Developing a Content-Aligned Eleven-Item Food Insecurity Scale Using a Polytomous Rasch Model

Jing Li[1,*] and George Engelhard, Jr.[2]

[1] Pittsburg State University, Pittsburg, KS 66762, USA
[2] University of Georgia, Athens, GA 30602, USA
*jing.li@pittstate.edu

Abstract. This study develops and evaluates an eleven-item scale for measuring food insecurity using a content-aligned subset of items from the U.S. Household Food Security Survey Module. The scale was calibrated using a polytomous Rasch model framework, specifically the Partial Credit Model, with data from families with children who participated in the 2019 Current Population Survey Food Security Supplement (N=1,248). The resulting HFI-11 scale provides a continuous measure of food insecurity, while preserving the theoretical severity structure of the original HFSSM. Overall, results indicate good model-data fit, with the model accounting for 64.2% of variance in food insecurity. The Wright map reveals a clear hierarchical ordering of items, with household-level reflecting less severe conditions and adult and child items capturing increasing food deprivation, which provides substantive validity evidence for the scale. In addition to supporting continuous measurement, empirically estimated cut scores allow the HFI-11 scale to be linked to USDA food classifications, which promotes both research and policy-relevant application. By directly modeling polytomous responses categories rather than relying on dichotomized scoring, the HFI-11 scale improves measurement precision and sensitivity to changes in food insecurity severity. This study offers a valid and efficient alternative to the full HFSSM and provides a flexible framework for food insecurity measurement and related applications in other fields.

Keywords: Rasch Measurement Theory, Food Insecurity, Polytomous Rasch Model, Substantive Validity, Household Food Security Survey Module.

1 Introduction

1.1 Background and Motivation

Issues related to food security have always been a concern. But recent events, like the COVID-19 pandemic, rising prices, and increasing global conflicts, have made problems worse. These events have disrupted supply chains, lowered household income, and raised the cost of basic food. As a result, many families now are facing greater risk of food insecurity than in previous years (FAO et al., 2022).

© The Author(s) 2026
C. Y. Lye et al. (eds.), *Proceedings of the Pacific Rim Objective Measurement Symposium (PROMS 2025)*, Atlantis Highlights in Social Sciences, Education and Humanities 56,
https://doi.org/10.2991/978-94-6239-834-4_23

Under these circumstances, ensuring people have enough safe and nutritious food is a critical policy goal. The goal speaks to basic human needs. It also speaks to our shared responsibility to support well-being. Access to reliable food sources does more than reduce hardship. It also helps prevent the negative health outcomes linked to food insecurity (Gundersen & Ziliak, 2015).

Because food insecurity carries health and social consequences, we need good tools that help us understand how people are doing. We need clear and consistent ways to measure food conditions at both individual level and household levels. In the U.S., there are several instruments that serve this purpose. The most generally used one is the Household Food Security Survey Module (HFSSM), which is included in the annual Current Population Survey Food Security Supplement (CPS-FSS) and administered by United States Department of Agriculture (U.S. Department of Agriculture Economic Research Service, 2025). Based on the results of HFSSM, households are categorized into four ranges: high food security, marginal food security, low food security and very low food security (Hamilton et al., 1997).

Although the HFSSM has demonstrated strong psychometric properties and is widely used in national monitoring (Bickel at al., 2000; Hamilton et al., 1997), its length and administration burden can pose challenges for applied research, program evaluation, and longitudinal studies. In practice, shorter instruments are often preferred when survey time is limited or when food insecurity is assessed together with multiple outcomes. However, scale reduction raises concerns about whether the shortened versions adequately preserve the theoretical structure and severity continuum of food insecurity. This highlights the need for abbreviated instruments that maintain content coverage and construct representation while improving efficiency.

1.2 Limitations of Categorical Measures

This approach reflects a broader practice in measurement. It's common to see people being categorized into categories for further diagnosis or action, just like what the HFSSM does and what's being suggested to help determine food security status (Bickel at al., 2000). It's also prevalent to see people use categories to help them make classifications. But, as Bowker & Star (2000) suggested in their book related to classification and its consequences, classifications can be "ad hoc, limited to an individual or a local community, and/or of limited duration." They shape how we see people. They can highlight important differences, but they can also hide complexity or fail to detect transitions.

In the case of measuring food security status, households are assigned into categories based on their responses. A household may receive an intervention and show real improvement, but the change may still be too small to move them into a different category. This creates a critical question: Is it accurate to judge the effectiveness or necessity of the intervention solely because the household didn't shift to a higher

category? And is it sufficient to use a single category as a full summary of their food security status, even when their food conditions have improved?

An additional limitation of categorical classification is that it masks the hierarchical nature of food insecurity experiences. The HFSSM was originally designed to reflect a progression from less severe household-level conditions to more severe food deprivation, including household level, adult level and child level. Preserving this severity ordering is central to meaningful interpretation, but it is rarely evaluated explicitly when scales are shortened or when responses are collapsed into categories. Measurement approaches that allow examination of item ordering along a latent continuum are therefore critical for evaluating whether theoretical severity structures are retained.

These issues leave researchers with a broader concern. Are categories a good indicator of determining food insecurity conditions? Or maybe there is a better way to do it. With this question held in mind, this study explores whether there is a better way to capture changes in households' food conditions, one that reduces the limitation that comes from relying on categorical classifications alone.

1.3 Rasch Measurement Theory as an Alternative

Together, these limitations suggest the need for a measurement framework that preserves theoretical structure while allowing for continuous measurement. Because of these limitations, a measurement approach that produces continuous scores and captures small but meaningful changes is needed. Rasch measurement theory offers such an approach, and has been applied in several studies of HFSSM, either to examine its psychometric properties or validate the scale for different populations (Rabbitt & Coleman-Jensen, 2017; Tanaka et al., 2020; Li et al., 2024). This body of work shows why a strong measurement framework is needed when studying food insecurity.

To measure household-level food insecurity well, we need to examine whether the instrument has strong psychometric qualities to ensure the results are accurate and meaningful for further analysis. Several studies have assessed HFSSM from this perspective. Their findings show this survey module demonstrates strong reliability and validity characteristics across different samples and contexts (Tanaka et al., 2020; Marques et al., 2015; Engelhard et al., 2018).

These results suggest that the HFSSM provides a solid foundation for understanding food insecurity. At the same time, they also point to opportunities for refining the scale, especially when researchers need measures that detect gradual changes in households' food conditions. This creates need for continued work using Rasch models to explore item functioning, improve scale structure, and strengthen the overall quality of food insecurity measurement. Knowing what the HFSSM does well and where it could be improved leads to an important question about what kinds of validity evidence are needed to support meaningful score interpretation.

In this context, scale development is not only a statistical exercise, but also a validity argument about how food insecurity is conceptualized, measured and interpreted. Validity is a key indicator of psychometric quality, and validity evidence provides support for both the meaning of test scores and the appropriate uses of an assessment (Cizek, 2020; Embretson, 1983, 2010; Messick, 1995). It is important to remember that validity evidence is always tied to interpretation and use. Messick (1995) described six different aspects of construct validity: content, substantive, structural, generalizability, external, and consequential. Each aspect highlights a different way of evaluating how well a measure represents the construct it is intended to assess.

The current study focuses on validity evidence related to score meaning with an emphasis on the substantive aspect of construct validity. The substantive aspect refers to theoretical rationales for the observed consistencies in test responses, including process models of task performance (Embretson, 1983), along with empirical evidence that the theoretical processes are actually engaged by respondents in the assessment tasks (Messick, 1995). Given the importance of building measures that reflect the processes underlying household food insecurity, the next step is to apply these validity principles to create a scale that captures these processes more effectively.

Building on this work, this study applied Rasch methods to develop an 11-item food insecurity scale supported by strong substantive validity evidence. This evidence comes from careful evaluation of item content, including the use of parallel adult and child items in the modified scale. The scale is calibrated using a polytomous Rasch model, which differs from the dichotomous approach used in USDA classifications and allows for more detailed information about household food conditions.

Rasch (1960/1980) originally introduced a model for dichotomous responses as in the 1950s as part of his work on constructing invariant measurement. The dichotomous Rasch model provides a probabilistic framework for describing the relationship between a person's ability and an item's difficulty. The model can be written as:

$$\phi_{ni1} = \frac{P_{ni1}}{P_{ni0}+P_{ni1}} = \frac{exp(\theta_n-\delta_{i1})}{1+exp(\theta_n-\delta_{i1})} \quad (1)$$

In this case, P_{ni1} is the probability of scoring 1 on item i, P_{ni0} is the probability of scoring 0 on item i. θ_n represents the location of person n on the latent variable, and δ_{i1} is item location on the latent variable where the probability of responding 0 or 1 is equal (Engelhard, 2013).

The USDA applies the Rasch model each year to calibrate the Household Food Security Survey Module. A key concern with this method is that it relies on collapsing rating-scale responses into dichotomous categories. Here is an example of an item that is dichotomized: "The children were not eating enough because we just couldn't afford enough food. Was that often, sometimes, or never true for you in the last 12 months?"

In the dichotomized version, "never true" is coded as 0, while "often" and "sometimes" are collapsed into a single category coded as 1. In this study, we preserve more information by coding the response categories separately: 0 for "never true", 1 for "sometimes", and 2 for "often."

In this study, we use the Partial Credit Model (PCM) is an alternative to the dichotomous Rasch Model (Masters, 1982). In its log-odds form, the PCM can be written as:

$$\phi_{nik} = \frac{P_{nik}}{P_{nik-1}+P_{nik}} = \frac{\exp(\theta_n-\delta_{ik})}{1+\exp(\theta_n-\delta_{ik})} \quad (2)$$

In which P_{nik} is the probability of household n scoring in category k on item i; P_{nik-1} is the probability of household n scoring (k-1) on item i; θ_n is the logit-scale location of household n on the latent variable, and δ_{ik} is the step difficulty for item i at category k (Engelhard, 2013). The Partial Credit Model is widely used calibrate measures with polytomous responses. It allows items to have a different number of response categories rather being restricted to dichotomous scoring. The PCM is appropriate in this study because the items in the scale include both dichotomous and polytomous response formats.

2 Purpose of the Study

Building on prior Rasch-based evaluations of HFSSM, this study develops and evaluates a theoretically aligned, shortened food insecurity scale using a polytomous Rasch framework. Specifically, this study applies PCM to an 11-item subset of HFSSM, referred to as HFI-11 scale, designed to capture household-, adult- and child-level food insecurity. The selected items were drawn from three domains (Household, Adult, and Child) to preserve parallel content and support substantive validity. Across domains, the items capture parallel food insecurity experiences, including whether household/adult/child ate less, cut meal size, hungry or even not eat for whole day. The selected items and parallel content structure are shown in Table 1.

Table 1. Taxonomy for HFI-11 scale.

Reference Group	Content
Household	
1. Worried food would run out	Worry
2. Food bought would not last	Worry
3. Could not afford to eat balanced meal	Worry
Adult Items	
4. Respondent ate less than should have	Ate less
5. Adult(s) cut size or skipped meals	Cut meal size
6. Respondent hungry but did not eat	Hungry
7. Adult(s) not eat for whole day	Not eat for whole day
Child Items	
8. Child(ren) not eating enough	Ate less
9. Cut size of child(ren)s meals	Cut meal size
10. Child(ren) hungry	Hungry
11. Child(ren) not eat for whole day	Not eat for whole day

The purpose of this study is to develop a content-aligned continuous measure of food insecurity for households with children using the Partial Credit Model. This study addresses the following research questions:

- Is the HFI-11 scale psychometrically sound?
- Do the items form an interpretable hierarchy consistent with substantive validity evidence?

3 Methodology

3.1 Participants

This study focused on households with children. All households with children who responded to the selected items in 2019 were included. The final sample includes 1,248 households.

3.2 Instrument

Eleven items were selected from the Household Food Security Survey Module (HFSSM) administered in the US Current Population Survey (Bickel et al., 2000). The original survey module includes 18 items and 11 were selected for the newly developed 11-item scale: three household items, four adult items, and 4 child items. The first three items focus broadly on concerns about food availability. As pointed out earlier, the adult and child items were chosen because they share parallel content areas: ate less,

cut meal size, hungry, and not eat for whole day. Table 1 shows the wording and content for the 11 items.

3.3 Analysis

The data were analyzed using the Facets computer program to fit the Partial Credit Model (Linacre, 2015). Specifically, we examined the model-data fit for the HFI-11 scale, including both person and item fit. In addition, we imputed cut scores on the scale to classify households into four levels of food insecurity: very low food security, low food security, marginal food security, high food security.

4 Results

The HFI-11 scale demonstrates good model-data fit. The Partial Credit Model accounts for 64.2% of the variance in food insecurity among 1,248 households in the sample. The Wright map also showed a hierarchical ordering of item difficulty along the continuum with the child items representing the most difficult to agree with. Both person fit and item fit statistics indicate acceptable fit to the model, suggesting that the scale functions well psychometrically and is suitable for further analysis and measurement.

Table 2 presents the summary statistics for households, items, and cut scores estimated using Partial Credit Model. The mean item location is centered at 0 logits (SD=2.34), while the mean household location is -2.88 logits (SD=1.53). The mean cut score location is also centered at 0 logits, with a standard deviation of 1.24. The overall model-data fit is quite good with both Infit and Outfit statistics close to the expected values of 1. The reliability of separation for households was .51, which is relatively low and reflects the skewed distribution of food insecurity in the sample. In contrast, the reliability of separation of items and cut scores were close to 1, which indicates strong reliability.

Table 2. Rasch summary statistics.

	Households	Items	Cut scores
Measure			
Mean	-2.88	0.00	.00
Standard deviation	1.53	2.34	1.24
Infit mean square error			
Mean	1.01	.99	1.01
Standard deviation	.67	.17	.05
Outfit mean square error			
Mean	.72	.97	.82
Standard deviation	1.09	.15	.19
Reliability of separation	.51	.99	>.99
Number of observations	1248	11	3
Percentage of variance	64.2%		

The Wright Map was constructed and shown in Figure 1. The first column is the measurement scale in logits, followed by the distribution of the households on the food insecurity scale (counts), the third column shows the food insecurity categories, and the final column lists the names of the items according to their location on the latent trait. Item location ranges from Item 18 ("Child(ren) not eat for whole day"), which is the most difficulty behavior to endorse, to Item 1 ("Worried food would run out") which is the easiest. Household, adult, and child items were aligned in the Wright map to allow direct comparison across item contents.

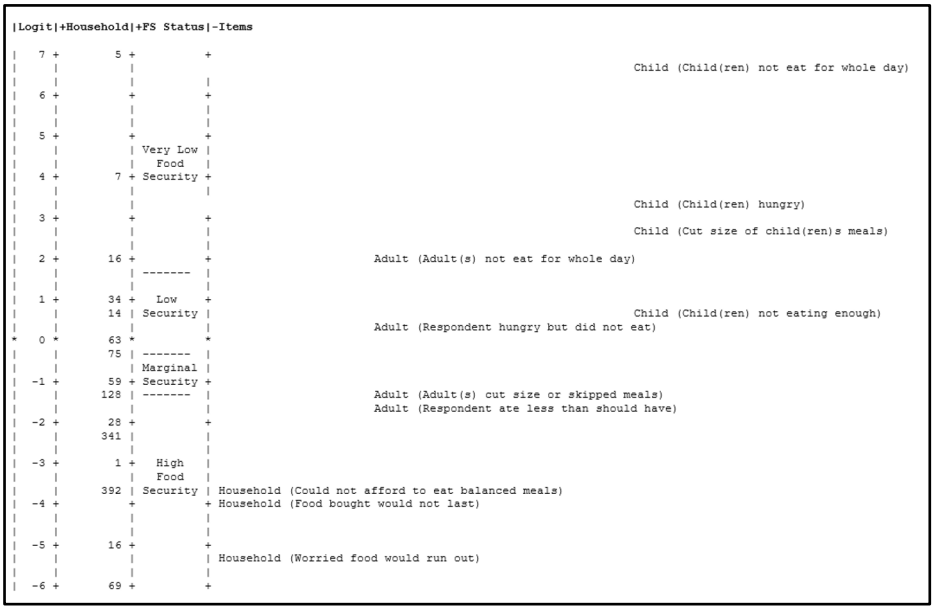


Fig. 1. Wright map

The Wright map shows that child items are consistently more difficult to endorse compared to adult items, while household items appeared to be the easiest. The ordering of the items follows the parallel structure expected based on their content. Specifically, items assessing not eating for whole day (Items 7 and 11) are located at the highest level of severity, followed by items related to hunger (Items 6 and 10), cut meal size (Items 5 and 9), and eating less (Items 4 and 8). This hierarchy supports the intended substantive alignment across the adult and child versions of the items.

Table 3 shows the summary statistics for the items on the HFI-11 scale estimated using Partial Credit Model. As shown in this table, all of the Household, Adult, and Child items exhibit good fit according to their Infit MSE value. However, the Outfit MSE values indicate that Item 8 (“Child(ren) not eating enough”) shows more degree of misfit as it has outfit MSE category belongs to C. The fit values were classified using four categories: MSE values between 0.50 and 1.5 (including 0.50) were placed in category A; values below 0.05 were placed in category B; values between 1.50 and 2.00 (including 1.50) were categorized as C; and values greater than 2.00 (including 2.00) were categorized as D (Wright & Linacre, 1994).

Table 3. Summary statistics for household food insecurity scale (11 items).

Items	Item Location	Standard Error	Infit MSE	Fit Category	Outfit MSE	Fit Category
Household						
1. Worried food would run out	-2.99	.07	1.08	A	1.12	A
2. Food bought would not last	-1.89	.06	1.00	A	1.02	A
3. Could not afford to eat balanced meals	-1.91	.06	1.22	A	1.30	A
Adult						
4. Adult(s) cut size or skipped meals	-1.68	.08	.86	A	.74	A
5. Respondent ate less than should have	-1.99	.08	.85	A	.65	A
6. Respondent hungry but did not eat	-.41	.10	.80	A	.46	A
7. Adult(s) not eat for whole day	1.14	.13	.93	A	.29	B
Child						
8. Child(ren) not eating enough	1.21	.09	1.18	A	1.93	C
9. Cut size of child(ren) 's meals	1.72	.16	1.06	A	1.03	A
10. Child(ren) hungry	2.23	.18	.80	A	.22	B
11. Child(ren) not eat for whole day	4.56	.43	.88	A	.10	B
Mean	.00	.16	.99		1.28	
Standard deviation	3.64	.16	.17		1.16	

In this study, items assigned to Categories A and B were considered productive values for measurement, whereas items in Categories C and D were viewed as unproductive and possibly distorting of measure (Wright & Linacre, 1994). With the exception of Item 8, all items in HFI-11 scale demonstrated good fit based on these criteria. The overall fit of the HFI-11 scale was strong, with the mean values of both Infit MSE and Outfit MSE close to 1.00, indicating good agreement between the model and the data. At person level, the mean infit statistic was 0.72 (SD=1.09), and the mean outfit statistic was 1.01 (SD=.67). Across the range of household measures, between 4.4% and 8.2% were classified as unproductive and possibly distorting for measurement, indicating generally acceptable person fit with a small proportion of unexpected response patterns.

Table 4 summarizes the cut scores estimated using a many-facet Rasch model (MFRM) based on Partial Credit Model, with food insecurity severity categories

specified as an ordered facet for estimating category cut points. Four food insecurity categories were included: high food insecurity, marginal food security, low food security, and very low food security, and the corresponding estimated cut-score locations on the latent continuum were -1.17, -0.12, and 1.29 logits. All three cut points demonstrated good model-data fit, as indicated by Infit and Outfit mean square values close to 1, placing them in fit category A. These corresponding cut-score locations are also displayed in the Wright map in Figure 1.

Table 4. Household cut scores.

	Cut Score	**SE**	**Infit MSE**	**Fit**	**Outfit MSE**	**Fit**
Very low food security						
	1.29	.05	.97	A	.95	A
Low food security						
	-.12	.04	1.06	A	.90	A
Marginal food security						
	-1.17	.06	1.00	A	.60	A
High food security						

Table 5 presents a score table for the HFI-11 scale that links raw sum scores (0-15) to their corresponding locations on the latent continuum of food insecurity and the associated USDA food insecurity classifications, using the cut scores estimated from MFRM-PCM. Specifically, the first column listed the raw sum scores that can be used to categorize households according to selected criteria. The second column shows the corresponding logit locations for each possible sum score, and the third column identifies the cut scores used to assign USDA food security categories. The fourth column displays the linear locations of the four food insecurity levels in relation to the item measures. This table translates latent model estimates into an interpretable and practical scoring framework, illustrating how increasing raw scores connect to higher levels of food insecurity along the latent scale.

Table 5. Scoring table.

Score	Measure	Cut Score	Food Security Status
15	5.89		
14	4.28		
13	3.09		Very Low Food Security
12	2.29		
11	1.65		
		1.29	
10	1.08		
9	0.56		Low Food Security
8	0.06		
		-.12	
7	-0.45		
6	-0.99		Marginal Food Security
		-1.17	
5	-1.60		
4	-2.30		
3	-3.15		High Food Security
2	-4.14		
1	-5.40		
0	-6.97		

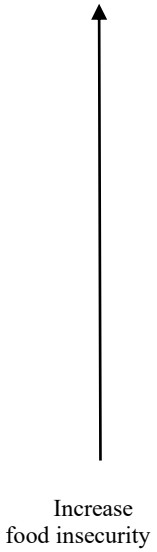


Figure 2 (Panel A) presents the test characteristic curve (TCC), which illustrates the relationship between the latent Rasch measure (logit) and the expected sum raw scores on the HFI-11 scale. It shows how increases in the latent trait correspond to the increase in expected observed cores, which also provides a basis for transforming raw scores into inter-level measures. This information is also summarized in the scoring table in Table 5, specifically, it shows how raw scores from 0 to 15 are transformed onto the scale's metric. Panel B in Figure 2 displays the test information curve for the HFI-11 scale, which shows how much measurement information (precision) the test provides across different levels of the latent trait. The peak of the curve indicates where the instrument measures more precisely, while lower values at the extremes indicate reduced precision for very low or very high levels of the trait (Bond & Fox, 2015). As shown in the figure, the scale provides the maximum information near the center of the continuum, when logit values are close to zero. These findings provide the basis for evaluating the interpretability, validity and practical utility of the HFI-11 scale.

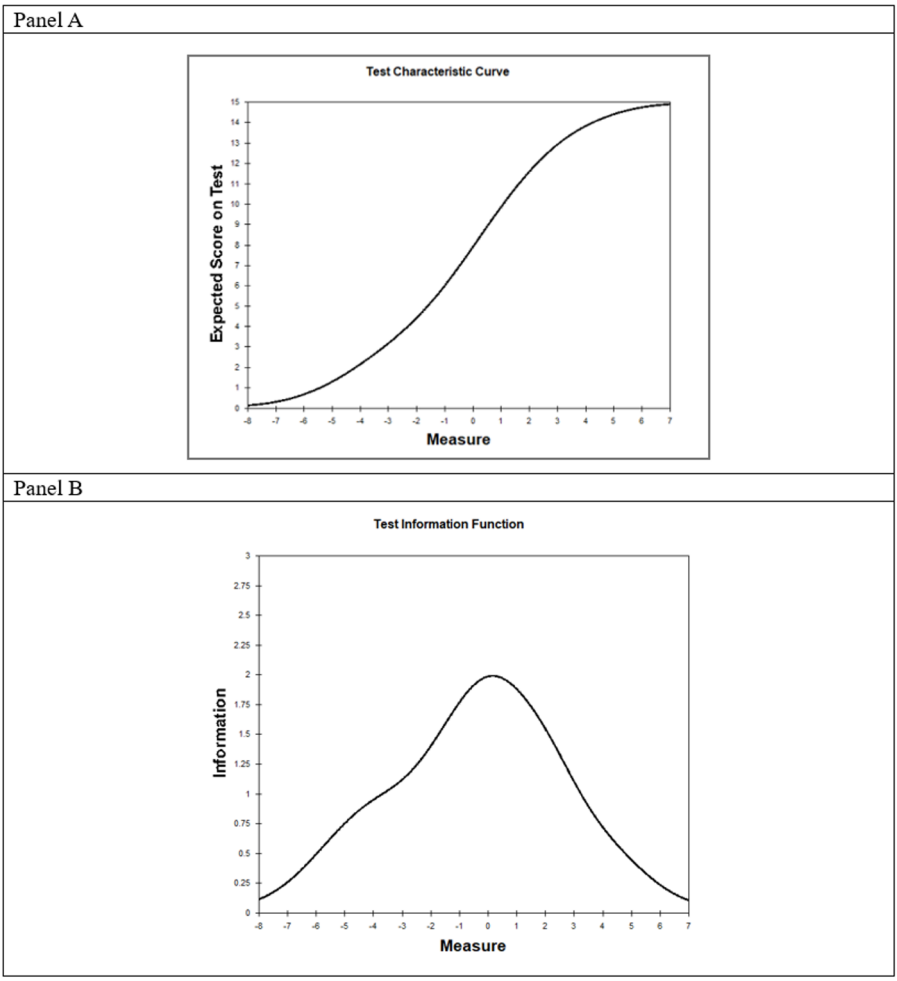


Fig. 2. Test characteristic curve and information function

5 Discussion

This study demonstrates that a theoretically aligned, shortened version of the HFSSM, the HFI-11 scale, can provide psychometrically sound, interpretable and policy-relevant measurement of food insecurity when modeled using a polytomous Rasch framework. The psychometric properties of the HFI-11 scale were evaluated using PCM. Previous work by Tanaka and her colleagues (2020) examined the full 18-item HFSSM and found that the PCM yields better model-data fit and more precise measurement of food insecurity. Building on this evidence, the present study applies PCM to a reduced set of 11 items capturing household, adult and child level food insecurity. Results indicate that the HFI-11 scale demonstrates good model-data fit,

accounting for 64.2% of the variance in food insecurity. The full set of items and response categories for the newly developed HFI-11 scale is provided in Table 6.

Table 6. HFI-11 scale items and response formats.

Domain	Items	Responses
Household items	1. "We worried whether our food would run out before we got money to buy more."	Often / Sometimes / Never true
	2. "The food that we bought just didn't last, and we didn't have money to get more."	Often / Sometimes / Never true
	3. "We couldn't afford to eat balanced meals."	Often / Sometimes / Never true
Adult items	4. In the last 12 months, did you or other adults in the household ever cut the size of your meals or skip meals because there wasn't enough money for food?	Yes / No
	5. In the last 12 months, did you ever eat less than you felt you should because there wasn't enough money for food?	Yes / No
	6. In the last 12 months, were you ever hungry but didn't eat because there wasn't enough money for food?	Yes / No
	7. In the last 12 months, did you or other adults in your household ever not eat for a whole day because there wasn't enough money for food?	Yes / No
Child items	8. "The children were not eating enough because we just couldn't afford enough food."	Often / Sometimes / Never true
	9. In the last 12 months, did you ever cut the size of any of the children's meals because there wasn't enough money for food?	Yes / No
	10. In the last 12 months, were the children ever hungry but you just couldn't afford more food?	Yes / No
	11. In the last 12 months, did your child or any of the children ever not eat for a whole day because there wasn't enough money for food?	Yes / No

This study is based on a content-aligned shorten scale that fits the PCM model and provides a continuous scale for the measurement of food insecurity. In this study,

household-level food insecurity primarily reflects concerns about food availability, whereas adult and child items capture increasing severe experiential deprivation. The Wright map revealed a clear ordering of item difficulty consistent with theoretical expectations, with household-level items representing the least severe manifestations of food insecurity, followed by adult items and child items representing the most severe conditions. This parallel structure across item content domains provides substantive validity evidence for the scale and supports the conceptual framework underlying the original HFSSM. The preservation of this hierarchy in a shortened scale further justifies item reduction without compromising construct representation.

The HFI-11 scale also offers practical advantages for applied research and policy evaluation. By linking raw sum scores to Rasch measures and USDA food security categories through empirically estimated cut scores, the scale allows researchers to obtain both continuous and categorical indicators of food insecurity. The continuous Rasch scale provides greater sensitivity to small changes in food insecurity that may not be detectable using standard categorical classifications alone, making the HFI-11 particularly useful for longitudinal studies and program evaluations.

Item and person fit statistics indicated generally good agreement between the model and the data. With the exception of Item 8 ("Child(ren) not eating enough") shows some evidence of misfit in this set of respondents, all the rest of items demonstrated productive fit for measurement. The observed misfit for item 8 may reflect substantively meaningful response behavior rather than measurement error. Prior research suggests that adults/caregivers may prioritize protecting their children from food insecurity, and they would feed their children before themselves. This potentially leads to underreporting of child food-related hardship even in households experiencing low food security (Ralston et al., 2017). From a measurement perspective, such aberrant response patterns may serve as diagnostic indicators of differential reporting behavior and need further investigation using explanatory Rasch models or differential item functioning analyses.

The relatively low household separation reliability (.51) should be interpreted considering the highly skewed distribution of food insecurity in the sample, with a large proportion of households clustered at lower severity levels. In Rach measurement framework, separation reliability is sensitive to sample variability (Linacre, 1997). Therefore, this value reflects limited spread in household food insecurity rather than poor scale functioning. In contrast, the near-perfect item and cut-score separation reliability indicate that the scale structure itself is highly stable and well-defined.

In summary, the HFI-11 scale offers a valid and efficient alternative to the full HFSSM. Its key strengths include the modeling of the rating scale structure, strong hierarchical alignment across item content domains, a continuous measurement metric with corresponding categorical classifications, and a simple, interpretable scoring framework. Together, these features support the use of the HFI-11 as a flexible tool for

both research and applied contexts where precise and efficient measurement of food insecurity is needed.

Beyond the specific application to food insecurity context, the analytic framework used in this study has broader implications for measurement development and validation in social science and public health. Many instruments in these fields are designed to classify respondents into ordered categories (e.g. risk levels, proficiency levels) while also relying on polytomous response formats. The combination of content-driven item selection, polytomous Rasch modeling, and empirically estimated cut scores provides a valuable approach for linking latent continuous measures to meaningful categorical classifications. This proposed framework offers a flexible and replicable strategy for scale development, validation, and classification across a wide range of educational and psychological measurement applications.

6 Limitations and Further Implications

Several limitations of this study should be acknowledged. First, although the HFI-11 scale preserves the intended severity hierarchy across household-, adult- and child-level items, the household-items included in the shortened scale primarily capture the least severe dimension of food insecurity, concerns or worries about food running out. These items were retained to anchor the lower end of the security continuum and to maintain conceptual continuity with the original HFSSM.

Unlike the adult and child domains, which span a broader range of severity from reduced intake to hunger or not eating for the whole day, the household items do not include indicators of more severe deprivation. As a result, the household domain may provide less differentiation at higher level of food insecurity. Further research could explore the inclusion of additional household-level indicators that reflect more severe food access constraints or examine whether the adult and child items sufficiently capture household-level severity in applied context.

Second, the scale was calibrated using data from households with children participating in the 2019 CPS-FSS, which may limit generalizability to household without children or to population experiencing different economic or policy contexts after 2019. Additional validation studies using diverse samples and alternative data sources are encouraged to evaluate the robustness and generalizability of HFI-11 scale. Finally, although the PCM allows for more detailed use of polytomous responses, the cross-sectional nature of the data limits the ability to directly assess sensitivity to change over time. Longitudinal applications would provide stronger evidence of the scale's utility for monitoring changes in food insecurity severity.

Acknowledgments. The authors declare that this research received no external funding.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bickel, G., Nord, M., Price, C., Hamilton, W., & Cook, J. (2000). Guide to measuring household food security: Revised 2000. U.S. Department of Agriculture, Food and Nutrition Service. Retrieved from https://www.fns.usda.gov/guide-measuring-household-food-security-revised-2000.
2. Bond, T. G., & Fox, C. M. (2015). Applying the Rasch model: Fundamental measurement in the human sciences (3rd ed.). Routledge. https://doi.org/10.4324/9781315814698
3. Bowker, G. C., & Star, S. L. (2000). Sorting things out: Classification and its consequences. MIT press.
4. Cizek, G.J. (2020). Validity: An integrated approach to test score meaning and use. New York: Routledge.
5. Embretson, S. (1983). Construct validity: Construct representation versus nomothetic span. Psychological Bulletin, 93, 179–197.
6. Embretson, S.E. (2010). (Ed.). Measuring psychological constructs: Advances in model-based approaches. Washington, DC: American Psychological Association.
7. Engelhard Jr, G. (2013). Invariant measurement: Using Rasch models in the social, behavioral, and health sciences. Routledge.
8. Engelhard Jr, G., Rabbitt, M. P., & Engelhard, E. M. (2018). Using household fit indices to examine the psychometric quality of food insecurity measures. Educational and Psychological Measurement, 78(6), 1089–1107.
9. FAO, IFAD, UNICEF, WFP, & WHO. (2022). The state of food security and nutrition in the world 2022: Repurposing food and agricultural policies to make healthy diets more affordable. FAO. https://doi.org/10.4060/cc0639en
10. Gundersen, C., & Ziliak, J. P. (2015). Food insecurity and health outcomes. Health affairs, 34(11), 1830–1839. https://doi.org/10.1377/hlthaff.2015.0645
11. Hamilton, W. L., Cook, J. T., Thompson, W. W., Buron, L. F., Frongillo, E. A., Jr., Olson, C. M., & Wehler, C. A. (1997). Household food security in the United States in 1995: Technical report of the Food Security Measurement Project (USDA Miscellaneous Publication No. 344955). U.S. Department of Agriculture. https://doi.org/10.22004/ag.econ.344955
12. Li, J., Kim, S. H., & Engelhard, G. (2024). Validation of the Household Food Security Survey Module (HFSSM) using factor analysis and Rasch measurement theory. In H. Hwang, H. Wu, & T. Sweet (Eds.), Quantitative psychology: IMPS 2023 (Vol. 452, pp. 381–398). Springer. https://doi.org/10.1007/978-3-031-55548-0_22
13. Linacre, J. M. (1997). KR-20/Cronbach alpha or Rasch person reliability: Which tells the "truth"? Rasch Measurement Transactions, 11(3), 580–581. https://www.rasch.org/rmt/rmt113l.htm
14. Linacre, J. M. (2015). Facets computer program for many-facet Rasch measurement, version 3.71.4 [Computer software]. Beaverton, OR: Winsteps.com.
15. Marques, E. S., Reichenheim, M. E., de Moraes, C. L., Antunes, M. M., & Salles-Costa, R. (2015). Household food insecurity: A systematic review of the measuring instruments used in epidemiological studies. Public Health Nutrition, 18, 877–892.
16. Masters, G. N. (1982). A Rasch model for partial credit scoring. Psychometrika, 47(2), 149–174. https://doi.org/10.1007/BF02296272
17. Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741–749.

18. Rabbitt, M. P., & Coleman-Jensen, A. (2017). Rasch analyses of the standardized Spanish translation of the US household food security survey module. Journal of Economic and Social Measurement, 42(2), 171–187.
19. Rasch, G. (1960/1980). Probabilistic models for some intelligence and attainment tests. Copenhagen: Danish Institute for Educational Research. (Expanded edition, Chicago: University of Chicago Press, 1980).
20. Ralston, K., Treen, K., Coleman-Jensen, A., & Guthrie, J. (2017). Children's food security and USDA child nutrition programs (Economic Information Bulletin No. EIB-174). U.S. Department of Agriculture, Economic Research Service.
21. Tanaka, V. T., Engelhard, G., & Rabbitt, M. P. (2020). Modeling household food insecurity with a polytomous Rasch model. In M. Wiberg, D. Molenaar, J. González, U. Böckenholt, & J. S. Kim (Eds.), *Quantitative psychology* (IMPS 2019; Springer Proceedings in Mathematics & Statistics, Vol. 322, pp. 265–276). Springer. https://doi.org/10.1007/978-3-030-43469-4_24
22. U.S. Department of Agriculture Economic Research Service. (2025). Food security in the U.S.-Survey tools. https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/survey-tools
23. Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. Rasch Measurement Transactions, 8(3), 370.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

