Reviewer's report

Title: Reliability and validity of the short form household food security scale in a Caribbean community

Version: 1 Date: 3 December 2003

Reviewer: Stephen J Blumberg

Reviewer's report:

General Comments on the Study Design

This manuscript was written in response to a Commentary by E. A. Frongillo (Int J Epidemiol 2003) that questioned the validity of the Short Form Household Food Security Scale when used without revision in Trinidad and Tobago. Frongillo’s questions arose from (1) the observation that the percentage of affirmative responses to the six scale items were generally higher than previously observed for US samples, from (2) unexplained ethnic differences in food insecurity, and from (3) the observation that the percentage of affirmative responses for the “balanced meals” item was greater than the percentage of affirmative responses for the “food last” item. The latter would not be expected given the food insecurity severity level assigned to each item during the US validation process.

Frongillo suggested that cognitive testing was necessary “to help evaluate whether the performance of the US measure was consistent with that expected” (Frongillo, page 517). Such cognitive testing is not the focus of the present manuscript. Rather than conducting the qualitative analysis requested by Frongillo, this manuscript attempts a quantitative analysis. The majority of the quantitative analysis conducted would be presented well as an evaluation of the reliability rather than validity (with the notable exception of the criterion-related analyses of income and vegetable consumption). By the end of the manuscript, I was not yet convinced that the Short Form was valid in Trinidad and Tobago, though I was convinced that it is reliable.

The Short Form was a simplification of the longer 18-item food security measure. The strength of the longer measure was not only its criterion-related validity, but also its fit to the Rasch measurement model. This model goes one step beyond the properties noted by the authors, that “the items can be ranked according to decreasing frequency of affirmative responses and increasing severity of food insecurity.” (page 3) Rather, this model also assigns a severity score to each item so that each item is placed along the unidimensional continuum of food security/insecurity is such a way that the probability that the item will be endorsed can be determined based on a household’s severity. Similarly, the household severity can be quantitatively assessed on an interval scale (rather than short form’s 0-6 ordinal scale) based on which items were endorsed. These item calibrations, when combined with in-fit and out-fit statistics, can provide invaluable information about the relationships between the items and the food insecurity construct. Whether these calibrations and fit statistics match the calibrations and fit statistics observed in the US would go a long way toward validating the use of the short form in Trinidad and Tobago. As such, I wonder why the authors did not attempt to apply the Rasch model to their data.

[Note: Instructions for applying the Rasch model to validate data from special populations is included in the Guide to Measuring Household Food Security, online at <http://www.fns.usda.gov/fsec/FILES/FSGuide.pdf>. This Guide also includes the standard scores for the Short Form of the scale, based on US data.]
Other General Comments

The sample is adequately described, though the reader is required to consult a previous publication to learn how participants were recruited and to learn the demographic characteristics and other descriptive data. The sample size is adequate for testing the proposed hypotheses.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

a) In Table 3, I was unclear what data were presented numerically, and what data were presented in parentheses. In all tables, percentages are more helpful than frequencies. I recommend dropping the frequencies.

b) The authors should avoid the use of the term “prevalence” throughout the manuscript. It is unclear that the unweighted sample (with some households represented by two raters) is representative of any particular population.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1) The Methods section indicates that each item in the household food security scale was dichotomized. Were appropriate ordinal methods (e.g., Kendall’s tau, point-biserial correlations) used for the correlations and other analyses reported? Is Cronbach’s alpha valid for ordinal data?

2) To what extent are Cronbach’s alpha and the other internal consistency measures still valid given the within-household clustering? (More than one rater was completing the measure for the same household.)

3) Item-total correlations and the frequency of food insecurity obtained after omitting single items may not be appropriate statistical tests for a Guttman or Rasch scale. Item-total correlations and other “omit one” statistics are based on the assumption that all items contribute equally to the measure of the construct. According to this assumption, persons rating one item highly should highly rate all other items. This is not the assumption of the Rasch model. Rather, the Rasch model states that a person giving an affirmative response to one item should also give affirmative responses to all less severe items (within the usual bounds of measurement error). Item-total correlations, then, would be expected to increase as the severity of the item increased. More description of the benefits, limitations, and assumptions of these analyses would be helpful.

4) Given the assumption of the Rasch model, it would be helpful to report the response profiles for persons answering each question affirmatively. Such a table would have a column and row for each item. The column would include data only for households (or raters) with an affirmative response to that item, and the values for each row would reflect the percentage of households (or raters) in that group who affirmatively answered the other items.

5) Why was “green vegetables and salads” the only food frequency variable used in Table 4? Based on the author’s 2003 article, data on frequency of eating fruit, rice, bread, legumes, ground provisions, fish, sweets, and other foods are also available.

6) The authors stated that one objective of this analysis was to investigate “possible ethnic differences in item responses.” This was a clear concern in Frongillo’s Commentary. Yet, the only ethnic differences noted in the results concern Cronbach’s alpha. This is not sufficient to achieve the objective stated by the authors, nor to support the authors’ statement that “our data suggest that the pattern of responses was similar in each ethnic group.”
7) In the U.S., affirmative responses to the “food last” item are not sufficient to indicate that a household is food insecure. In fact, a household with an affirmative response to just that item would be considered food secure. By contrast, a report of not eating balanced meals is considered an indicator of food insecurity. The authors state that “it is not clear that affirmative responses to an item concerning inability to afford balanced meals should always indicate a greater severity of food insecurity than affirmative responses to an item about food not lasting and being unable to buy more.” In fact, based on Rasch scaling, there is clear evidence for that difference in the U.S. Such a difference may not exist in Trinidad and Tobago, and this would be an important finding. But the data presented in this manuscript do not provide “clear” evidence that the difference does not exist. Applying the Rasch model to the data would provide that evidence, as would the cognitive interviews recommended by Frongillo.

8) Finally, following the new analyses recommended above, if the Trinidad and Tobago data still show that the “balanced meals” item is less severe than the “food last” item, the authors are encouraged to consider and report on the implications for researchers using the Short Form.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article of limited interest

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:
None