Journal of Family and Economic Issues
https://doi.org/10.1007/s10834-026-10102-x

INTRODUCTION

Introduction to the Special Issue on Measurement

Dee Warmath[1] · Ashley B. LeBaron-Black[2] · Melissa Wilmarth[3]

Received: 29 April 2026 / Accepted: 23 May 2026

Abstract
Quality measurement is a requirement for good science and an important consideration in whether the science should be relied on for policy, program development, innovation, and future research. Despite the fundamental importance of measurement, it receives relatively little attention in the field of family finance. This special issue seeks to stimulate conversation on the subject and communicate JFEI's commitment to supporting measurement efforts.

Measurement Matters

Measurement is the act of quantifying or classifying a physical, social, or psychological phenomenon (Finkelstein & Learning, 1984). The quality of a measure is determined by its validity (i.e., accuracy) and reliability (i.e., consistency). The development of quality measures requires an understanding of the construct being operationalized to ensure that the measure accurately reflects the construct. Quality measure development involves iteration and continuous study across arguments, research settings, study populations, and time rather than a one-and-done activity. The best available measures, though flawed reflections of a construct, can be used effectively in good science if the flaws are understood and considered intentionally.

The challenges of measurement are magnified with latent constructs, meaning abstract constructs known theoretically but difficult to measure directly. For example, "happiness" is generally understood as positive inner feelings. An accurate definition is an important first step. But how can happiness then be measured accurately and consistently across cultures, people, and time independent of other factors (e.g., luck, events) that influence the level of happiness an individual experiences? The answer, of course, is with great care and effort, and not without some error (Oishi et al., 2013). The study of family finance uses many such sociopsychological measures to assess constructs such as financial wellbeing, financial literacy, approaches to money management, financial values, financial satisfaction, relationship quality, depression, and countless others.

Psychometrics provides the processes and procedures to measure such latent constructs (DeVellis & Thorpe, 2022). Working from a theoretical construct, psychometric methods guide the development of candidate items, exploratory analysis to define the scale, and confirmatory analysis to assess the validity and reliability of the scale. However, quality measures result not merely from following established steps but from the care taken in those steps. For example, cognitive interviewing can be a helpful step in scale development to ensure that potential items are clearly understood and answered as intended by members of the target population (Boateng et al., 2018). Scale validity further increases when researchers ensure that the interviewees are sufficiently diverse to truly represent the target population and are not inadvertently a sociodemographic subset (e.g., college students, White individuals, etc.) (Beatty & Willis, 2007). Sociopsychological constructs deal with attributes, properties, and phenomena that are difficult to measure, socially constructed, and vary across contexts and time (Duncan, 1984). Measures developed in one setting and time with a given population may or may not be appropriate for use in other settings or times with different populations. Psychometrics can support the researcher in making that determination. Financial wellbeing, as defined and

Dee Warmath
warmath@uga.edu

Ashley B. LeBaron-Black
lebaronashley@gmail.com

Melissa Wilmarth
mwilmarth@ches.ua.edu

1 University of Georgia, Athens, Georgia

2 Brigham Young University, Provo, UT, USA

3 University of Alabama, Tuscaloosa, AL, USA

Published online: 30 May 2026

measured by the Consumer Financial Protection Bureau, provides an example using working age and retirement age adults (CFPB, 2018). In their research, they discovered that retirement age adults weighted elements of the financial wellbeing construct differently than working age adults. Further, when the measure is used in other cultures, with a different subset of a U.S. population, or perhaps even post-COVID, it is worth assessing whether the measure is valid and reliable in those circumstances. Measurement should be an iterative process rather than a fixed or one-time process.

Quantitative empirical study requires measurement as a fundamental activity. Quality measurement is a requirement for good science and an important consideration in whether the science should be relied on for policy, program development, innovation, and future research.

The measures selected determine what is truly being studied regardless of the conceptual argument being made. For example, a paper that derives hypotheses regarding financial wellbeing but selects measures of financial distress is not a study of financial wellbeing–it's a study of financial distress. Additionally, if a paper uses a scale named "financial wellbeing" but the scale does not accurately or fully capture the phenomenon, then this paper is also not truly a study of financial wellbeing–it's a study of whatever the scale actually measures. Measures also determine whether planned comparisons are valid. Scale items may reflect the latent construct very well in one setting but not in another. An example here is the Snyder Hope Scale (Snyder et al., 1991) item: *I can think of many ways to get out of a jam.* When the scale was first developed in the United States, "get out of a jam" had clear meaning to the population being studied. When researchers tried to use this scale in Australia, this saying was not well understood so the scale did not function as well. Researchers have also found that other (primarily younger) subsets of the U.S. population do not naturally understand this saying either. When inaccurate and/or inconsistent measures are used, study findings may not actually support the researchers' claims. "An unfortunate but distressingly common occurrence is the conclusion that some *construct* is unimportant or that some *theory* is inconsistent based on the performance of a *measure* that may not reflect the variable assumed by the investigator" (DeVellis & Thorpe, 2022, p. 15).

Measures should be developed carefully and selected intentionally, and their ability to reflect the constructs in a given study accurately should be confirmed. It is irresponsible to select measures based solely on their convenience or frequency of use. Because no measure of latent constructs is perfect, knowing how a measure is flawed is crucial to interpreting findings appropriately (DeVellis & Thorpe, 2022). And yet, the realities of research unrelentingly strike. Deadlines loom, tenure requirements threaten, and energies are pulled in many directions; too often, measurement is low on the list of a researcher's priorities. Frankly, we can hardly expect more from researchers when measurement is low on the list of priorities for funders, reviewers, and editors, and when measurement is not sufficiently covered in graduate training. If researchers can expect little attention to or pushback about their measurement choices, there is little incentive to spend limited time and resources making careful choices. In fact, fields discourage careful measurement when few grants focus exclusively on measurement and/or it is difficult to get a paper accepted to a journal or conference if it examines measurement-related topics. Despite the fundamental importance of measurement, it receives relatively little attention in the field of family finance.

The Special Issue

This special issue is one step towards greater prioritization of measurement. We seek to stimulate conversation on the subject and communicate JFEI's commitment to supporting measurement efforts. We hope researchers, reviewers, editors, graduate advisors and instructors, funders, policymakers, educators, and practitioners (whose work is informed by research) will join us in giving increased attention and care to measurement issues. We are grateful to the National Endowment for Financial Education (NEFE) for partnering with us in this endeavor and for making measurement a funding priority over the past several years. In addition to the 10 research articles in the special issue, we also include an important commentary, "Adding Equity Considerations to Construct, Scale, and Model Development in the Field of Financial Education," written by Dr. Joshua Caraballo, Managing Director of Research & Evaluation for NEFE. The commentary highlights the need for greater attention to equity in measurement development, testing, and usage to accurately and responsibly study diverse, complex populations. Dr. Caraballo notes that "measurement is not inherently neutral" and discusses the far-reaching, real-world impact of measurement. He presents practical steps for researchers and an actionable strategy for the field for more equitable, culturally-responsive, and rigorous measurement.

Four of the 10 research articles included in the special issue came from a panel convened by NEFE to explore measurement topics in the field of consumer finance. One of these four articles is a consensus statement on measurement from the NEFE panel members. The other six articles included in the issue were submitted to the journal and met the intention of the special issue. All 10 research articles are explicitly concerned with the development, validation, refinement, or evaluation of measures in the family (and consumer) finance domain. All 10 articles went through the typical peer review process. One of the most interesting aspects of this process was seeing the conversations between authors and reviewers that unfolded across initial submissions and revisions. As those conversations unfolded, it became clear that there is insufficient opportunity for and encouragement of such conversations in the regular course of academic research, at least in the field of family finance.

As a set, the papers in this issue address several questions, including the following four.

1. (1) How should financial constructs be conceptualized and operationalized for specific populations?
2. (2) Do widely used financial measures actually function as intended across different groups?
3. (3) When are comparisons across groups substantively meaningful rather than artifacts of measurement bias?
4. (4) How do practical constraints (e.g., survey length, language proficiency) shape measurement quality?

Below, we provide a brief overview of each of the 10 papers.

Improving Measurement in the Study of Consumer Finance: Consensus from the NEFE Scale Validation Project

This paper reports the results of a modified Delphi consensus process aimed at improving measurement practices in consumer and family finance research. An expert panel identified 19 practical, field-wide recommendations for improved measurement including scale development, selection, and use organized into five conceptual domains. Rather than validating a specific measure or prescribing action, the paper positions itself as a methodological intervention designed to stimulate conversation and improved practices regarding measurement across the discipline.

Measuring Financial Constructs: The Predictive Power of Single Items and Scales

This paper investigates whether single-item measures drawn from multi-item financial scales can predict financial behaviors nearly as well as full scales. Using NEFE survey data, the authors compare the predictive validity of complete scales versus individual items selected based on factor loadings and Item Response Theory parameters, using logistic regression models. The findings confirm that full scales generally explain more variance, but also reveal that some single items—regardless of whether they have high or low factor loadings—can perform similarly in predicting outcomes. The paper contributes to methodological discourse by challenging assumptions about item selection, highlighting trade-offs between survey length and precision, and providing empirical evidence relevant to survey design under space constraints.

The Concept of Money Management Across Poverty Levels

This paper examines whether the Brief Money Management Scale (BMMS) measures the same underlying construct across individuals living at different poverty levels in the United States. Using data from the NEFE survey, the authors conduct confirmatory factor analysis, reliability testing, and measurement invariance analyses across three poverty levels. The results support factorial validity, reliability, and measurement invariance of the BMMS, indicating that the scale can be meaningfully used to compare money management behaviors across poverty groups. The study's contribution lies in demonstrating that observed differences in money management scores reflect substantive behavioral differences rather than measurement bias, strengthening the empirical foundation for income-based comparisons in financial behavior research.

An Overlooked Population: Why Linguistic Background Matters in Personal Finance Survey Design

This study explores how linguistic background and reading difficulty interact to influence responses to personal finance survey items in the United States. Using data from the NEFE study, the authors analyze item-level reading difficulty, response variance, and response styles among participants who primarily use English versus other languages at home. The results suggest that reading difficulty plays a modest but meaningful role in shaping survey responses and that this relationship differs by language background. The paper's contribution is highlighting linguistic diversity as a critical but underexamined dimension of measurement validity in financial research, and it offers practical recommendations—such as cognitive interviewing—for improving survey inclusivity and accuracy.

Financial Identity Scale: Testing the International Validity of Its Variable-Centered and Person-Centered Models

This study evaluates the Financial Identity Scale (FIS) in an international sample of emerging adults from ten countries, testing both variable-centered (factor structure) and person-centered (identity profiles) models. Using confirmatory factor analysis, approximate measurement invariance, and latent profile analysis, the authors examine whether the scale functions similarly across cultural contexts. Results indicate that the four-factor structure is supported in most—but not all—countries, and that meaningful identity profiles can be identified across cultures. The paper's contribution lies in extending financial identity research beyond U.S. samples, demonstrating partial cross-national comparability, and advancing methodological integration of variable- and person-centered approaches in financial measurement research.

Nine Versions of the Parent Financial Socialization Scale (PFSS)

This paper develops and validates nine versions of the Parent Financial Socialization Scale (PFSS), designed to measure how parents socialize children around finances through modeling, discussion, and experiential learning. The authors create long, short, and minimal versions of the scale for three populations—emerging adults, adolescents, and parents—grounded in family financial socialization theory. Using multiple samples, the study evaluates reliability, factorial validity, and measurement invariance across demographic groups. The findings provide strong evidence of internal consistency and acceptable validity for most versions of the PFSS, though some limitations in measurement invariance (particularly across racial/ethnic groups) are noted. The paper's primary contribution is offering a flexible, theory-based measurement toolkit that allows researchers to assess financial socialization across developmental stages and reporting perspectives, while transparently documenting the trade-offs associated with abbreviated scales.

The Work-Family Enrichment and Work-Family Conflict Scales in the Indonesian Context

This study evaluates the psychometric properties of the Indonesian versions of the Work-Family Enrichment Scale (WFES) and Work-Family Conflict Scale (WFCS) using data from 747 employees, focusing on factor structure, reliability, and gender invariance. The findings show that WFES is best represented by a six-factor multidimensional model and WFCS by a four-factor model, both demonstrating strong internal consistency and good model fit. Importantly, both scales exhibit measurement invariance across gender, indicating they are free from gender bias and suitable for comparing male and female employees. The study's main contribution is providing robust validation of culturally adapted instruments for Indonesia, addressing prior gaps in measurement rigor, and supporting the conceptualization of work-family enrichment and conflict as multidimensional constructs applicable in non-Western contexts.

Evaluating the Klontz Money Script Inventory–Revised (KMSI-R)

This paper assesses the factorial validity, internal consistency, and measurement invariance of the Klontz Money Script Inventory–Revised (KMSI-R) using a large, diverse sample. The authors conduct confirmatory factor analyses and test whether the established four-factor structure holds across racial and ethnic groups. Findings reveal acceptable internal consistency but poor initial model fit and a lack of measurement invariance, even after theoretically justified modifications. The study concludes that the KMSI-R requires substantial revision before it can be reliably used with diverse populations. Its contribution lies in rigorously documenting the limitations of a widely used clinical and research tool, thereby preventing uncritical application and encouraging further scale redevelopment.

Developing and Validating New Measures of High School, College, and Graduate Students' General Financial Knowledge in the United States

This paper develops and validates six new financial knowledge scales tailored to U.S. students across educational stages, covering domains such as budgeting, credit, income, risk, and saving. Using large samples of high school, undergraduate, and graduate students, the authors apply confirmatory factor analysis, Item Response Theory, measurement invariance testing, and known-groups validity analyses. The results show strong psychometric support for the scales and demonstrate systematic differences in financial knowledge by academic level. The paper contributes by providing domain-specific, developmentally appropriate measures of financial knowledge that are especially useful for formative assessment and educational evaluation within the U.S. context.

Validating Financial Well-Being Scales Across Cultures: Evidence from Eastern Cultural Contexts

This study evaluates whether a widely used Western-developed financial well-being scale (Netemeyer et al., 2018) functions equivalently in Eastern cultural contexts, specifically Taiwan and South Korea. Using secondary survey data, the authors assess reliability, factorial validity, and measurement invariance for two dimensions of financial well-being: current money management stress and expected future financial security. Results indicate that while the scale demonstrates acceptable psychometric properties overall, several items are interpreted differently across cultural contexts, and full measurement invariance is not achieved. The paper contributes to cross-cultural measurement research by demonstrating both the promise and limitations of applying Western financial well-being scales internationally, and by emphasizing the need for further cultural adaptation or construct re-examination.

Conclusion

Measurement papers are not ancillary or supplemental; they are essential. Measurement papers, review papers, theory papers, and empirical papers each provide distinct and irreplaceable functions in the progression of high-quality science. The *Journal of Family and Economic Issues* is committed to an increased prioritization of measurement. We invite researchers, reviewers, editors, graduate advisors and instructors, funders, policymakers, educators, and practitioners to give greater attention to measurement issues and thus elevate the validity and effectiveness of family finance research and application.

Author contributions All authors contributed equally to the preparation of this manuscript.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing interest The authors declare no competing interest.

References

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, *71*(2), 287–311. https://doi.org/10.1093/poq/nfm006

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Qui˜nonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, *6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Consumer Financial Protection Bureau (CFPB) (2018). *Measuring financial skill*. https://www.consumerfinance.gov/data-research/research-reports/measuring-financial-skill/

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage.

Duncan, O. D. (1984). *Notes on social measurement: Historical and critical*. Russell Sage Foundation.

Finkelstein, L., & Leaning, M. S. (1984). A review of the fundamental concepts of measurement. *Measurement*, *2*(1), 25–34.

Oishi, S., Graham, J., Kesebir, S., & Galinha, I. C. (2013). Concepts of happiness across time and cultures. *Personality and Social Psychology Bulletin*, *39*(5), 559–577.

Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., & Harney, P. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology*, *60*(4), 570.