

The Children's Food Security Scale

The 18 items in the household food security scale include:

- 3 items that ask about experiences of the entire household
- 7 items that ask about experiences and behaviors of the adult members of the household as a group, or of the adult respondent individually
- 8 items that ask about experiences and conditions of the children in the household as a group.

The eight child-referenced items are used to calculate the children's food security scale, a graduated, continuous measure of the severity of food insecurity among children in the household. The scale provides the framework for a corresponding categorical measure of households with children's hunger. This latter measure classifies households as to whether or not the severity of food deprivation reported over the previous 12 months gives clear evidence of hunger among the children in the household. We begin this chapter with a conceptual description of the children's food security scale and of the categorical threshold that identifies households with children's hunger. We then describe statistical assessments of the scale items and of the overall reliability of the scale. The conceptual description is intended to be accessible to all interested readers. The description of the statistical model and its applications is somewhat technical and a moderate level of statistical knowledge is required to understand that section. Detailed information on implementing the children's food security scale is presented in appendix B.

Conceptual Basis of the Children's Food Security Scale

An essential characteristic of both the household food security scale and the children's food security scale is that the items comprising them vary across a wide range of severity of food insecurity. The precise severity level of each item (the "item calibration," discussed below) is estimated empirically from the overall pattern of response to the scale items by U.S. households. However, the range of severity of the

conditions identified by the items is also intuitively evident from inspection of the items. In the children's food security scale, for example, the item, *children not eating for a whole day*, is a more severe manifestation of food insecurity than is the item, *cutting the size of children's meals*, and the latter indicates a more severe level of food insecurity than does the item, *relied on a few kinds of low-cost food to feed the children*. These differences in severity are observed in the response patterns of surveyed households. The more severe items are less frequently affirmed than less severe items. Moreover, a household that affirms an item of midrange severity is likely to have also affirmed all items that are less severe. Similarly, a household that denies an item at midrange is likely to deny all items that are more severe. These typical response patterns are not universal, but they are predominant. And among households that do deviate from the typical patterns, the extent of deviation tends to be slight.

This highly regular pattern of severity ordering of the food security indicator items is the basis of the statistical model used to calculate and assess the food security scales. The statistical model expresses and summarizes in succinct quantitative form the consistently ordered pattern that exists in the U.S. population of the diverse conditions, experiences, and behavioral responses that identify, and to some extent constitute, the phenomenon of food insecurity.⁸ This underlying ordering of the items measuring food insecurity through successive levels of severity also provides a framework for relating

⁸ While the indicator items are diverse in nature, reflecting the inherent multidimensionality of the phenomenon of food insecurity, the highly ordered response pattern among the items reveals an underlying unifying dimension as well—the severity of the phenomenon, as experienced by the surveyed households. This empirical property—the ordered character of the responses—enables the response data to be fitted to unidimensional scale measures of severity both at the household and child-specific levels. The importance of these unidimensional severity scales is, first, that they represent a higher order of measurement than simple categorical or ordered measures, and second, that they provide a methodologically sound basis for estimating the prevalence of specified ranges of severity of food insecurity.

the total number of items affirmed by a household (the household's "raw score") to the conditions existing within the household. For example, households with a raw score of 5 can be characterized as having affirmed the five least severe items and having denied the remaining three more severe items. We can draw on this property of the item set to select an appropriate threshold, in terms of household raw score, to identify households with children's hunger. Table 4 lists the 18 items in the household food security scale, ordered by severity as measured in 1998. The item calibration (described in more detail later in this chapter) is a measure of the severity of each item as inferred statistically from the pattern of responses of all households to all items.

The items in the children's food security scale are particularly strongly ordered. In 1995, among households that affirmed at least one item but not all items, 78 percent conformed exactly to the most typical, or

modal, pattern described above.⁹ That is, these "modal households" affirmed one or more items, beginning with the least severe and continuing in order of increasing severity until, beginning at some particular level of severity for each household, they then denied all of the more severe items. Furthermore, among all households with children, few of the responses that were not perfectly ordered departed very far from this characteristic pattern.

Identifying Households with Children's Hunger—The Categorical Measure

The strong ordering of the indicator items justifies using the household raw scores to identify households with children's hunger. We set the threshold for identifying

⁹ Data from 1995 were used for this assessment of modality because the 1998 and 1999 questionnaires included internal screens to reduce respondent burden that also had the effect of artificially suppressing highly nonmodal responses.

Table 4—Item calibrations based on August 1998 Current Population Survey Food Security Supplement data

Reference	Item description	Item calibration ¹
Household	Worried food would run out	1.488
Household	Food bought didn't last	2.793
<i>Food-insecurity threshold on household scale</i>		
Child	Relied on a few kinds of low-cost food for children	3.268
Household	Couldn't afford to eat balanced meals	3.669
Child	Couldn't feed the children a balanced meal	5.040
Adult	Adult cut size of meals or skipped meals	5.374
Adult/respondent	Respondent ate less than felt he or she should	5.534
<i>Hunger threshold on household scale (identifies households with hunger among adults)</i>		
Adult	Adult cut or skipped meals, 3 or more months	6.424
Child	Children were not eating enough	6.661
Adult/respondent	Respondent hungry but didn't eat	7.545
Adult/respondent	Respondent lost weight	8.613
Child	Cut size of child's meals	8.791
<i>Severe hunger threshold on household scale; children's hunger threshold on children's food security scale</i>		
Adult	Adult did not eat for whole day	9.122
Child	Child hungry but couldn't afford more food	9.240
Adult	Adult did not eat for whole day, 3 or more months	9.934
Child	Child skipped meal	9.935
Child	Child skipped meals, 3 or more months	10.627
Child	Child did not eat for whole day	11.944

¹ Calibrations are based on the standard computational metric (i.e., with discrimination parameter of 1 and mean item calibration of 7; see Bickel et al., 2000). Source: Calculated by ERS based on August 1998 Current Population Survey Food Security Supplement data.

children's hunger at a raw score of five items affirmed. Thus, a household just below this threshold level of severity—almost, but not quite, severe enough to be classified as having hunger among children—will have affirmed four items, typically the four least severe items. That is, the household will report that:

- They often or sometimes relied on only a few kinds of low-cost food to feed the children because they were running out of money to buy food.
- They often or sometimes couldn't feed the children a balanced meal because they couldn't afford that.
- The children were not eating enough because they just couldn't afford enough food.
- They cut the size of the children's meals because there wasn't enough money for food.

A typical household just beyond the threshold—just severe enough to be classified as having hunger among children—will have affirmed those same four items, *and in addition* will have reported that:

- The children were hungry but they just couldn't afford more food.

This last item may be described as the “threshold item” for the range of severity on the measurement scale that corresponds to the category, “food insecure with children's hunger.” The subsequent items within that range, children skipping meals and children not eating for a whole day, are measurably more severe as indicators of children's hunger than the threshold item. Thus, all of the households classified by this method as having children's hunger will have affirmed one or more child-hunger items that are at least as severe as the threshold item. A small proportion (around one-fourth) of the borderline households—those whose response pattern places them just within the children's hunger category—will have denied the threshold item and instead will have affirmed an item that is *more severe* than the threshold item.

In the original work to develop the U.S. household food security scale, Hamilton et al. (1997a) made the first effort to identify households with hunger among children by defining a “severe hunger” category on the 18-item household scale. This household-level severe-hunger category was developed by first identifying the range of severity of children's hunger in households with children, establishing the same line of reasoning that is followed in the current study. That is, a “shadow” threshold for children's hunger was first identified

in the sequence of child-referenced items, as these were combined with the household- and adult-referenced items in the 18-item scale. The “shadow” threshold item was *The children were hungry but we just couldn't afford more food*, which is the same item we have identified as the threshold for the children's hunger range. Then, since an objective of the household scale was that it be consistently applied across all households—both with and without children—the adult-referenced item most nearly corresponding in severity to the “shadow” threshold was designated as the household-level threshold indicator defining the severe-hunger range, or severe-hunger category, for all household types. The threshold item selected was *Adult did not eat for whole day*.

With regard to the child-referenced items, the original severe-hunger threshold corresponds exactly to the children's hunger threshold specified in the current report. However, even though the thresholds are the same in the two scales, the children's food security scale identifies a substantially larger number of households with hungry children. In chapter 3, we discuss the reasons for this difference in prevalence estimates and classification between the severe-hunger category on the household scale and the children's hunger category on the children's food security scale.

Not only is the children's hunger threshold specified in this report consistent with the corresponding severe-hunger threshold in the household-level scale, it also is conceptually consistent with the operational principles developed in setting the household-level hunger threshold on the 18-item scale—in effect, the corresponding adult-hunger threshold. The operational rule of thumb that emerged in the development of the original scale and its associated categorical measure (Hamilton et al., 1997a, 1997b) required that a household affirm multiple indicators of reduced food intake to be classified “food insecure with hunger.” This extent of reduced food intake and disruption of eating patterns indicates the likelihood of the direct experience of hunger—i.e., “the uneasy or painful sensation caused by a lack of food” (Anderson, 1990, following the *Oxford English Dictionary*). This is a methodologically conservative decision rule, designed to minimize false-positive classifications. Similarly, to pass the children's hunger threshold, at least three indicators of reduced food intake among children are required: the children were not eating enough, the size of children's meals was cut, and the children were hungry.

It is important to keep this methodological conservatism in mind when interpreting statistics from either the children's food security scale or the household-level food security scale. Although households beyond the respective thresholds can be labeled quite confidently as "food insecure," or "food insecure with hunger," or "food insecure with hunger among children," households with scores just below the thresholds cannot be described with as much confidence as "food secure," or "not having hunger," or "not having hunger among children." For example, some proportion of the households classified as not having hunger among children do, in fact, affirm one or two items indicating reduced food intake among children due to constrained household resources. In 1998, just under 1 percent of households were classified as food insecure with hunger among children. An additional 4.6 percent of households affirmed one or more items indicating reduced food intake among children. In some of these households, children may have been hungry at times because the family lacked money to buy food, yet these households are classified as not having hunger among children.

There are several detailed respects in which the child-hunger threshold is not precisely parallel to the adult-hunger threshold, due in each case to slight differences between the adult- and child-hunger indicator items available in the CPS food security data. One of these is that the third indicator of reduced food intake among adults—the threshold item for adult hunger—is an item indicating a *repeated pattern* of reduced food intake in 3 or more of the past 12 months. This represents an additional, temporal dimension indicative of increased severity of food insecurity. It also reflects a secondary formal conceptual definition of hunger presented in the 1990 Life Sciences Research Office report: "The recurrent and involuntary lack of access to food" (Anderson, 1990). A strictly parallel procedure in identifying the children's hunger threshold was not possible given the available data. Among the child-specific items, frequency information was collected only for the "skipped meals" question, which is comparatively a more severe indicator item than the threshold item for adult hunger, cutting the size of meals or skipping meals in 3 or more months.¹⁰ In another sense, the children's hunger threshold might

¹⁰ The item that asks about adult hunger also differs from the parallel child item in being referenced only to the respondent, not to "you or other adults in the household."

be viewed as relatively more severe than the adult threshold in that the children's hunger threshold item explicitly affirms that children were hungry, whereas the adult hunger threshold does not require this explicit affirmation for adults.

The discussion of the children's hunger threshold has, to this point, focused primarily on households located just beyond the threshold, those that affirmed exactly five items. However, most of the households classified as having children's hunger affirmed more than five items, indicating that they had experienced still higher levels of food stress. The distribution of response patterns of the households classified as having children's hunger provides a more accurate picture of the conditions and experiences of children in these households (table 5). Almost all households classified as having hunger among children affirmed the three least severe scale items. Nearly 83 percent affirmed the item that asked directly about whether children were hungry; 71 percent reported that children skipped meals because there wasn't enough money for food; and in one-fifth of the households with children's hunger, children went a whole day without eating.

Identifying a Less Severe Threshold of Food Insecurity Among Children

Children may be adversely affected by levels of food insecurity that are less severe than the hunger level (Cook et al., 2001), and may also be affected indirectly by food insecurity and hunger experienced by adults in the household. In an earlier paper (Nord and Bickel, 1999), we specified experimentally a less severe threshold on the children's food security scale that identified households in which the quality and variety of children's diets were reduced. Households that affirmed two or more of the child-referenced items—typically that they relied on a few kinds of low-cost food for the children and that they couldn't feed the children a balanced meal—were classified in that paper as having "reduced quality and variety of children's diet." Based on that threshold, 8.9 percent of households with children would be assigned to the category of reduced quality and variety of children's diets in 1998-99. We have not presented detailed statistics based on that threshold in the current report because there is as yet no expert consensus on whether such a boundary represents a meaningful increment in severity of children's food insecurity, and if it does, what language should be used to describe it. For example, it is not clear that the *nutritional quality* of

Table 5—Proportion of households that affirmed child-referenced items, average 1998-99

Item	In all households with children	In households with hunger among children
	<i>Percent</i>	
Relied on a few kinds of low-cost food for children	15.4	98.8
Couldn't feed the children a balanced meal	8.9	96.0
Children were not eating enough	4.8	91.0
Cut size of child's meals	1.3	78.1
Child hungry but couldn't afford more food	1.0	82.8
Child skipped meal	0.7	70.8
Child skipped meals, 3 or more months	.5	58.4
Child did not eat for whole day	.2	20.2
Affirmed no items, or were screened out	83.5	0
Affirmed 1 item	7.6	0
Affirmed 2 items	4.2	0
Affirmed 3 items	3.2	0
Affirmed 4 items	.7	0
Affirmed 5 items	.3	43.8
Affirmed 6 items	.2	24.8
Affirmed 7 items	.2	24.1
Affirmed 8 items	.1	7.3

Note: Prevalences are based on household supplement weights to represent the national population.

Source: Calculated by ERS based on August 1998 and April 1999 Current Population Survey Food Security Supplement data.

children's diets is necessarily or significantly reduced beyond that particular threshold. More research and discussion are needed to establish a threshold that adequately represents a meaningful range of severity of food stress among children.¹¹

An alternative approach, used in chapter 1 of this report, is to combine information from the household food security scale with information from the children's food security scale to provide a more comprehensive representation of the entire range of food stress and food deprivation in households with children. Households with hunger among children are a subset of households with hunger among adults, just as the latter

are a subset of all food-insecure households.¹² Thus, combining the two scales provides a four-level categorical measure: food secure, food insecure without hunger, food insecure with hunger among adults but not children, and food insecure with hunger among both adults and children.

It is important to note, however, that this tabulation does not represent or reflect a single underlying continuous scale combining household-level and child-specific items. No such scale can adequately represent both household and child-specific food insecurity because these two aspects, while closely related, lie on slightly different dimensions of the general phenomenon. See chapter 3 for a full discussion of the bidimensionality in the 18-item set.

¹¹ The thresholds, or severity-range boundaries, specified with respect to the household food security scale were set, and language was selected to describe the resulting categories, by a team of experts on nutrition and food behavior (Carlson et al., 1999). These thresholds are necessary to create the categorical form of the food security measure (i.e., food secure, food insecure without hunger, food insecure with hunger) needed for monitoring changes in the prevalence of food insecurity and hunger. For this monitoring use, the designated categorical ranges are meaningful in their own right. However, whether these identifiable ranges of severity in the underlying phenomenon provide meaningful indicators for variation in the nutritional and health consequences of food stress and food deprivation is not yet known.

¹² It is logically possible for a household to register hunger among children but not among adults, but in the data sets examined to date—including 6 CPS Food Security Supplements and 3 other nationally representative data sets—no instance of this has been observed, provided that the full 18-item scale is used to identify households with hunger among adults. However, if only the 10 adult and household items are used to identify hunger among adults (using the scale that is normally applied to households without children), then a small proportion of households are classified as having hunger among children but not among adults.

Rasch Measurement Model: Basic Concepts

The 18-item household food security scale was developed through the use of Rasch modeling methods, and we used the same methods to create the 8-item children's food security scale. The Rasch measurement model provides powerful analytic tools to assess the suitability of a set of items for scale construction. We used it to assess the scalability of the child-referenced items and to assess the stability and reliability of the children's food security scale. In essence, the Rasch model formalizes the concept of the severity-ordering of items as discussed above and provides standard statistical measures of the extent to which the response patterns observed in a data set are consistent with this concept. A basic understanding of the Rasch statistical model will help to clarify the meaning of item severity and of other statistics used to assess the children's food security scale and the items that comprise it.

The Rasch model was developed primarily in the educational testing field, where multiple correct/incorrect items, varying in difficulty, are used to measure an individual's level of knowledge or skill. More generally, the model can be used to assess the location of an individual or household along a continuum—in the current case, a continuum of the severity of deprivation in the basic need for food—by combining information from multiple dichotomous (yes/no) items that vary as to the point on the continuum that each one uniquely reflects. This corresponds exactly to the character of the food insecurity/hunger measurement construct. There is no commonly used language that describes the entire continuum of food insecurity and hunger. People do not ask, “On a scale of 1 to 10, how food insecure is your household?” But people do speak readily about specific experiences, such as running out of money for food, and the specific behaviors and conditions that result, such as being forced to cut back on quality or quantity of food. Information about these experiences, behaviors, and conditions then, can be elicited by well-designed survey questions.

The food security scale works well as a measure because the behaviors and experiences represented by the 18 items in the scale correspond closely to the most prevalent experiences and responses of the U.S. population in coping with inadequate resources for food. This result was achieved by basing the questions upon a substantial body of research among low-income U.S. families regarding their experiences of

food deprivation and how they described and coped with them. (Wehler et al., 1992; Radimer et al., 1990; Radimer et al., 1992; Fitchen, 1981; Fitchen, 1988). The questions reflect familiar conditions, experiences, and behaviors, and use natural language derived from the qualitative research to describe them. This set of conditions, experiences, and behaviors reflects the range of variation in severity across the continuum of food insecurity and hunger, just as the set of questions in a well-designed test vary in difficulty across an appropriate range in order to reflect accurately the level of knowledge of the test-taker.

Using the Rasch model to create a measure of food insecurity and hunger assumes that both the indicator items making up the scale and the households responding to the items can be located on the same underlying continuum of severity of food insecurity. It further assumes that the probability of a specific household affirming a specific item depends on the *difference* between the severity level of that household and the severity of that item. The single-parameter Rasch model, which is used to create the household food security scale as well as the children's food security scale, assumes specifically that the log of the odds of a household affirming an item is proportional to the difference between the “true” severity level of the household and the “true” severity level of the item. Thus, the odds that a household at severity level h will affirm an item at severity level i is:

$$p_{h,i}/q_{h,i}=e^{(h-i)}$$

where p is the probability of affirming the item, q is $1-p$, that is, the probability the household will deny the item, and e is the base of the natural logarithms. The probability that the household will affirm the item is:

$$p_{h,i}=e^{(h-i)}/(1+e^{(h-i)})$$

The severity of an item, then, can be thought of as the severity level of households that are just at the threshold of affirming or denying that item. The odds that a household will affirm an item right at the “true” severity level of the household is 1, corresponding to a probability of 0.5. The odds that a household will affirm an item with a severity score 1 unit lower than that of the household is e^{-1} , or about 0.37, corresponding to a probability of 0.27 [i.e., $0.27/(1+0.27)$]. The probability that the household will affirm an item 2 units lower than its own severity measure is 0.12, and for an

item 3 units lower, it is 0.95. Since it is the *difference* between the household and item severity levels that determines the probability of affirmation, it is clear that the metric of the severity scale can be transformed by adding a constant to both household and item scores without changing the character of the scale. That is, the size of the intervals on the scale conveys meaningful information, but the zero point is arbitrary. The Federal interagency food security measurement project has adopted a standard computational metric for the 18-item scale based on a mean item score of 7 for the 18 items in order to keep all item and household scores positive (Bickel et al., 2000). This results in household scores that range from about 1.5 to 13.¹³

Software that implements the Rasch model begins with the household-by-item matrix of responses. Maximum-likelihood methods are then used to estimate simultaneously the household and item severity scores most consistent with the observed responses under the Rasch assumptions. The resulting household scores are a continuous interval-level measure of the extent of food insecurity or hunger in the households. These scores are appropriate for associative analyses such as correlation and regression, with the caveat that the score for households that denied all items or were screened out cannot be estimated by the Rasch model.

¹³ For some communication purposes, the Federal project further transforms the scale to a 0-10 metric by multiplying the scores from the standard computational metric by 10/14. While the 0-10 scale is more intuitively accessible, researchers need to keep in mind that the natural logistic interval length has been modified in this transformation.

The score of zero assigned to such households in some datasets is purely nominal. While Rasch modeling produces an interval scale, the size of the interval between households that denied all items and those that affirmed one item cannot be determined statistically. The Rasch model also provides the basis for “fit” statistics that assess how well each item, each household, and the overall data conform to the assumptions of the measurement model.

Severity of Items in the Children’s Food Security Scale

To create the continuous children’s food security scale, responses to the eight child-referenced items by all households with children were submitted to Rasch modeling software without the adult-specific and household items. The severity order of the eight child-referenced items when scaled without the adult and household items was identical to their order in the all-items scale (table 6; fig. 2).¹⁴ This was expected, since the calibrations of these items in both scales depend only on the responses of households with children. The relative severities of the eight items were also nearly perfectly proportional in the two scales. The correlation of the item calibrations in the two scales in 1995 was 0.997. However, the range of severity of the items was greater when they were scaled alone. This can be seen in figure 2 and can be expressed statistically by

¹⁴ The metric of the child scale in figure 2 was adjusted by adding a constant so that the mean of the calibrations of the 8 child items was equal to the mean of the same 8 items in the 18-item scale.

Table 6—Calibration of items in the children’s food security scale¹

Item	1995		1998	
	Scaled with all 18 Items	Child items scaled alone	Scaled with all 18 items	Child items scaled alone
Relied on a few kinds of low-cost food for children	3.92	1.36	3.27	-.23
Couldn’t feed the children a balanced meal	5.35	3.95	5.04	3.49
Children were not eating enough	6.78	6.32	6.66	6.43
Cut size of child’s meals	8.59	8.95	8.79	9.39
Child hungry but couldn’t afford more food	8.86	9.33	9.24	9.99
Child skipped meal	9.97	10.87	9.94	10.93
Child skipped meals, 3 or more months	10.57	11.70	10.63	11.87
Child did not eat for whole day	12.01	13.58	11.94	13.65
Mean	8.26	8.26	8.19	8.19
Standard deviation	2.56	3.85	2.76	4.35

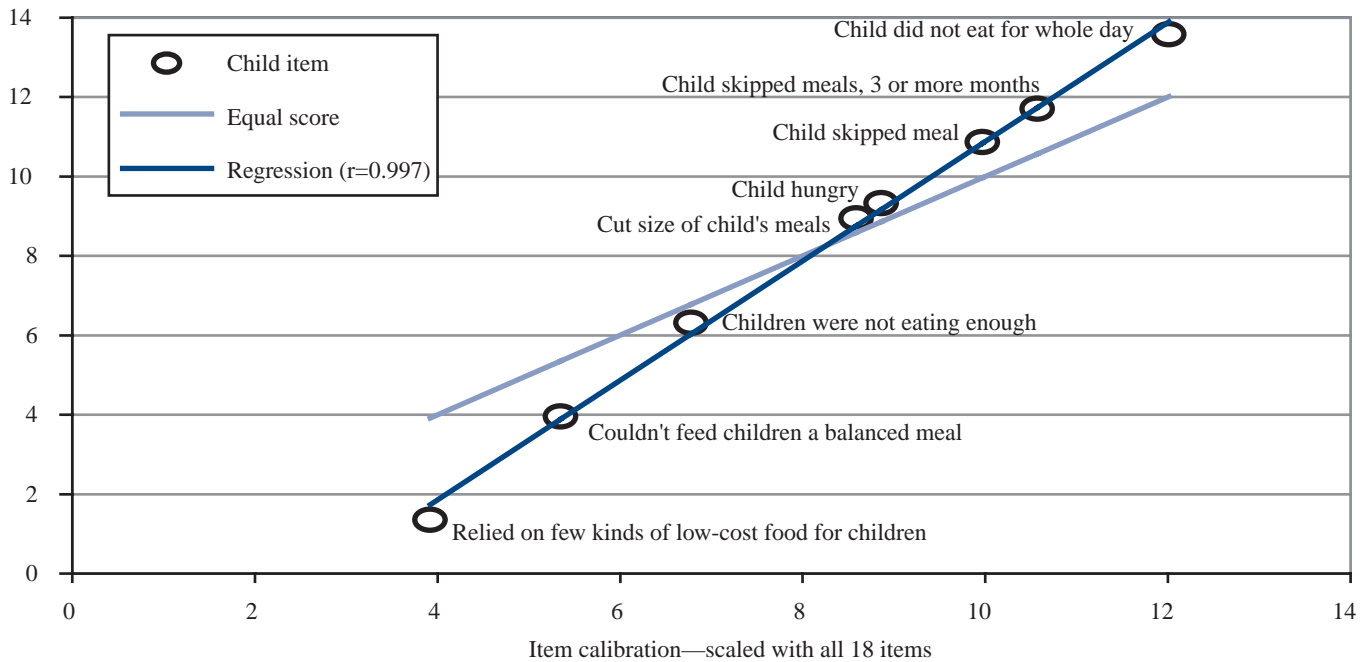
¹Calibrations were calculated under Rasch-model assumptions using joint maximum likelihood methods. The discrimination parameter was set at unity, and mean item score was set at the mean of the eight child items in the standard scale as described in Bickel and Nord (2000).

Source: Calculated by ERS based on April 1995 and August 1998 Current Population Survey Food Security Supplement data.

Figure 2

Item calibration (severity) of child-referenced items scaled alone compared with calibration when scaled with all 18 items

Item calibration—child items scaled alone

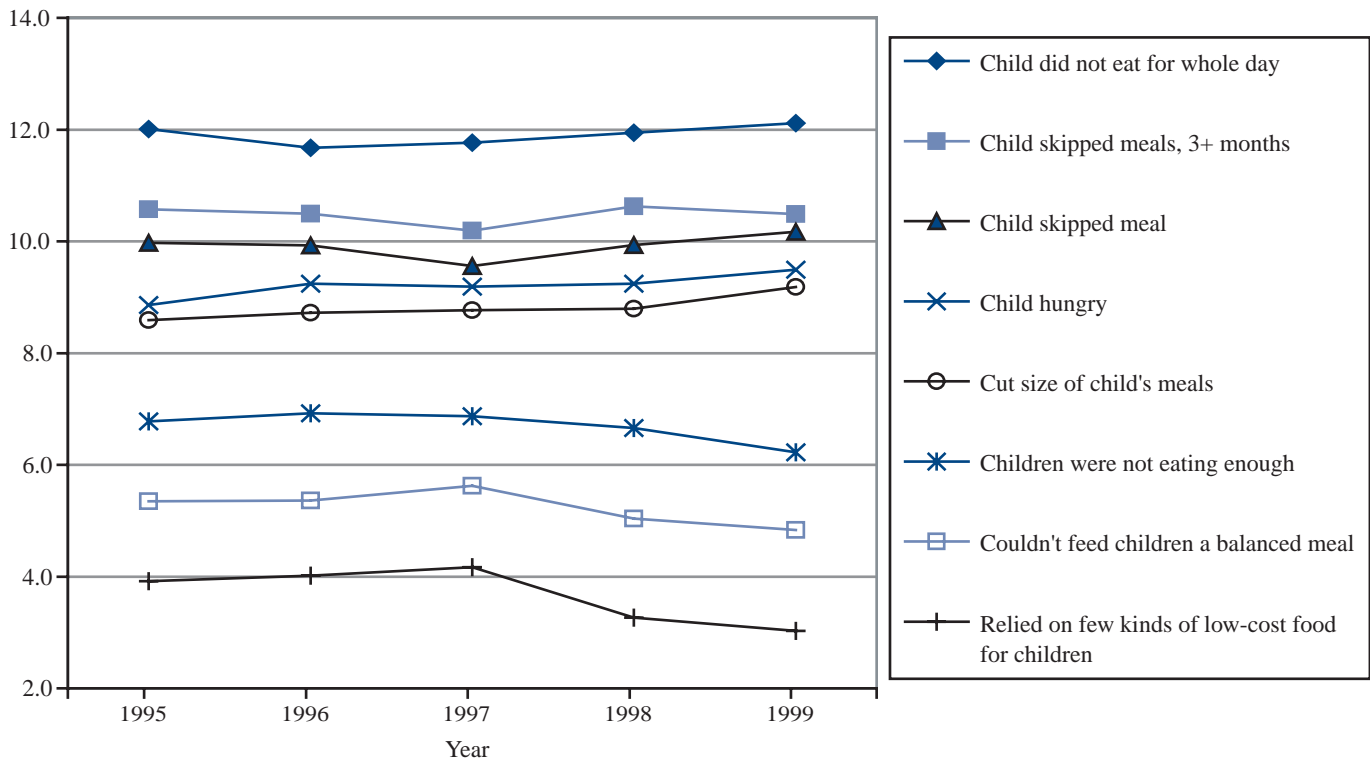


Source: Calculated by ERS based on data from Current Population Survey Food Security Supplements 1995-1999.

Figure 3

Item scores of child items scaled with all 18 items, by year

Item score



Source: Calculated by ERS based on data from Current Population Survey Food Security Supplements 1995-1999.

comparing the standard deviation of the items in the two scales. In 1995, the standard deviation of the child-specific items when scaled with all 18 items was 2.56. When scaled alone, the standard deviation of the same items was about 50 percent greater at 3.85. A similar pattern is seen in 1998. This greater range indicates a more consistent ordering of the child items when they are scaled alone, which translates statistically to higher “discrimination”—i.e., sensitivity—of the items. This is a result of the greater homogeneity of these items as a set compared with the modest bidimensionality (i.e., an adult hunger dimension and a children’s hunger dimension) present in the 18-item set. This slight bidimensionality is, in fact, the reason the severe hunger category of the household scale does not perform well in identifying households with hunger among their child members. We explore this issue further in chapter 3.

The children’s food security scale was highly stable from year to year. The severity-order of the items was invariant across years and relative item severities were quite consistent (fig. 3). The stability of the measurement construct over time increases confidence in the stable relationship of the items to the underlying phenomenon of food insecurity and validates cross-year comparisons of prevalence estimates based on the scale. The greater item range in 1998 and 1999 as compared with earlier years probably resulted largely or entirely from a change in the questionnaire structure of the CPS Food Security Survey and in the screening procedures followed during survey administration. Beginning in 1998, items were administered in approximate order of severity. Households that did not affirm any item in the least severe block (five items for households with children) were skipped over the remaining items and deemed to have denied them. A second, similar screener was inserted after an additional six items. This screening protocol reduces respondent burden and avoids the awkwardness of asking questions that are inappropriate in light of earlier responses. However, screening out a small proportion of highly atypical response patterns that would otherwise be observed also improves the fit of items to Rasch model assumptions.¹⁵

¹⁵ The patterns that are screened out are both improbable statistically and highly unlikely given what we know about households’ normal patterns of behavioral responses to food insecurity. However, it is not known to what extent these are data or survey errors in the form of miscoding by the interviewer or incorrect responses due to misunderstanding or inattention by the respondent, and to what extent they represent actual, but highly unusual, circumstances.

The Children’s Food Security Scale: A Graduated Measure of Food Deprivation Among Children

The categorical measure of hunger among children described earlier in this chapter is appropriate for calculating prevalence rates of children’s hunger and for some analytic purposes. However, a continuous, interval-level measure providing more precise information about the level of severity of food deprivation among children in households is desirable for many associative analyses (correlation, regression, etc.) of the causes and consequences of food deprivation among children. The children’s food security scale—the household severity score derived from fitting the Rasch model to the child-specific items—provides this more detailed and precise measure of food insecurity among children.

A characteristic of the single-parameter Rasch model is that, for households with no missing items, each raw score corresponds to a unique household severity score, irrespective of which items are affirmed to achieve the raw score. The advantage of using the Rasch scores rather than raw scores is that the Rasch-based scores are linear with respect to the underlying phenomenon of food insecurity, while the raw scores are not. As such, the Rasch-based scores are more appropriate for use in linear models such as correlation and regression analyses.

It is desirable to put the children’s food security scale on the same metric, or unit of measure, as used in the household food security scale. We accomplished this by “anchoring” the item calibrations at their values from the 18-item scale and calculating household scores based on these fixed item scores.¹⁶ This assures maximum comparability of the two scales and eliminates the need for a separate set of item calibrations for the children’s food security scale. It introduces no distortion into the children’s food security scale, since the relative item severities of the child items were found to be the same whether they were scaled with or without the adult and household items. Household scale scores and categorical assignments from the children’s food security scale based on data from the 1998 national CPS Food Security Survey are presented

¹⁶ For this calculation, we set the item discrimination at the ratio of the standard deviations of the items in the two scales to reflect the higher discrimination of the child-referenced items taken alone.

in table 7. We selected 1998 data because the core food security module recommended by the *Guide to Measuring Household Food Security, Revised 2000* incorporates the item order and screening first introduced to the national CPS Food Security Survey in 1998, and the household food security scale scores provided by the *Guide* are also based on the 1998 data.

Technically, table 7 is only applicable for households that provide valid responses to all items in the child hunger scale. For households that have missing responses for one or more items, the scale score depends on the number of items affirmed and on which items were omitted. In practice, missing responses to items in the core food security module are rare, so this limitation has slight importance. Further, in most surveys, (including the national CPS Food Security Survey), missing items can be imputed based on valid responses to other items (appendix B), raw scores can be recalculated based on these imputed responses, and household scores can then be assigned from table 7. Thus, for most surveys, the standard methods described in appendix B can be used to assign food security scale scores and status categories, and no Rasch modeling software is required.

Statistical Assessment of the Children’s Food Security Scale

Two statistics commonly used to assess how well responses to items correspond to the Rasch-model

assumptions (or “fit” the model) are “infit” and “outfit.” After item calibrations and household scores have been estimated, the probability of an affirmative response in each cell of the household-by-item matrix is calculated. The infit and outfit statistics are then calculated by comparing the actual responses to the probabilistically expected responses in each cell of the matrix. Infit is an “information weighted” fit statistic for each item that is sensitive to general item fit within the range near the severity level of the particular item. (See appendix A for further information about Rasch model fit statistics.) Outfit is sensitive to unexpected responses from households with severities much higher or lower than that of the item—that is, to highly improbable responses (outliers). Both statistics compare observed deviations of responses from those expected under Rasch assumptions, so the expected value of the statistics is 1. The single-parameter Rasch model, which is used in creating the food security scales, assumes that all items discriminate equally sharply, so fit-statistic values (especially infit) that are far from unity call into question the suitability of the item for use in the scale.¹⁷

¹⁷ The discrimination of an item refers to the rapidity with which the probability of affirmation of the item increases or decreases as household scores diverge from the item score. For the average item in a scale that is fit on the logistic metric, the probability of affirmation by a household with severity one unit above the severity of the item is 0.73. For an item in the same scale that has unusually high discrimination, the probability might be 0.9.

Table 7—Household scores on the children’s food security scale

Number of items affirmed (raw score)	Scale score		Children’s hunger status
	Standard computational metric	Standard 0-10 metric	
0 items (or screened out) ¹	0	0	
1 item	4.1	2.9	No evidence, or insufficient evidence, of hunger among children
2 items	5.9	4.2	
3 items	7.5	5.4	
4 items	8.8	6.3	
5 items	9.6	6.9	Clear evidence of hunger among children
6 items	10.4	7.5	
7 items	11.5	8.2	
8 items ¹	12.2	8.7	

¹Scores for households that affirmed no items or all eight items are not defined under the assumptions of the Rasch model. Here the score for households that affirmed no items is, by convention, shown as zero, but researchers should make allowance in associative models for the fact that the distance between this score and that of households that affirmed one item is not meaningful. The score for households that affirmed all eight items is calculated at a raw score of 7.5. This score can be used for most analyses without distorting the fit of associative models, since these households typically represent a very small proportion of all households.

Source: Calculated by ERS based on August 1998 Current Population Survey Food Security Supplement data to be consistent with the national benchmark scores presented in Bickel et al., 2000.

Item fit statistics for the child items scaled with all 18 items and scaled alone are presented in table 8. Data from the 1995 Food Security Supplement were used to calculate these statistics because there were no internal screeners in that year that might affect fit statistics. Infit can be thought of as a measure of how sharply the item discriminates compared with other items in the model. Except for the two items about skipping meals (a special case discussed below), infits of the child-referenced items when scaled as part of the 18-item scale ranged from 0.92 to 1.14. As a conventional criterion, infits between 0.8 and 1.2 are considered quite good, and 0.7 to 1.3 are acceptable. Thus, the observed infits of the child items indicate that their discrimination characteristics were all near the average of the entire 18-item set. The item that asks directly about whether the children were hungry discriminated somewhat more sharply than the average item. That is, affirmations by households with severity scores lower than that of the item and denials by households with scores higher than that of the item were somewhat rarer than for the average item in the 18-item scale.¹⁸

Infit statistics for the child items scaled alone (i.e., without the household and adult items) were somewhat less consistent. The statistics are not quite comparable between the two scales because many households with children that are included in calculating the 18-item scale are excluded from calculation of the children's food security scale because they are "extreme" on the latter scale. That is, they did not affirm any item in the child scale, so their responses provide no information

¹⁸ It might be thought tempting to use just this single item to categorize households as to children's hunger. However, the discrimination of the scale at a selected threshold is substantially higher than the discrimination of any single item taken alone.

about relative item severities and are excluded from the Rasch model fitting. Further, the children's food security scale, because it is based on fewer items, provides a less precise measure of severity (rather like rounding off a length measurement to the nearest inch). This lower precision also tends to inflate fit statistics. Four items in the children's food security scale, the first two, the fourth, and the fifth had infits near unity. The item about children not eating enough discriminated somewhat more sharply than the average, and the most severe item, children not eating for whole days, did not discriminate as sharply as the other items.

The lower infit statistics (on both scales) for the two items on children skipping meals result in part from a violation of Rasch model assumptions. The Rasch model assumes that items are independent. These two items clearly are not. If a household denies that children skipped meals, they are not asked how often this occurred, and that item is imputed as a denial. There are three sets of such dependencies in the 18-item scale. Analysis (not shown) demonstrates these dependencies have only a negligible effect on item calibrations, but they depress the item fit statistics for the dependent items, and slightly inflate the item fit statistics for the other items in the scale (by artificially increasing the average discrimination). To calculate unbiased fit statistics for the two items about children skipping meals, we reestimated the children's food security scale and fit statistics with each of the two dependent items excluded in turn. The resulting infit statistics were 0.91 for *Child skipped meal* and 0.85 *Child skipped meals in 3 or more months*, much nearer unity than the values shown in table 8 (0.78 and 0.69 respectively). These analyses also reduced the infit statistic of the most severe item, *Child did not eat for whole day*, to about 1.25. Thus, when the biases associated with the item

Table 8—Fit statistics of items in children's food security scale, 1995

Item	Scaled with all 18 Items, all households		Child items scaled alone	
	Infit	Outfit	Infit	Outfit
Relied on a few kinds of low-cost food for children	1.14	1.25	1.08	14.65
Couldn't feed the children a balanced meal	1.04	.88	.85	2.71
Children were not eating enough	1.07	.87	.74	.81
Cut size of child's meals	1.00	1.07	1.00	2.71
Child hungry but couldn't afford more food	.92	.82	.95	3.29
Child skipped meal	.88	.64	.78	4.56
Child skipped meals, 3 or more months	.81	.25	.69	.56
Child did not eat for whole day	1.14	7.95	1.42	30.70

Source: Calculated by ERS based on April 1995 Current Population Survey Food Security Supplement data.

dependencies are removed, infit statistics for the children's food security scale were quite good for six items and acceptable for all eight.

Outfit statistics for the child-referenced items in the 18-item scale were, with one exception, smaller than 1.3, indicating that highly improbable responses were rare. The exception was the most severe item. The higher outfit for that item indicates that a few households affirmed that a child went a whole day without eating, but affirmed very few other items. When the child items were scaled alone, outfits were much less consistent, and those for the least severe and most severe items were quite high. However, outfit can be very sensitive to a very few highly improbable responses. Further analysis (not shown) found that the high outfit of the least severe item, *Relied on a few kinds of low-cost food for children*, was almost entirely a result of responses of five households out of the 2,583 non-extreme households used for calculating the scale. These households affirmed 4, 5, or 6 items, but denied this particular item. Similarly, the high outfit for the most severe item, *Child did not eat for whole day*, was entirely the result of responses of just two households. One affirmed this item but denied all other child items. The second affirmed only this item and the least severe item. These highly improbable responses may result from miscoding by an interviewer, inattention or misunderstanding on the part of the respondent, or from genuine differences in how food deprivation is experienced in different households. Further investigation of characteristics of these households may cast more light on their responses. The rarity of these responses, however, suggests that even the highest outfit statistics seen in the children's food security scale do not seriously undermine the reliability of the overall scale.

Another statistic that can be compared across the two scales is the mean household measurement error. This is the expected mean measurement error of the severity score assigned to households with a given raw score if the data conformed probabilistically to the model expectations. To make the statistics for the two scales comparable, we calculated them across a common group of households—those that were non-extreme on the children's food security scale in 1995. The model error is expressed in the measurement metric, so we adjusted the metric of the children's food security scale so that the item calibrations had the same standard deviation as the child-specific items on the household-level scale, thus making the error estimates comparable.

The larger number of items in the household-level scale provides a more reliable measure of households' food security status, but this is partially offset by the higher discrimination of the child-specific items when they are scaled alone. The mean household standard error was 0.80 for the 18-item scale and 1.02 for the children's hunger scale. The size of the household measurement error varies across the range of each scale, however, and on both scales the threshold relevant for the identification of children's hunger falls in the range of minimum measurement error. Thus, the discrimination of both scales is highest at the severity level most critical for estimating the prevalence of children's hunger, and in this range, the measurement errors were nearly the same, 0.70 for the household-level scale and 0.73 for the child-specific scale.

Summary

A continuous, interval-level, scale measuring the extent of food insecurity among children, and also a categorical measure that identifies households with hunger among children, can be calculated from the eight items in the Food Security Core Survey Module that ask specifically about experiences and conditions of children in the household. The interval-level children's food security scale is appropriate for use as a research tool in associative analyses, while the categorical measure is appropriate for prevalence estimation and for comparing the prevalence of children's hunger among subpopulations.

The children's food security scale is reasonably reliable and is quite stable across years. The eight items in the children's food security scale are strongly ordered in terms of severity. Both descriptive and Rasch-model-based statistical analyses indicate that these items form a coherent scale for measuring the extent of food insecurity—i.e., deprivation in basic food need—among children in households. The severity-order and relative severity of the items were stable over 5 years in spite of a major restructuring of the questionnaire.

The threshold for identifying households with children's hunger is consistent—both conceptually and in terms of the specific response patterns of households near the threshold—with earlier work on identifying households with hunger among children. The measure is methodologically conservative: households classified as having hunger among children have affirmed at least five items, including at least three items indicating reduced food intake by the children.