested on the general population.</p> <p>Analyses were conducted on each of three age groups (under 25, 25-64, over 64), the two sexes, two races (Black and White), three levels of education (less than high school, high school, greater than high school), and the two "need help" groups ("need help," "not need help") With few exceptions, the results for the total population were confirmed in all subgroups. (tables available on request). In all subgroups, coefficient alpha was .80 or above (Radloff, 1977).</p> <p>The MacArthur Foundation's Midlife in the United States</p>	20 item; short form is 12 items	<p>Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. <i>Applied Psychological Measurement</i>, 1(3), 385–401. https://doi.org/10.1177/014662167700100306.</p> <p>Cosco, T. D., Prina, M., Stubbs, B., & Wu, Y. T. (2017). Reliability and Validity of the Center for Epidemiologic Studies Depression Scale in a Population-Based Cohort of Middle-Aged U.S. Adults. <i>Journal of nursing measurement</i>, 25(3), 476–485. https://doi.org/10.1891/1061-3749.25.3.476.</p> <p>Whiteside-Mansell, L., Ayoub, C., McKelvey, L., Faldowski, R. A., Hart, A., & Shears, J. (2007). Parenting stress of low-income parents of toddlers and preschoolers: Psychometric properties of a short form of the Parenting Stress Index. <i>Parenting: Science and Practice</i>, 7(1), 26-56. https://doi.org/10.1080/15295190709336775.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
				Study (MIDUS), is a nationally representative cohort study of 7,108 noninstitutionalized, English-speaking adults aged 24–74 years through a random digit dialing. Average age is 54.5. (Cosco et al., 2017).		
Family Conflict	Family Environment Scale	Moos & Moos, 2002	<p>Items were recoded and averaged so that 4 indicated high levels of conflict. Cronbach’s α was computed as .65 (15 months), .67 (25 months) and .68 (37 months) (Whiteside-Mansell et al., 2007).</p> <p>The FES subscales generally show adequate internal consistency reliability and stability over time when applied in samples that are diverse; the items also have good content and face validity. An extensive body of research supports the construct, concurrent, and predictive validity of the FES. More generally, reliability and validity are a joint function of scale items and response formats and of the characteristics and diversity of specific samples (Moos, 1990).</p>	Moos (1990) stated these samples are not as diverse as the original normative sample; specifically, the new samples include fewer individuals of low socioeconomic status, fewer minority individuals, fewer adolescents, and fewer currently distressed families. Given the reduced variability in these new samples, the FES subscale internal consistencies are quite acceptable.	90 items	<p>Moos, R. H., & Moos, B. S. (2002). Family environment scale. Redwood City, CA: Mind Garden.</p> <p>Moos R. H. (1990). Conceptual and empirical approaches to developing family-based assessment procedures: resolving the case of the Family Environment Scale. <i>Family process</i>, 29(2), 199–211. https://doi.org/10.1111/j.1545-5300.1990.00199.x.</p> <p>Whiteside-Mansell, L., Ayoub, C., McKelvey, L., Faldowski, R. A., Hart, A., & Shears, J. (2007). Parenting stress of low-income parents of toddlers and preschoolers: Psychometric properties of a short form of the Parenting Stress Index. <i>Parenting: Science and Practice</i>, 7(1), 26-56.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
						https://doi.org/10.1080/15295190709336775 .
Adult Psychological Symptoms	Symptom Checklist–90–Revised (SCL–90–R)	Derogatis, 1994	<p>Convergent-discriminant validity studies indicate SCL–90–R scales have moderate to high correlations with corresponding Minnesota Multiphasic Personality Inventory scales; for example, correlations between the analogous depression scales range from .68 to .75. Internal consistency of scales for this sample ($\alpha = .74-.88$) was consistent with the range reported by the author (Haskett et al., 2006).</p> <p>The study presents several significant findings related to the appropriateness of the SCL-90-R for use with African American women. Most notably, the current study is one of the few to date to examine the factor structure of a well-known self-report instrument in an exclusively African American sample. Based on the results of the current study, results indicate that the SCL-90-R is an adequate measure for assessing psychological distress, specifically in communitydwelling African American women (Chapman, Petrie, & Vines, 2012).</p>	<p>Ninety-one African American women completed the SCL-90-R as part of a larger investigation of anxiety and related disorders in African American parent-child dyads. Participants were 91 community dwelling African American females. Participants ranged in age from 23 to 55, with a mean age of 37 (S.D. $\sqrt{7.28}$). The majority of participants in the current study were single (n$\sqrt{470}$), although 23% were married at the time the study took place. The number of children that participants in the current study had ranged from one to seven, with a mean of approximately 2.5 children (S.D. $\sqrt{41.27}$). Ninety-three percent of the participants were high school</p>	90 items	<p>Haskett, M. E., Ahern, L. S., Ward, C. S., & Allaire, J. C. (2006). Factor structure and validity of the parenting stress index-short form. <i>Journal of Clinical Child & Adolescent Psychology</i>, 35(2), 302-312. https://doi.org/10.1207/s15374424jccp3502_14.</p> <p>Derogatis L.R. (1994). SCL-90-R: Administration, scoring and procedures manual (3rd ed.). Minneapolis, MN: NCS Pearson</p> <p>Kevin Chapman, L., Petrie, J., & Vines, L. (2012). Factor structure analysis of the SCL-90-R in a community-based sample of African American women. <i>Psychiatry research</i>, 199(2), 140–144. https://doi.org/10.1016/j.psychres.2012.03.028.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
				graduates, although 62% of the sample earned less than \$30,000 annually. (Chapman, Petrie, & Vines, 2012).		
Parent perceptions of child adjustment	Eyberg Child Behavior Inventory	Eyberg & Pincus, 1999	<p>The ECBI is a widely used measure of behavior problems, is sensitive to effects of parent training interventions, and has strong psychometric properties. Twelve-week test–retest stability is high (.75), and correspondence between mothers and fathers on the Intensity score is good (.69). Scores on the ECBI are correlated with observational measures of child affect and behavior during mother–child interactions and with parent-reported temperament (for a summary of psychometric properties, see Eyberg & Pincus, 1999). For this sample, internal consistency of the Intensity score was .92 (Haskett et al., 2006).</p> <p>ECBI scales had high internal consistency reliabilities and good convergence with the Child Behavior Checklist/1-5. Some racial/ethnic and income effects were found. There were no mean differences by ECBI language version or by child gender. Using confirmatory factor analysis, a single-factor invariant model of the ECBI Intensity Scale provided a good fit with</p>	The ECBI was originally standardized on parents of preadolescent children in 1980. It was standardized on parents of adolescents in 1983. Primarily, these children were from lower- and lower-middle SES White families recruited from a pediatric outpatient clinic located in a large urban medical school in the Northwest U.S. (Eyberg & Robinson, 1983; Robinson, Eyberg, & Ross, 1980).	36 items	<p>Haskett, M. E., Ahern, L. S., Ward, C. S., & Allaire, J. C. (2006). Factor structure and validity of the parenting stress index-short form. <i>Journal of Clinical Child & Adolescent Psychology</i>, 35(2), 302-312. https://doi.org/10.1207/s15374424jccp3502_14.</p> <p>Gross, D., Fogg, L., Young, M., Ridge, A., Cowell, J., Sivan, A., & Richardson, R. (2007). Reliability and validity of the Eyberg Child Behavior Inventory with African-American and Latino parents of young children. <i>Research in nursing & health</i>, 30(2), 213–223. https://doi.org/10.1002/nur.20181.</p> <p>Eyberg, S., & Pincus, D. (1999). <i>Eyberg Child Behavior Inventory & Sutter-Eyberg Student Behavior Inventory-Revised: Professional Manual</i>. Odessa, FL: Psychological Assessment Resources.</p>

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			the data across racial/ethnic and income groups (Gross et al., 2007).			Eyberg, S.M., & Robinson, E.A. (1983). Conduct problem behavior: Standardization of a behavioral rating scale with adolescents. <i>Journal of Clinical Child Psychology</i> , 12(3), 347-354.
Emotional maladjustment	Mental Health Inventory	Veit & Ware, 1983	Cronbach's α is .71 for emotional maladjustment; Depression Cronbach's α is .91; Anxiety Cronbach's α is .83	Representative of Swedish children in urban areas	38 items: 4 items for depressive mood; 8 items referred to anxiety	Overbeek, G., Stattin, H., Vermulst, A., Ha, T., & Engels, R. C. M. E. (2007). Parent-child relationships, partner relationships, and emotional adjustment: A birth-to-maturity prospective study. <i>Developmental Psychology</i> , 43(2), 429-437.
General satisfaction with life	Life Satisfaction Index	Neugarten, Havighurst, & Tobin, 1961	Dissatisfaction with life Cronbach's α is .73	Representative of Swedish children in urban areas	8 items	Overbeek, G., Stattin, H., Vermulst, A., Ha, T., & Engels, R. C. M. E. (2007). Parent-child relationships, partner relationships, and emotional adjustment: A birth-to-maturity prospective study. <i>Developmental Psychology</i> , 43(2), 429-437.
Coparenting	N/A	(Feinberg & Kan, 2008)	Scale development was based on a combination of theoretically based conceptualization, exploratory factor analysis, and reliability analysis. All items utilized 7-point Likert response scales. Cronbach's alphas ranged from .80 to .83 for mothers and from .66 to .80 for fathers.	Heterosexual adult couples who, at the time of recruitment, were expecting their first child and were living together. Resided in rural areas. Eighty-two percent of couples were married, and the majority of participants (91% of mothers and 90% of fathers) were non-	3 scales: Coparental Support, Parenting-Based Closeness, and Coparental Undermining. Each scale	Feinberg, M. E., & Kan, M. L. (2008). Establishing family foundations: Intervention effects on coparenting, parent/infant well-being, and parent-child relations. <i>Journal of Family Psychology</i> , 22(2), 253-263. https://doi.org/10.1037/0893-3200.22.2.253 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
				Hispanic White. Median annual family income was \$65,000 (<i>SD</i> = \$34,372), with a range of \$2,500 to \$162,500.	comprised 5 items.	

Cross-Domain Measures

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
Quality of life	Quality of Life Scale (QOLS)	Carol Burckhardt	-Strong internal consistency (Cronbach's alpha ranged from 0.82-0.92) and high test-retest reliability over 3 weeks (<i>r</i> ranged from 0.78 to 0.84) (Burckhardt & Anderson, 2003)	-Healthy adults and patients with rheumatic diseases, fibromyalgia, chronic obstructive pulmonary disease, gastrointestinal disorders, cardiac disease, spinal cord injury, psoriasis, urinary stress incontinence, posttraumatic stress disorder, and diabetes (Burckhardt & Anderson, 2003)	15- and 16-item versions	Burckhardt, C. S., & Anderson, K. L. (2003). The Quality of Life Scale (QOLS): Reliability, Validity, and Utilization. <i>Health and Quality of Life Outcomes</i> , 1(1), 60. https://doi.org/10.1186/1477-7525-1-60 .
Experiences of everyday discrimination	Everyday Discrimination Scale (Short Version)	Dr. David R. Williams (scale available at https://scholar.harvard.edu/davidwilliams/node/32397)	Cronbach's alpha of .77	Data from the Chicago Community Adult Health Study (CCAHS), a household probability sample of 3,105 adults aged 18 and over living in Chicago, Illinois stratified into 343 neighborhood clusters.	6 items	Sternthal, M. J., Slopen, N., & Williams, D. R. (2011). Racial disparities in health: how much does stress really matter? 1. <i>Du Bois review: social science research on race</i> , 8(1), 95-113.
Experiences of unfair	Chronic Work	Dr. David R. Williams	Cronbach's alpha of .73	Data from the Chicago Community Adult Health	3 items	Sternthal, M. J., Slopen, N., & Williams, D. R. (2011). Racial

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
treatment at work	Discrimination and Harassment (Abbreviated)	(scale available at https://scholar.harvard.edu/davidwilliams/node/32397)		Study (CCAHS), a household probability sample of 3,105 adults aged 18 and over living in Chicago, Illinois stratified into 343 neighborhood clusters.		disparities in health: how much does stress really matter? 1. <i>Du Bois review: social science research on race</i> , 8(1), 95-113.
Heightened vigilance	Heightened Vigilance Scale (Abbreviated)	Dr. David R. Williams (scale available at https://scholar.harvard.edu/davidwilliams/node/32397)	Cronbach's alpha of .72	Data from the Chicago Community Adult Health Study (CCAHS), a household probability sample of 3,105 adults aged 18 and over living in Chicago, Illinois stratified into 343 neighborhood clusters.	4 items	Sternthal, M. J., Slopen, N., & Williams, D. R. (2011). Racial disparities in health: how much does stress really matter? 1. <i>Du Bois review: social science research on race</i> , 8(1), 95-113.
Experiences of discrimination	Major Experiences of Discrimination (Abbreviated)	Dr. David R. Williams	Cronbach's alpha of .46	Data from the Chicago Community Adult Health Study (CCAHS), a household probability sample of 3,105 adults aged 18 and over living in Chicago, Illinois stratified into 343 neighborhood clusters.	4 item summary measure	Sternthal, M. J., Slopen, N., & Williams, D. R. (2011). Racial disparities in health: how much does stress really matter? 1. <i>Du Bois review: social science research on race</i> , 8(1), 95-113.
Experiences of discrimination	Major Experiences of Discrimination	Dr. David R. Williams (scale available at https://scholar.harvard.edu/davidwilliams/node/32397)	- Developed based on qualitative studies of discrimination. Not intended for use as scale; correlations exist between items but are modest and principal axis factor analysis did not find evidence of underlying structure. - Cronbach's alpha of .71 for African American working class sample and .52 for Latino	Used widely on populations of all races in the U.S. and elsewhere.	9 items (with 3 follow up questions for each "yes" response)	Kessler, R. C., Mickelson, K. D., & Williams, D. R. (1999). The prevalence, distribution, and mental health correlates of perceived discrimination in the United States. <i>Journal of health and social behavior</i> , 208-230. https://doi.org/10.2307/2676349 .

Construct	Measure Name	Publisher	Psychometrics	Populations measure has been normed on/used with	Number of items	Citation
			working class sample for situations with racial cause of discrimination (Krieger et al. 2005)			Krieger, N., Smith, K., Naishadham, D., Hartman, C., & Barbeau, E. M. (2005). Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health. <i>Social science & medicine</i> , 61(7), 1576-1596. https://doi.org/10.1016/j.socscimed.2005.03.006 .