United States Department of Agriculture

Food and Consumer Service

Office of Analysis and Evaluation

Household Food Security in the United States in 1995

Technical Report of the Food Security Measurement Project

September 1997

Prepared for:

Gary W. Bickel, Project Officer
U.S. Department of Agriculture
Food and Consumer Service
3101 Park Center Drive
Alexandria, VA 22302
under contract no. 53-3198-5-028

Prepared by:

William L. Hamilton, Project Director[a]
John T. Cook, Principal Investigator[b]
William W. Thompson[a]
Lawrence F. Buron[a]
Edward A. Frongillo, Jr.[c]
Christine M. Olson[c]
Cheryl A. Wehler[d]

[a] Abt Associates, Inc.
[b] Tufts University Center on Hunger, Poverty, and Nutrition Policy
[c] Cornell University Division of Nutritional Sciences
[d] C.A.W. and Associates

TABLE OF CONTENTS

Chapter One INTRODUCTION 1

Chapter Two METHODS AND RESULTS OF FITTING LINEAR AND NON-LINEAR FACTOR ANALYSIS MODELS TO CPS DATA 5

2.1 Preliminary Linear Factor Analysis 8
2.2 Exploratory Two-Parameter Non-linear Factor Analysis Model . . 10
2.3 Unidimensional One-Parameter Non-linear Factor Analysis Models 13
2.4 Summary 28

Chapter Three RELIABILITY ESTIMATES FOR THE FOOD SECURITY SCALES 29

3.1 Spearman-Brown Split-half Reliability Estimates 31
3.2 Rulon's Split-Half Reliability Estimates 33
3.3 Cronbach's Alpha Reliability Estimates 34
3.4 Rasch Model Reliability Estimates 35
3.5 Reliability in Identifying Cases with No Food Insecurity Problems 38
3.6 Summary 40

Chapter Four DEFINING RANGES OF THE FOOD SECURITY SCALE 43

4.1 Conceptual Basis for a Categorical Food Security Status Variable 43
4.2 Defining Ranges and Selecting Scale Cutpoints 45
4.3 Evidence of Food Insecurity 50
4.4 Subjective Reporting of Hunger 53
4.5 Evidence of Child Hunger and Severe Adult Hunger 56
4.6 Summary 58

Chapter Five THE RESOURCE AUGMENTATION QUESTIONS 61

5.1 Two Dimensions of Food Insecurity 61
5.2 The Composite Resource Augmentation Index 65
5.3 Effects of Using the Composite Resource Augmentation Index . . 66
5.4 Summary 67

Chapter Six EXTERNAL CONSTRUCT VALIDATION OF THE FOOD SECURITY MEASURES 69

6.1 Relationship of Construct Validation Items to Food Security . . . 69
6.2 Weekly Food Expenditures per Household Member 70
6.3 Household Income 72
6.4 Food Sufficiency 76
6.5 Summary 77

Chapter Seven PROCEDURES FOR CALCULATING STANDARD ERRORS FOR FOOD SECURITY PREVALENCE ESTIMATES 79

7.1 CPS Sample Design 79
7.2 Adjustment Factor for Between-PSU Variance 80
7.3 Estimation of Within-PSU Variance 81
7.4 Calculation of the Standard Errors 84

Chapter Eight POTENTIAL SOURCES OF BIAS IN PREVALENCE ESTIMATES . 85

8.1 Screening Bias 86
8.2 Response Bias 88
8.3 Random Error in Survey Responses 89
8.4 Summary 92

REFERENCES 93

Appendix A REVIEW OF LITERATURE FROM PHYSIOLOGY AND CLINICAL NUTRITION RESEARCH ADDRESSING THE NATURE OF HUNGER

Appendix B PREVALENCE OF HOUSEHOLD FOOD SECURITY STATUS (30-DAY SCALE)

Appendix C PARTICIPANTS IN FEDERAL INTERAGENCY WORKING GROUP FOR FOOD SECURITY MEASUREMENT

CHAPTER ONE
INTRODUCTION

In April 1995, the U.S. Bureau of the Census conducted the first collection of comprehensive food security data as a supplement to its regular Current Population Survey (CPS). With about 45,000 household interviews, this survey is the first to collect the special data needed to measure food insecurity and hunger in a nationally-representative sample of U.S. households.

The Food and Consumer Service (FCS) of the U.S. Department of Agriculture led the effort to develop the Food Security Supplement to the CPS, building on research conducted at universities and elsewhere over the past decade. After the survey was conducted, the next step was to analyze the data to create measurement scales that gauge households' levels of severity of food insecurity and hunger. FCS contracted with Abt Associates Inc. and three subcontractors — the Tufts University Center on Hunger, Poverty, and Nutrition Policy; the Cornell University Division of Nutritional Sciences; and CAW and Associates — to carry out the scale construction analysis.

The results of that analysis are presented in *Household Food Security in the United States in 1995: Summary Report of the Food Security Measurement Project*, to which this report is a companion volume. The purpose of this report is to describe the analyses through which the food security scales and food security status variable were developed, as well as related tests of the reliability and validity of these measures.

Two scales were developed to measure the degree of food insecurity and hunger in American households. One measures food insecurity and hunger over the period of the 12 months prior to the survey interview, and the second measures these conditions in the 30 days immediately preceding the interview. After a number of exploratory analyses, a type of non-linear factor analysis known as a Rasch model was used to form the scales. This methodology and the procedures through which it was applied are described in Chapter Two.

The two scales were subjected to a variety of tests of reliability, including tests specific to the Rasch model and more traditional tests commonly used with scales developed through linear factor analysis. The results, presented in Chapter Three, generally indicate good reliability for the 12-month scale. The 30-day scale, because it is based on a smaller number of questions

and provides detailed measurement for a narrower portion of the food insecurity spectrum, has somewhat lower reliability.

The two scales serve as the basis for defining two corresponding food security status variables. The 12-month variable has four categories: (1) Food Secure; (2) Food Insecure with No Hunger Evident; (3) Food Insecure with Moderate Hunger Evident; and (4) Food Insecure with Severe Hunger Evident. The 30-day scale has three categories: (1) No Hunger Evident; (2) Food Insecure with Moderate Hunger Evident; and (3) Food Insecure with Severe Hunger Evident.

To classify households into the various categories, it was necessary to define ranges on the 12-month and 30-day scales that correspond to each category. The rationale for the range definitions is described in Chapter Four.

The food security scale and the food security status indicator represent a central dimension of food insecurity: availability of enough food for the household to meet basic needs. The concept of food insecurity has other dimensions, however, including the specification that households should be able to acquire food in socially acceptable ways. Because the CPS Supplement includes several indicators of "coping" or "resource augmentation" behaviors related to this dimension of food insecurity, the possibility was explored of supplementing the primary food security scale with an index of resource augmentation actions. The analysis, described in Chapter Five, suggests that such an index should not be used in classifying households' food security status at this time.

A key question for any new scale is how accurately it represents the condition it attempts to measure. Ideally, one would compare the food security scales and status variables to some more definitive measure or measures of food insecurity and hunger. Because no such definitive measure exists, the best way to judge the measure is to assess its relationship to other measures thought to be related to food insecurity and hunger, such as the household's level of food expenditures or its total income. Chapter Six presents the results of such analyses, which show relationships of the sort that would be expected with a valid measure of food insecurity and hunger.

The central purpose of the food security scales and the status variables is to assess the food security of the U.S. population and of subgroups within the population. Estimates of the prevalence of food insecurity and hunger are presented in the study's main report, based on the

April 1995 data. Because these data come from a sample of households, prevalence estimates are subject to sampling error, and the report therefore presents estimated standard errors corresponding to the estimated prevalences. The estimation of standard errors is complicated by the multi-stage sampling design used by the CPS. Chapter Seven describes the methodology used in the estimation of standard errors.

Finally, Chapter Eight discusses the potential sources of bias in prevalence estimates that might result from the sample design of the CPS, from household response behaviors to the Food Security Supplement, and from the fact that only a small proportion of the population experiences food insecurity. The analysis indicates that the various potential sources of bias probably lead to quite small levels of estimation error in counterbalancing directions.

CHAPTER TWO

METHODS AND RESULTS OF SCALING ANALYSIS OF CPS DATA

This section describes the rationale and the results of conducting preliminary linear factor analyses and subsequently fitting a series of non-linear factor analysis models to the CPS food security data. This latter analysis approach more accurately characterizes the covariation among items in the CPS data set than more traditional linear factor analysis models. Most items available for analysis in the CPS data set were severely skewed and dichotomous or categorical in nature. Therefore, a number of statistical assumptions were violated using the linear factor analysis methods with the CPS items, such as the assumption of normally distributed error variance. Such situations can be dealt with more appropriately using non-linear scaling techniques.

Item Response Theory (IRT) describes a general model that was developed by the educational testing industry to assist in creating valid and reliable aptitude tests, such as the Scholastic Aptitude Test (SAT) and the American College Testing Program (ACT) test. When applying a particular IRT model to data, the test designer usually assumes that the responses to a set of items can be accounted for by latent traits or factors that are fewer in number than the test items. The primary goal is to determine how an individual with a certain ability level will respond to an item associated with a particular difficulty level. There are a number of alternative forms the IRT model can take, depending on the assumptions regarding how the underlying data were generated.

The three most frequently discussed IRT models in the literature are (1) the three-parameter logistic model, (2) the two-parameter logistic model, and (3) the one-parameter logistic model. The three-parameter logistic IRT model is the most complex, and can include varying discrimination parameters, varying difficulty levels, and varying guessing parameters. Using the notation of Hambleton (1983),[1] the three-parameter logistic model can be written as follows:

[1] Hambleton, R.K. (ed.), *Application of Item Response Theory*, Vancouver: Educational Research Institute of British Columbia, 1983.

$$P_{ni}(\theta_n) = c_i + (1-c_i)\frac{e^{Da_i(\theta_n - b_i)}}{1 + e^{Da_i(\theta_n - b_i)}} \quad (1)$$

where

θ_n = latent trait score of person *n*,

a_i = item discrimination parameter for item *I*

b_i = item difficulty for item *I*,

c_i = guessing parameter for item *I*,

n = person, and

I = item.

The two-parameter logistic model assumes that guessing does not occur, and therefore the guessing term is dropped from the model. The two-parameter logistic model can be expressed as follows:

$$P_{ni}(\theta_n) = \frac{e^{Da_i(\theta_n - b_i)}}{1 + e^{Da_i(\theta_n - b_i)}} \quad (2)$$

where

θ_n = latent trait score of person *n*,

a_i = item discrimination parameter for item *I*,

b_i = item difficulty for item *I*,

n = person, and

I = item.

Finally, the one-parameter logistic model is a more straightforward model relative to the two previous models, because the model (1) has no guessing parameters, and (2) specifies that all items have the same discrimination parameter ($\bar{a}$). That is, the slopes of the item-characteristic curves are constrained to be equal for all items. The model can be written as follows:

$$P_{ni}(\theta_n) = \frac{e^{D\bar{a}(\theta_n - b_i)}}{1 + e^{D\bar{a}(\theta_n - b_i)}} \quad (3)$$

where

θ_n = latent trait score of person *n*,

$\bar{a}$ = average item discrimination parameter for item *I*,

b_i = item difficulty for item *I*,

n = person, and

I = item.

Because *D* and $\bar{a}$ are constants in the model, the one-parameter logistic model can be written in a more simplified form:

$$P_{ni}(\theta_n) = \frac{e^{(\theta_n^* - b_i^*)}}{1 + e^{(\theta_n^* - b_i^*)}} \quad (3)$$

We can also express this model using the notation of Wright and Masters (1982):

$$P_{nik}(\theta) = \frac{e^{[\beta_n - (\delta_i + \tau_k)]}}{1 + e^{[\beta_n - (\delta_i + \tau_k)]}} \quad (4)$$

where

β_n = latent trait score of person *n*,

δ_i = item difficulty for item *I*,

τ_k = threshold parameter for step k of item *I*,

n = person,

I = item, and

k = step,

and include a threshold parameter that is associated with the rating scale model developed by Andrich (1978, 1979).

2.1 PRELIMINARY LINEAR FACTOR ANALYSIS

The CPS Food Security Supplement builds on a substantial amount of recent research on the measurement of food insecurity, some of which included scaling analysis.[2] The first analytic step was to replicate some of the prior analyses to determine whether the general patterns and relationships in the data were similar to those seen in prior work.

A series of linear factor analyses were fit to the CPS data. One illustrative model, summarized in Exhibit 2-1, was fit for households with children (because this group was asked all questions in the Supplement). The factor model incorporated a Procrustes rotation, which allows one to rotate to a pre-specified factor solution, where the solution was specified to represent the dominant themes of the prior research. Fitting the factor analysis model resulted in three factors with eigenvalues greater than 1.0 prior to rotation (15.0, 1.6, and 1.4), with factor loadings as shown in the exhibit. The first factor includes primarily items related to child food intake reductions and hunger, the second consists mainly of household-level food insecurity items, and the third comprises mainly items related to adult food intake reduction and hunger.

In sum, the results generally confirmed that the response patterns in the CPS data were similar to those seen in prior research and that similar relationships might be expected to exist. In addition, the large positive factor intercorrelations suggested the possibility that non-linear factor analysis methods might result in the items loading onto a single factor (i.e., that the separation of factors could occur in part because of the limitations of linear factor analysis in handling low-frequency dichotomous items). Finally, exploratory analyses of groups of households without children suggested that, for those items applicable to all groups, the factors might be relatively invariant across groups.

2.2 EXPLORATORY TWO-PARAMETER NON-LINEAR FACTOR ANALYSIS MODEL

[2] Two key prior studies are Olson, Frongillo, and Kendall (1995), and Scott, Wehler, and Anderson (1995). The first study estimated a factor analysis model including four items from the Community Childhood Hunger Identification Project (CCHIP) and ten items from two previous Cornell surveys. The analysis identified two key factors, one associated with household-level food insecurity and one associated with hunger. The second study, analyzing data from multiple CCHIP studies, found a first factor comprising mainly household-level food insecurity items and adult hunger items, whereas the second factor included mainly child hunger items.

Exhibit 2-1

SUMMARY OF FACTOR LOADINGS FOR LINEAR FACTOR ANALYSIS MODEL
(*n*=2,991)

Items	Standardized Regression Coefficients		
	F_1	F_2	F_3
Q11		38	
Q15		59	
Q16		63	
Q20		52	
Q24			45
Q28			52
Q32			47
Q35			48
Q38			43
Q40	50		
Q43	42		
Q47	60		
Q50	40		
Q53		78	
Q54		76	
Q55		78	
Q56		73	
Q57	49		
Q58		75	

Initially, we fit a series of exploratory non-linear factor analysis models to determine the dimensionality of the Food Security Survey items.[3] From these alternative models, we selected one representative non-linear model, labeled M121, which best describes the consistent findings across the various alternative models. M121 was fit as a two-parameter logistic model that included estimates for both factor loadings (discrimination parameters) and uniquenesses (error term).[4] Descriptive statistics for the subsample of 994 subjects and 21 items are presented in Exhibit 2-2. The items ranged in proportion of positive responses from .850 (item 15) to .004 (item 50), where the higher the proportion, the lower the severity of food insecurity indicated by the particular item.

The results of the non-linear factor analysis model are presented in Exhibit 2-3. The primary fit statistic, the root mean square residual (RMSR) suggested that the one-factor model adequately fit the data (RMSR = .0074). That is, the RMSR was well within the acceptable range with a single factor, and was not materially improved by adding further factors, making the single-factor model the most parsimonious solution. As with the linear factor analysis model, items 15 and 23 were poor-fitting, with low factor loadings (.31 and .22, respectively). Item 22 had a moderately positive factor loading (L = .43), whereas the rest of the items all had large positive loadings above .50. The findings support the linear factor analysis results with respect to item fits, but suggest that items 15 and 23 should be removed from subsequent models.

2.3 UNIDIMENSIONAL ONE-PARAMETER NON-LINEAR FACTOR ANALYSIS MODELS

The exploratory non-linear factor analysis models indicated that the Food Security Survey items could be described efficiently as a unidimensional construct. Therefore, we pursued a specific non-linear factor model called the Rasch model. The Rasch model is a concise one-

[3] Exploratory non-linear factor analysis models were fit using two software packages: LISCOMP and NOHARM. LISCOMP is a structural equation modeling program that is designed to work with dichotomous and/or ordinal data. NOHARM is a non-linear factor analysis program that analyzes moment matrices. Both programs allow one to fit a two-parameter item response theory model (non-linear factor analysis model) to the data. Exploratory analysis focused on households with children in random 25 percent subsamples of the Food Security Supplement sample. Households that did not pass the series of screening questions (i.e., higher-income households with no indication of food insecurity), and consequently were not asked the full series of food insecurity and hunger questions, were excluded from the analysis.

[4] The two-parameter model can be fit with either item difficulty or uniqueness as the second parameter. The specification shown here chose the uniqueness parameter.

Exhibit 2-2

DESCRIPTIVE STATISTICS FOR MODEL M121

Variable	Mean	Std	Sum
Q11	.231	.421	231
Q15	.850	.356	850
Q16	.450	.497	450
Q18	.325	.468	325
Q19	.095	.293	95
Q20	.274	.446	274
Q21	.585	.492	585
Q22	.122	.327	122
Q23	.016	.125	16
Q24	.244	.429	244
Q28	.054	.226	54
Q32	.233	.423	233
Q35	.123	.328	123
Q38	.047	.211	47
Q40	.048	.213	48
Q43	.023	.150	23
Q47	.049	.216	49
Q50	.004	.063	4
Q53	.600	.490	600
Q54	.434	.495	434
Q55	.398	.489	398
Q56	.267	.442	267
Q57	.137	.344	137
Q58	.377	.484	377

Exhibit 2-3

SUMMARY OF FACTOR LOADINGS FOR MODEL M121

Item	Item Label	Standardized Regression Coefficients
		F_1
Q11	General food sufficiency question	70
Q15	Try to make food or money go further	31
Q16	Run out of foods needed to make meal	70
Q18	Borrow food or money to make meal	56
Q19	Take child to other home for meal	68
Q20	Serve few low-cost foods several days in a row	73
Q21	Put off paying bills to buy food	51
Q22	Get emergency food from church or food bank	43
Q23	Eat meal at soup kitchen	22
Q24	Adults cut or skip meals because not enough money for food	89
Q28	Adults don't eat for whole day	79
Q32	Eat less than should because not enough money to buy food	88
Q35	Hungry but don't eat because can't afford to	85
Q38	Lost weight because not enough food	75
Q40	Child's meal size cut because not enough money for food	76
Q43	Child skip meal because not enough money for food	60
Q47	Child hungry but can't afford more food	80
Q50	Child did not eat for a whole day	71
Q53	Worry food will run out before getting money for more	79
Q54	Food doesn't last and don't have money to get more	89
Q55	Can't afford to eat balanced meals	88
Q56	Can't feed children a balanced meal	85
Q57	Child not eating enough because can't afford more food	83
Q58	Child fed only few low-cost foods, running out of money	82

factor model that constrains the discrimination parameters (factor loadings) to be equal across all items. The statistical constraints of the Rasch model result in several desirable properties for the measurement scale, especially its robustness across multiple samples and multiple variations of the test (Wright and Masters, 1982). Furthermore, the preliminary exploratory models indicated that most of the items had very similar discrimination parameters when the discrimination parameters were allowed to vary.[5]

The computer program BIGSTEPS was designed specifically to fit the unidimensional Rasch model. All subsequent models described in this section were fit using BIGSTEPS.

Five alternative measurement models based on existing theoretical frameworks were generated for the Food Security Survey items. The five alternative models are summarized in Exhibit 2-4. For most of the models, the items were divided into two subsets based on the specific time frame that the items referenced. For models R101, R102, and R103, the first subset of items references behaviors and events that occurred in the last 12 months, whereas the second subset references behaviors and events that occurred in the last 30 days. Models were fit separately for the 12-month and 30-day time periods.

[5] Note in Exhibit 2-3 that nearly all factor loadings fall in the fairly narrow range from 70 to 88. The questions with loadings substantially outside this range (Q15, Q18, Q21, Q22, Q23) are all ultimately excluded from the scale.

Exhibit 2-4

ALTERNATIVE NON-LINEAR FACTOR ANALYSIS MODELS

Model	12-Month Scale	30-Day Scale
R101	Scale includes items that referenced events that occurred in the last 12 months. Items 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 28, 29, 32, 35, 38, 40, 43, 44, 47, 50, 53b, 54b, 55b, 56b, 57b, 58b.	Scale includes items that referenced events that occurred in the last 30 days. Items 17, 26, 27, 30, 31, 33, 34, 36, 37, 39, 41, 42, 45, 46, 48, 49, 51, 52.
R102	Scale includes items that referenced events that occurred in the last 12 months, and excludes resource augmenting behaviors (18, 19, 21, 22, and 23). Items 15, 16, 20, 24, 25, 28, 29, 32, 35, 38, 40, 43, 44, 47, 50, 53b, 54b, 55b, 56b, 57b, 58b.	Scale includes items that referenced events that occurred in the last 30 days, and excludes resource augmenting behaviors. Items 17, 26, 27, 30, 31, 33, 34, 36, 37, 39, 41, 42, 45, 46, 48, 49, 51, 52.
R103	Scale includes food insecurity items based on the CCHIP model. Items 15, 18, 19, 20, 21, 22, 23, 53a, 55a, 56a, 58a.	Scale includes food insufficiency and hunger items based on the CCHIP model. Items 16, 17, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 54a, 57a.
R104	NA	Scale includes items that reference events that occurred in the last 30 days. When no 30-day reference was available, items that referenced the last 12-month period are included. Items 15, 17, 18, 19, 20, 21, 22, 23, 26, 27, 30, 31, 33, 34, 36, 37, 39, 41, 42, 45, 46, 48, 49, 51, 52, 53a, 54a, 55a, 56a, 57a, 58a.
R105	NA	Scale includes items that referenced 30-day period and number of days in the last month. Also includes items that reference "often true" in the last 12 months. Items 17, 26, 27, 30, 31, 33, 34, 36, 37, 39, 41, 42, 45, 46, 48, 49, 51, 52, 53a, 54a, 55a, 56a, 57a, 58a.

NOTES:

(1) For items that referenced number of days, one dummy code was created based on whether the behavior or experience occurred five or more times in the last month.

(2) For items that referenced number of months, one dummy code was created by combining the two more extreme categories of the variable, indicating the experience occurred in three or more of the past 12 months.

(3) For items Q53 through Q58, ‘a' denotes a dummy code that represents ‘often true,' whereas ‘b' denotes a dummy code that combines ‘sometimes true' and ‘often true.'

Exhibit 2-5

SUMMARY OF RESULTS FROM ALTERNATIVE NON-LINEAR FACTOR ANALYSIS MODELS

Model	12-Month Scale		30-Day Scale	
	Poorly Fitting Items	**Redundant Items**	**Poorly Fitting Items**	**Redundant Items**
R101	Q21, Q18, Q15, Q22	Q54b	Q17	No redundant items
R102	Q15, Q16, Q20	No redundant items	Q17	No redundant items
R103	No poor fitting items.	No redundant items	Q16, Q17, Q43	Q26
R104	NA	NA	Q22, Q23	Q33
R105	NA	NA	Q58a, Q17	No redundant items

A general summary of item fits for the alternative models is presented in Exhibit 2-5. The identification of poorly-fitting items and/or redundant items is based on item in-fit and out-fit statistics. The out-fit statistic, μ_i, is an unweighted fit statistic. It is based on a standardized residual, written as:

$$z_{ni} = \frac{y_{ni}}{W_{ni}^{.5}}$$

where y_{ni} is the score residual for household n on item i, and W_{ni} is the variance of the score residual. The standardized residual is then squared and averaged to obtain a mean estimate of item fit.

$$\mu_i = \frac{\Sigma z_{ni}^2}{N}$$

The in-fit statistic, ν_i, is a weighted fit statistic that includes the same squared standardized residual as μ_i, and is written as:

$$v_i = \frac{\Sigma W_{ni} z_{ni}^2}{W_{ni}}$$

Both the in-fit and out-fit statistics have an expected value of 1.0. As they deviate from 1.0, the associated items become candidates for removal from the scale. Generally speaking, a mean square fit statistic that is greater than 1.20 indicates a poor fitting item, whereas a mean square fit statistic that is less than .80 indicates an item is redundant with other similar types of items in the scale. ***Items that have both an in-fit and out-fit statistic above 1.2 are targeted for removal from the scale***. Items with both in-fit and out-fit statistics below .80 are redundant with respect to the information they share with other items in the scale. Items that were shown to be redundant items were also considered for removal and/or combined with other items. Below we focus on describing the results of the 12-month and 30-day scale for M102, because these two specific models were subsequently considered the most parsimonious by the study team.

12-Month Food Security Scale

As with the linear factor analysis models, all Rasch models were initially tested using only households with children, because they comprised the subsample of households that were administered the entire set of food security items. The results for Model M102 are presented in Exhibit 2-6. The summary table contains a large amount of information, briefly described below.

The order of items in the table is determined by their ***item calibration***, shown in the fourth column of Exhibit 2-6. A question's item calibration represents the point on the scale at which there is a 50 percent probability that any given household will respond "yes" to the item. That is, households with higher values on the scale than a particular item's calibration score have a greater than 50 percent probability of answering that item positively; households with lower values have a less than 50 percent probability of a positive response to the item in question. The items are listed from high calibration at the top of the table to low calibration at the bottom.

The item calibration is a function of (1) the total number of individuals that have responded to any item in the scale (1,687); (2) the number of individuals that responded to the particular item in the scale (*n*); and (3) the number of positive responses to the particular item (raw score). For example, item 50 refers to the item "child did not eat for a whole day." The item has an item calibration of 4.56, which is the highest in the table. This event occurs rarely

Exhibit 2-6

SUMMARY OF MODEL R102A

Item	*n*	Raw Score	Item Calibration	Real SE	In-fit		Out-fit		Point Biserial Corr.
					Mean Sq	Z	Mean Sq	Z	
Q50	1,684	12	4.81	.30	.99	0.0	.28	-0.4	.19
Q44	1,684	23	4.01	.22	1.00	0.0	.41	-0.5	.24
Q43	1,684	38	3.36	.18	1.04	0.3	1.73	0.5	.28
Q29	1,683	62	2.68	.14	.89	-1.1	.28	-1.3	.39
Q40	1,683	86	2.21	.13	1.01	0.1	1.99	1.2	.40
Q47	1,684	89	2.15	.12	.88	-1.5	.56	-0.8	.44
Q38	1,683	91	2.12	.13	1.07	0.8	.46	-1.1	.40
Q28	1,684	95	2.06	.12	.95	-0.6	.41	-1.3	.44
Q35	1,685	212	.65	.09	.91	-1.6	.83	-0.6	.57
Q57	1,680	246	.36	.09	1.00	0.1	.60	-1.8	.57
Q25	1,683	293	-.01	.08	.94	-1.3	.56	-2.4	.61
Q32	1,683	442	-.98	.07	.94	-1.5	.67	-2.7	.64
Q24	1,685	449	-1.01	.07	.86	-3.5	.67	-2.8	.67
Q56	1,679	466	-1.12	.07	1.04	0.9	.75	-2.1	.61
Q20	1,686	480	-1.19	.08	1.24	5.5	1.50	3.5	.52
Q58	1,680	671	-2.18	.07	.99	-0.4	.96	-0.4	.60
Q55	1,678	706	-2.36	.07	.87	-3.6	.68	-3.5	.64
Q54	1,679	785	-2.73	.06	.82	-5.2	.74	-2.5	.64
Q16	1,687	795	-2.77	.07	1.23	5.9	1.22	1.9	.50
Q53	1,680	1,066	-4.01	.06	.95	-1.6	.85	-0.8	.49
Q15	1,686	1,469	-6.06	.09	1.31	6.7	7.70	5.5	.10
Mean	1,683	408	.00	.11	1.00	-0.1	1.14	-0.6	
SD	2	382	2.82	.06	.13	2.9	1.53	2.1	

NOTE:

Sample includes households with children only. Items are ordered on terms of severity.

in any household. For this specific subsample, this event occurred for only 12 of the 1,684 households that responded to the item. At the other end of the scale, item 15 ("run short of money and try to make food or food money go further") is the least severe item included in the analysis. The item has the low calibration of -5.74, based on 1,469 positive responses out of the 1,686 households that answered the question.

The column headed "Real SE" shows the standard error of the items, which can be used to create a confidence interval for the item calibration. Items located at the severe end of the scale tend to have the largest standard errors, because they tend to have larger variances compared to items throughout the center and less-severe end of the distribution.

For the 12-month scale presented in Exhibit 2-6, there are three items with both in-fit and out-fit statistics that exceed 1.20 (Q15, Q16, and Q20). Therefore, these three items were removed from the scale, and the model re-estimated. The results of the revised model are presented in Exhibit 2-7. The effective sample size for the revised model is reduced (n = 1,276) because two of the least severe items were removed from the analysis. This results in fewer subjects who have responded yes to any particular item.

For the revised model, there are no items with both in-fit and out-fit statistics that exceed 1.20. Similarly, there are no items with both in-fit and out-fit statistics below .80. Some of the out-fit statistics were small, due primarily to dependencies in some item pairs. For example, item 29 has a low out-fit statistic (mean square = .36), but the item is associated with item 28. We examined several alternative models with these items modeled as trichotomies rather than the multiple dichotomies, but the basic results of the models did not change.

Final 12-Month Food Security Scale

The analyses for the 12-month scale were replicated on subsequent subsamples of the data set.[6] The model replications provided clear support for the invariance of the primary measurement model across subsamples, as well as across different types of households. In each replication, the item calibrations gave identical or near-identical rankings of item severity and

[6] The overall sample was initially divided into four random subsamples. Initial model estimation was carried out for households with children within one subsample. Tests for invariance were performed for households with children in the other three random subsamples. Invariance tests were also performed for households without children, subdividing them into households with any elderly members (age 60 or over) and households with no elderly members.

Exhibit 2-7

SUMMARY OF REVISED MODEL R102A

Item	*n*	Raw Score	Item Calibration	Real SE	In-fit		Out-fit		Point Biserial Corr.
					Mean Sq	Z	Mean Sq	Z	
Q50	1,275	12	4.38	.30	.96	-0.2	.32	-0.5	.21
Q44	1,275	23	3.59	.22	.99	-0.1	.50	-0.5	.25
Q43	1,275	38	2.93	.18	1.01	0.1	1.50	0.5	.29
Q29	1,274	62	2.26	.14	.90	-1.0	.36	-1.4	.40
Q40	1,274	86	1.77	.13	1.02	0.2	2.34	2.0	.39
Q47	1,275	89	1.72	.12	.88	-1.4	.70	-0.7	.45
Q38	1,274	91	1.69	.13	1.09	1.1	.65	-0.8	.39
Q28	1,275	95	1.63	.12	.96	-0.5	.52	-1.3	.44
Q35	1,276	212	.21	.09	.95	-0.9	1.09	0.4	.55
Q57	1,274	246	-.11	.09	.99	-0.2	.65	-2.1	.56
Q25	1,274	293	-.49	.08	.98	-0.4	.76	-1.6	.57
Q32	1,274	442	-1.53	.08	1.01	0.2	.99	-0.1	.57
Q24	1,276	449	-1.56	.08	.96	-1.0	1.01	0.1	.59
Q56	1,273	466	-1.68	.08	1.08	1.9	.97	-0.3	.54
Q58	1,274	671	-2.89	.08	1.11	2.6	1.28	2.1	.47
Q55	1,272	706	-3.09	.07	.94	-1.7	.84	-1.2	.53
Q54	1,273	785	-3.54	.07	.92	-2.2	.94	-0.4	.49
Q53	1,274	1,066	-5.28	.09	1.16	3.7	1.28	0.7	.23
Mean	1,274	324	.00	.12	.99	0.0	.93	-0.3	
SD	1	303	2.70	.06	.07	1.5	.46	1.1	

NOTE:

Sample includes households with children only. Items are ordered in terms of severity.

consistent clustering of closely-ranked items. Applying models fit on separate subsamples yielded household values that correlated at the .99 level.[7]

The final model estimates are based upon all households in the analysis sample; these are presented in Exhibit 2-8. Of the 18,370 households that passed the screener and responded to at least half of the questions applicable to them, there were 7,897 households in which the respondent answered "yes" to at least one of the 12-month scale items. The ordering of the items in the final model changes slightly relative to the ordering of the items described in Exhibit 2-7; however, these minor fluctuations in item severities are expected with different random subsamples of households.[8]

Exhibit 2-9 shows the frequency distribution for the number of responses to items in the survey. The two most frequent response patterns are 10 items and 18 items.[9] The response pattern of 10 items applies largely to the households without children, because these had the opportunity to respond to a maximum of 10 items. The response pattern of 18 items applies to households with children, who had an opportunity to respond to 18 items. These two response patterns account for 98.8 percent of the households, indicating a very low incidence of item nonresponse (1.2 percent of all respondents). Households, whether with or without children, that responded to less than half the items administered had their household score set to "missing."

The central function of the Rasch model is to assign to each responding household a value on the food security scale. The household scale value is fundamentally based on a count of the number of affirmative responses to questions included in the scale. At its simplest, if all households respond to the same set of questions, the household scale value is a constant

[7] In this procedure, we separately fit the model to each subpopulation, such as households with children, households with no children but with elderly members, and households with neither children nor elderly. Each of the separate models was then used to compute scale values for all households in the full sample. The values computed with the different models were then compared through plotting and correlation analysis.

[8] The Rasch model software initially assigns scale values in a range that yields a mean of zero. Because the presence of positive and negative values in the scale can be confusing or misleading, it is conventional to transform the values into a range such as 0-1, 0-10, or 0-100. Values of the 12-month scale presented in other reports from this project transform the original scale values to range from 0.0 to 10.0. The original value is multiplied by .8333 and added to 5.071 to obtain the transformed value. All respondents giving zero affirmative responses are assigned a value of zero, and respondents answering all questions affirmatively get a value of 10.0.

[9] Over half of all households in the sample were higher-income households that did not pass the screening questions, and therefore were not asked any of the questions included in the scales.

Exhibit 2-8

SUMMARY OF FINAL 12-MONTH SCALE

Item	*n*	Raw Score	Item Calibration	Real SE	In-fit		Out-fit		Point Biserial Corr.	Trans-formed Item Calibra-tion[a]
					Mean Sq	Z	Mean Sq	Z		
Q50	4,333	29	4.92	.20	1.09	0.5	6.02	1.8	.18	9.2
Q44	4,331	87	3.48	.12	.84	-1.8	.28	-1.6	.34	8.0
Q43	4,332	135	2.86	.10	.88	-1.7	.78	-0.5	.37	7.5
Q29	7,889	332	2.55	.06	.89	-2.5	.55	-1.8	.35	7.2
Q47	4,333	257	1.88	.07	.93	-1.3	.97	-0.1	.44	6.6
Q28	7,892	537	1.82	.05	.97	-1.0	1.16	0.8	.39	6.6
Q40	4,333	290	1.69	.07	1.01	0.3	1.28	1.0	.44	6.5
Q38	7,861	625	1.54	.05	1.10	3.1	1.31	1.6	.39	6.4
Q35	7,883	1,249	.27	.04	.91	-4.0	.77	-2.6	.54	5.3
Q57	4,324	779	-.15	.05	1.07	2.3	.86	-1.4	.53	5.0
Q25	7,879	1,919	-.70	.03	.93	-3.4	.76	-4.6	.58	4.5
Q32	7,885	2,661	-1.56	.03	.94	-3.5	.94	-1.5	.57	3.8
Q56	4,325	1,453	-1.64	.04	1.08	3.4	.94	-1.0	.54	3.7
Q24	7,893	2,824	-1.72	.03	.88	-7.3	.87	-3.2	.59	3.6
Q58	4,324	2,295	-3.10	.04	1.14	6.5	1.29	3.3	.43	2.5
Q55	7,862	4,627	-3.42	.03	1.03	2.1	1.61	7.9	.41	2.2
Q54	7,863	4,973	-3.73	.03	.92	-5.9	1.06	0.8	.42	2.0
Q53	7,870	6,312	-4.99	.03	1.16	9.9	3.04	9.4	.18	0.9
Mean	6,301	1,744	.00	.06	.99	-0.2	1.36	0.5		
SD	1,763	1,833	2.71	.04	.10	4.2	1.26	3.5		

[a] The transformed item calibration is a linear transform of the item calibration that places all values in the range from 0.0 to 10.0.

arithmetic transformation of the count of positive responses. For example, among households with children responding to all 18 questions in the scale, all households with three positive responses have a scale value of -4.13. Households with more affirmative responses have higher scale values; for example, households with children giving ten affirmative responses have a scale value of 0.62. The scale value does not depend on which questions the household answers

Exhibit 2-9

NUMBER OF QUESTIONS ANSWERED: QUESTIONS IN THE 12-MONTH SCALE

Number of Questions Answered	Frequency	Percent	Cumulative Frequency	Cumulative Percent
2	7	0.0	7	0.0
3	4	0.0	11	0.1
4	6	0.0	17	0.1
5	11	0.1	28	0.2
6	14	0.1	42	0.2
7	53	0.3	95	0.5
8	11	0.1	106	0.6
9	51	0.3	157	0.9
10	10293	55.9	10450	56.8
12	21	0.1	10471	56.9
13	2	0.0	10473	56.9
14	2	0.0	10475	56.9
15	3	0.0	10478	56.9
16	11	0.1	10489	57.0
17	29	0.2	10518	57.1
18	7888	42.9	18406[a]	100.0

[a] Households that answered fewer than half of the applicable questions are excluded from the main analysis, reducing the sample to 18,370.

affirmatively: all households with children who give three affirmative answers have the same scale value, even if they give affirmative answers to quite different questions.

If all respondents are given exactly the same set of questions, the scale value depends solely on the number of affirmative responses. If different respondents answer different sets of questions, however, scale values depend on the severity (as indicated by the item calibration) of the questions that the respondent answers. In the current situation, households with children are asked 18 questions, whereas those without children are asked only ten. Moreover, the questions asked only of households with children are disproportionately the more severe questions. The Rasch model takes these differences into account, assigning values to both types of household

that are comparable even though they responded to different types of questions. Similarly, the model adjusts the scale values assigned to households with or without children that failed to respond to one or more of the items applicable to them.

The frequency distribution of household values on the 12-month scale is presented in Exhibit 2-10. Household values for the 12-month scale range from −6.08 to 5.91 in the original model estimation (values transformed to a 0-10 range are also shown). Most households in the analysis sample responded "no" to all items in the scale, and received a scale value of −6.08 (10,276 households).[10] All other households responded "yes" to at least one item. Their assigned scale value is a non-linear transformation of the total number of items to which they responded affirmatively. If all households had responded to all 18 items, there would be 19 possible scale score values that could be assigned to households. Because households without children could respond to only 10 items, however, there are a number of additional scale scores that can be assigned to households based on a missing data adjustment that is part of the Rasch measurement model. The small proportion of households in either group that failed to respond to one or more questions also received distinct measure scores, depending on the number of items missed.

Final 30-Day Food Security Scale

The 30-day scale was developed in the same manner as the 12-month scale, though there were fewer 30-day items available for analysis. The 30-day scale also has a larger number of item dependencies than the 12-month scale. The results of the final Rasch model for the 30-day scale are presented in Exhibit 2-11. The 30-day scale includes 17 items, and the estimated item calibrations range from −4.37 to 4.00. For the most severe item (item 52), only five households responded affirmatively.

Exhibit 2-12 shows the number of responses households made to the 30-day items administered in the survey. Similar to the 12-month scale, there were two major response categories: 9 (households without children) and 17 (households with children). These two response patterns account for 99.3 percent of households. Here also, households that did not

[10] For analyses involving the full sample, households that did not pass the screen are assigned the minimum possible score (−6.08). This procedure is also used in classifying households on the food security status variables.

Exhibit 2-10

FREQUENCY DISTRIBUTION FOR HOUSEHOLD VALUES ON THE 12-MONTH SCALE

Value on Scale	Frequency	Percent	Cumulative Frequency	Cumulative Percent	Transformed Scale Value[a]
-6.08	10276	56.5	10276	56.5	0.0
-5.2	970	5.3	11246	61.9	0.7
-4.96	902	5.0	12148	66.8	0.9
-4.13	661	3.6	12809	70.5	1.6
-3.73	614	3.4	13423	73.8	2.0
-3.36	550	3.0	13973	76.9	2.3
-2.73	657	3.6	14630	80.5	2.8
-2.69	386	2.1	15016	82.6	2.8
-2.09	343	1.9	15359	84.5	3.3
-1.82	306	1.7	15665	86.2	3.6
-1.52	358	2.0	16023	88.1	3.8
-0.97	255	1.4	16278	89.5	4.3
-0.96	285	1.6	16563	91.1	4.3
-0.43	188	1.0	16751	92.1	4.7
-0.09	295	1.6	17046	93.8	5.0
0.1	176	1.0	17222	94.7	5.2
0.62	132	0.7	17354	95.5	5.6
0.81	231	1.3	17585	96.7	5.8
1.13	86	0.5	17671	97.2	6.0
1.62	59	0.3	17730	97.5	6.4
1.75	128	0.7	17858	98.2	6.5
2.12	59	0.3	17917	98.6	6.8
2.65	28	0.2	17945	98.7	7.3
2.88	85	0.5	18030	99.2	7.5
3.24	15	0.1	18045	99.3	7.8
3.77	103	0.6	18148	99.8	8.2
3.96	12	0.1	18160	99.9	8.4
5.02	13	0.1	18173	100.0	9.3
5.91	6	0.0	18179[b]	100.0	10.0

[a] The transformed scale value is a linear transform that places all values in the range from 0.0 to 10.0.

[b] Includes only households that responded to all applicable items.

Exhibit 2-11

SUMMARY OF FINAL 30-DAY SCALE

Item	*n*	Raw Score	Item Calibration	Real SE	In-fit		Out-fit		Point Biserial Corr.
					Mean Sq	Z	Mean Sq	Z	
Q52	990	5	4.00	.45	.83	-0.4	.22	-0.7	.23
Q51	990	13	2.91	.30	1.07	0.3	1.04	0.0	.20
Q46	988	21	2.33	.23	.92	-.4	.68	-0.5	.34
Q31	1992	83	1.61	.12	.83	-1.9	.27	-3.2	.34
Q49	990	45	1.37	.16	.80	-1.7	.44	-1.7	.47
Q42	990	64	.91	.14	.88	-1.1	.59	-1.5	.46
Q45	988	69	.80	.14	1.10	1.0	1.67	1.8	.32
Q37	1985	249	.10	.08	.84	-3.3	.51	-4.1	.46
Q48	990	129	-.09	.11	1.03	0.4	1.07	0.4	.40
Q30	1992	294	-.17	.07	1.08	1.8	1.22	1.6	.34
Q41	990	154	-.37	.11	1.14	2.1	1.42	2.3	.34
Q39	1958	344	-.48	.07	1.18	4.0	1.42	3.4	.29
Q34	1983	611	-1.52	.06	.92	-2.6	.73	-4.7	.46
Q36	1985	637	-1.61	.06	.94	-1.9	.91	-1.4	.44
Q27	1993	715	-1.86	.06	1.04	1.2	.96	-0.8	.37
Q33	1983	1285	-3.56	.05	.97	-1.3	.87	-1.5	.29
Q26	1993	1549	-4.37	.06	1.13	4.3	1.54	3.3	.14
Mean	1516	369	.00	.13	.98	0.0	.92	-0.4	
SD	497	444	2.12	.10	.12	2.1	.43	2.3	

respond to at least half the items administered had their scale value set to "missing."

Exhibit 2-13 provides the frequency distribution of the 30-day household scale scores. The scale scores range from −5.62 to 5.32. Almost 90 percent of the households that passed the series of screening questions responded "no" to all items in the 30-day scale.

The 30-day scale in its present form is not considered as useful as the 12-month scale, for both conceptual and statistical reasons. Conceptually, the 30-day scale provides detail on a narrower portion of the spectrum of food insecurity than the 12-month scale. Most of the less-

Exhibit 2-12

NUMBER OF QUESTIONS ANSWERED: QUESTIONS IN THE 30-DAY SCALE

Number of Responses	Frequency	Percent	Cumulative Frequency	Cumulative Percent
2	7	0.0	7	0.0
3	1	0.0	8	0.0
4	6	0.0	14	0.1
5	2	0.0	16	0.1
6	10	0.1	26	0.1
7	17	0.1	43	0.2
8	35	0.2	78	0.4
9	10369	56.3	10447	56.8
10	1	0.0	10448	56.8
11	2	0.0	10450	56.8
13	1	0.0	10451	56.8
15	16	0.1	10467	56.9
16	15	0.1	10482	57.0
17	7922	43.0	18404	100.0

severe conditions and behaviors incorporated in the 12-month scale were not measured in the 30-day time frame in the CPS Supplement. The 30-day measures thus focus on reductions of food intake and related indicators of hunger, providing little information on food insecurity with no hunger evident. The broader range of the 12-month scale makes it likely to be more useful both in describing the conditions of the population at a point in time and in monitoring changes.

Statistically, Chapter Three will show that the 30-day scale is considerably less reliable than the 12-month scale in its ability to discriminate between households at varying levels of food insecurity. This more limited reliability stems mainly from the smaller number of independent questions asked in the 30-day time frame. The 30-day scale has just nine

Exhibit 2-13

FREQUENCY DISTRIBUTION OF HOUSEHOLD VALUES ON THE 30-DAY SCALE

Scale Values	Frequency	Percent	Cumulative Frequency	Cumulative Percent
-5.62	16309	89.2	16309	89.2
-4.69	261	1.4	16570	90.6
-4.63	288	1.6	16858	92.2
-3.5	239	1.3	17097	93.5
-3.39	246	1.3	17343	94.8
-2.66	123	0.7	17466	95.5
-2.45	96	0.5	17562	96.0
-2.01	113	0.6	17675	96.7
-1.68	144	0.8	17819	97.5
-1.47	67	0.4	17886	97.8
-1	57	0.3	17943	98.1
-0.97	69	0.4	18012	98.5
-0.56	34	0.2	18046	98.7
-0.25	59	0.3	18105	99.0
-0.14	25	0.1	18130	99.2
0.27	23	0.1	18153	99.3
0.57	47	0.3	18200	99.5
0.68	9	0.0	18209	99.6
1.11	5	0.0	18214	99.6
1.57	5	0.0	18219	99.6
1.7	24	0.1	18243	99.8
2.08	4	0.0	18247	99.8
2.62	31	0.2	18278	100.0
2.66	2	0.0	18280	100.0
3.39	3	0.0	18283	100.0
4.44	1	0.0	18284	100.0
5.32	1	0.0	18285	100.0

independent items, and a total of 17 when follow-up items are included.[11] The 12-month scale has 15 independent questions, plus three follow-up items. In addition, the absence of questions measuring the less severe food insecurity conditions creates a situation in which an extremely small proportion of the population gives affirmative responses to any of the items, which makes it more difficult for the scale to discriminate reliably among different levels of food insecurity.

For these reasons, the main report of this study focuses almost exclusively on the 12-month scale, and this report provides less detail on the 30-day than the 12-month scale. Estimates of the prevalence of hunger based on the 30-day scale are presented in Appendix B.

2.4 SUMMARY

The scale development process involved five main steps:

- Exploratory linear factor analysis replicating key elements of prior research, which indicated that the response patterns and relationships in the CPS Food Security Supplement were largely similar to those seen previously.

- Estimation of two-parameter non-linear models, which indicated that a one-factor solution would be appropriate.

- Preliminary estimation of one-factor Rasch models on a one-fourth random subsample of the full CPS sample, resulting in the specification of an 18-item set for inclusion in the 12-month scale and a 17-item set for the 30-day scale.

- Tests of invariance of the model across other random subsamples of the full population and across three demographic subgroups (households with children, households without children but with elderly members, and households with neither children nor elderly members), which indicated that the models were quite invariant across groups.

- Estimation of the final scales on the full CPS sample.

Subsequent chapters of this report detail the steps taken to test the scales for reliability, construct validity, and estimation bias. Primary attention is given to the 12-month scale, which appears more useful than the 30-day scale on both conceptual and statistical grounds.

[11] The primary question typically asks if a particular behavior or condition occurred in the past 30 days. If the response is affirmative, the follow-up question then asks on how many of the 30 days the behavior or condition occurred.

CHAPTER THREE

RELIABILITY ESTIMATES FOR THE FOOD SECURITY SCALES

Whenever an instrument is used to measure some quality of a person — whether it be a heart rate, a psychological profile, or a level of food insecurity — researchers want to be assured that the instrument is reliable. A reliable instrument is one that, if it were administered to the same individual on two occasions under similar conditions, would provide similar results in both tests. Reliability indices therefore attempt to measure the degree to which an individual's score is expected to remain stable (relative to other individuals' scores) over repeated occasions using the same instrument.

Often it is not feasible to administer an instrument repeatedly to the same individuals under similar circumstances. Reliability indices have therefore been developed that attempt to approximate this result through a single administration of the instrument. Most reliability indices for multi-item scales attempt to provide an estimate of the ratio of the true score variance to the total variance for a particular instrument. The underlying concept is that an individual's score on a scale (x) is composed of the individual's "true" score (t) and an error component. A general equation for a measure indicating the reliability of a scale (ρ) can be written as:

$$\rho = \frac{\sigma_t^2}{\sigma_x^2},$$

where σ_t^2 is the variance of the households' true scores and σ_x^2 is the variance of the observed measure (i.e., the household scores on the scale).

There are a number of reliability indices available for characterizing the reliability of a measure. Because the food security scales are estimated using a Rasch modeling approach, the most appropriate index is the Rasch reliability index. Because the Rasch reliability index has not been used as often in the scale development literature as some other reliability estimators, however, we provide estimates using some of the more common reliability indices as well as the Rasch reliability index to characterize the reliability of the food security scale.

One major difference between the more traditional reliability indices and the Rasch reliability index is the treatment of cases with extreme scores. Cases with extreme scores are

those with either the maximum or minimum score possible on the measure (i.e., those that have responded affirmatively to all questions in the scale, or negatively to all questions). When scale scores are normally distributed over a population, very few cases have extreme scores and consequently they have very little impact on the reliability estimate. When the distribution is severely skewed, however, the treatment of cases with extreme scores can have a major impact on reliability estimates. This is very relevant to the food security scales, because over 80 percent of the population has the lowest possible score on the 12-month scale and over 90 percent on the 30-day scale.

Because of differences in estimation algorithms, the Rasch reliability estimate always decreases when extreme scores are included, whereas the more traditional reliability estimates always increase. The Rasch model typically provides two reliability estimates, one including and one excluding the cases with extreme scores. The conventional practice with the more traditional reliability indices is to include the extreme scores. The discussion below provides separate reliability estimates that include and exclude extreme scores. In general, the estimate excluding households with extreme scores can be taken as indicating the reliability of the scale in measuring the severity of food insecurity and hunger among households that have experienced at least one of the food insecurity or hunger conditions represented in the scale. The interpretation of the estimate when extreme scores are included is less clear.

Among the more traditional indices, Nunnally (1978) recommended that at least two types of reliability coefficients be reported: correlations between alternate test forms, and coefficient alpha. The discussion below presents the results using three traditional reliability indices, two of which are based on the correlation between alternate test forms (the Spearman-Brown split-half reliability estimate, and Rulon's split-half reliability estimate), and Cronbach's alpha. All three reliability indices are based on the use of linear composites, and therefore do not correspond exactly to the Rasch model (a non-linear model). Nonetheless, the indices provide a general indication of the reliability of the scale and familiar measures that may be compared to other work.

3.1 SPEARMAN-BROWN SPLIT-HALF RELIABILITY ESTIMATES

The general form of the Spearman-Brown prophecy formula can be written as:

$$\rho_{sp} = \frac{k\rho_{ii'}}{1+(k-1)\rho_{ii'}},$$

where ρ_{sp} represents the reliability of the composite measure with k parallel tests, and $\rho_{ii'}$ represents the reliability of any one particular test. A simplified form of the equation can be written as:

$$\rho_{sp} = \frac{2\rho_{ab}}{1+\rho_{ab}},$$

where ρ_{ab} represents the correlation coefficient between two parallel tests.

In order to create two somewhat parallel tests, the item pool (i.e., all the items used in the scale) is typically split in half randomly. Each subset of the items is considered a separate scale, and the results of the two scales are compared. When the number of available items is small, as in the present situation, a commonly used method is to order the items in terms of severity and assign odd-numbered items to one test and even-numbered items to another test. The two new scales should have the same number of items, so if the item pool contains an odd number of items, one is dropped before the pool is split.

To estimate ρ_{sp} for the 12-month scale, it was necessary to drop dependent items in order to generate unbiased reliability estimates.[1] It was also considered informative to generate reliability estimates separately for items that were administered only to households with children and for items that were administered to all households.

For households with children, there were 15 independent items available to create two parallel measures. Because there were an odd number of items, the most severe item was dropped from the list. For the first parallel scale, households' responses to items 43, 28, 38, 57, 56, 58, and 54 were summed to create the household score. For the second parallel scale, items 47, 40, 35, 32, 24, 55, and 53 were summed. Based on the correlation between household scores

[1] Dependent items are those that are follow-ups to previous items. A number of items in the food insecurity scales have an initial question (e.g., did this situation occur within the past 12 months?) and a follow-up (e.g., in how many of the past 12 months did the situation occur?)

on these two scales, the Spearman-Brown reliability estimate for the total scale was .852 with extreme scores excluded (see Exhibit 3-1). Including extreme scores raises the reliability index to .903.

Exhibit 3-1

SUMMARY OF RELIABILITY ESTIMATES USING TRADITIONAL INDICES

Household Type	Reliability Estimate	Extreme Scores Included	Extreme Scores Excluded
12-Month Scale			
All households	Spearman	.899	.794
	Rulon	.932	.878
	Alpha	.856	.743
Households with children	Spearman	.903	.852
	Rulon	.899	.813
	Alpha	.882	.814
30-Day Scale			
All households	Spearman	.840	.357
	Rulon	.888	.650
	Alpha	.789	.356
Households with children	Spearman	.852	.530
	Rulon	.844	.530
	Alpha	.799	.555

For all household types (i.e., households with any combination of either children, adults, and elderly), there were eight independent items available to create two parallel measures. For the first parallel scale, items 28, 35, 24, and 54 were summed. For the second parallel scale, items 38, 32, 55, and 53 were summed. The reliability estimate for the total scale is .794 with extreme scores excluded, and .899 with extreme scores included.

For the 30-day scale, the reliability estimate for households with children is .530 and the reliability estimate for all households is .357 with extreme scores excluded. Including extreme scores generates a striking increase in the reliability estimates, to .852 for households with children and .840 for all households.

Note that, although including cases with extreme scores increases the reliability estimate for both scales, the effect is particularly striking for the 30-day scale. This occurs for three reasons.

First, the number of items in the paired subscales is smaller for the 30-day scale. The 30-day scale contains just five independent items that apply to all households, and ten that apply to households with children. This means that the split-half scales each contain just two items in the analysis for all households, and five in the analysis of households with children. In contrast, the split-half 12-month scales contain four items for the analysis of all households and seven items for the analysis of households with children. Smaller numbers of items in general lead to lower reliability estimates.

The second factor is that the 30-day scale measures a narrower band of the spectrum of food insecurity than the 12-month scale. The least severe items in the 12-month scale were not asked in the 30-day time frame. This means that the 30-day scale not only contains fewer items, but that the scale is attempting to make distinctions within a narrower range than the 12-month scale. In effect, this means that the 30-day scale faces a more difficult challenge in distinguishing the varying levels of food insecurity and hunger among those households that have experienced one or more of the conditions measured.

The final distinction between the scales is that a far greater proportion of households answered negatively to all items on the 30-day scale than the 12-month scale (89 percent *vs.* 57 percent of households that passed the screening questions). Thus, including or excluding the households with extreme scores will have a greater effect on the 30-day than the 12-month scales.

3.2 RULON'S SPLIT-HALF RELIABILITY ESTIMATES

Rulon proposed an alternative method for estimating the reliability of a scale using the split-half tests.[2] The method involves estimating the difference between household scores on two parallel tests and estimating the ratio of the variance of the difference score to the variance of the total score. The equation for Rulon's method is written as:

[2] Rulon, P.J., "A Simplified Procedure for Determining the Reliability of a Test by Split Halves," *Harvard Educational Review* vol. 9, pp. 99-103, 1939.

$$\rho_{sh} = 1 - \left(\frac{\sigma_D^2}{\sigma_x^2} \right),$$

where σ_D^2 is the variance of the difference score and σ_x^2 is the variance of the total score.

To estimate the index, we used the same subsets of items described above for the Spearman test, again performing the computation both for households with children and for all households (see Exhibit 3-1). For the 12-month scale, the reliability estimate for households with children is .813 and the estimate for all households is .878 with extreme scores excluded. When extreme scores are included, the estimates increase to .899 for households with children and .932 for all households.

For the 30-day scale, the reliability estimate for households with children is .530 and the reliability estimate for all household types is .650 when extreme scores are excluded. Including the extreme scores raises the estimates to .844 and .888, respectively.

3.3 CRONBACH'S ALPHA RELIABILITY ESTIMATES

Cronbach's alpha and Kuder Richardson 20 (McDonald, 1985) produce identical results when using independent items that are dichotomous in form. Therefore, for the 12-month scale, these two equations are interchangeable. For simplicity, we will refer to Cronbach's alpha when describing these reliability estimates.

Cronbach's alpha was developed to circumvent problems associated with the non-random selection of subsets of items when using methods such as the Spearman-Brown or Rulon methods. Cronbach's alpha, α_{xx}, can be written as:

$$\alpha_{xx} = \left[\frac{k}{(k-1)} \right] \cdot \left[1 - \left(\frac{\Sigma \sigma_i^2}{\sigma_x^2} \right) \right],$$

where k represents the number of items in the test, σ_i^2 represents the variance of item i, and σ_x^2 represents the variance of the total test score. Alpha is considered to be the lower bound of the true theoretical reliability estimate, the coefficient of precision.

The overall reliability estimates, summarized in Exhibit 3-1, are similar to those seen with the prior tests. With extreme scores excluded, the values of α for the 12-month scale are .814 for households with children and .743 for all households. Including the households with extreme scores raises the estimates to .882 for households with children and .856 for all households.

For the 30-day scale, the α values are .555 for households with children and .356 for all households when cases with extreme values are excluded. When households with extreme values are included, the values are .799 for households with children and .789 for all households.

In addition to assessing the reliability of the total scale, Cronbach's alpha is often used to examine the appropriateness of including individual items in the scale. The usual rule is that if α_{xx} increases substantially when an item is removed from the scale, the item should be considered for removal. It is also possible to evaluate how the reliability of the scale changes when any one item is removed from the scale. Exhibits 3-2 and 3-3 show that in nearly all instances, removing an item would reduce the estimated reliability of the scale. The only potential exception would be item 53;[3] removing this item would generate a small increase in the reliability estimate with extreme scores excluded, but the loss of information at the end of the scale would be more detrimental to scale validity than is justified by this small increase in reliability.

3.4 RASCH MODEL RELIABILITY ESTIMATES

The Rasch reliability indices behave in a slightly different manner and yield somewhat lower estimates of reliability than the more traditional indices presented above. The reliability index for the Rasch Scale is defined as:

$$\rho_r = \frac{(\sigma_x^2 - MSE)}{\sigma_x^2}$$

where ρ_r is the reliability index, σ_x^2 is the variance of the scale, and MSE is the mean square error of the scale. Like the previously described reliability indices, ρ_r is intended to represent

[3] Removing item 28 with extreme scores included also generates an increase in α, but the difference is tiny (measured in the third decimal).

Exhibit 3-2

CRONBACH'S ALPHA FOR THE 12-MONTH SCALE FOR HOUSEHOLDS WITH CHILDREN

Item	Extreme Scores Included (α = .882; *n*=7,888)			Extreme Scores Excluded (α = .814; *n*=4,278)		
	Item Mean	Correlation with Total Score	α with Item Deleted	Item Mean	Correlation with Total Score	α with Item Deleted
43	.017	.338	.882	.028	.309	.812
47	.033	.433	.879	.057	.415	.806
28	.036	.397	.880	.063	.354	.809
40	.037	.433	.879	.064	.408	.806
38	.040	.429	.879	.071	.394	.806
35	.081	.565	.873	.146	.529	.796
57	.098	.587	.872	.177	.540	.795
32	.179	.669	.867	.327	.567	.791
56	.183	.664	.867	.333	.556	.793
24	.182	.642	.868	.332	.522	.796
58	.288	.656	.868	.528	.441	.804
55	.290	.709	.865	.532	.528	.796
54	.338	.692	.866	.621	.462	.801
53	.450	.607	.873	.827	.221	.818

the proportion of total variance in household scores that is caused by variance in households "true" scores.

In Exhibit 3-4, the reliability estimates for the 12-month and 30-day scale are presented. Separate estimates are presented for two treatments of the variables that involve follow-up questions. For example, the 12-month scale includes an item that indicates that adults have cut or skipped meals in the past 12 months, and a second (answered only by people who responded positively to the first item) that indicates that meals were cut or skipped in three or more months. In one treatment, these are considered as independent dichotomous items. In the second treatment, they are combined into a single trichotomous item (no meals cut/skipped in past 12 months; meals cut/skipped in one or two months; meals cut/skipped in three or more months).

Exhibit 3-3

CRONBACH'S ALPHA FOR THE 12-MONTH SCALE FOR ALL HOUSEHOLDS

Item	Extreme Scores Included (α = .856; *n*=18,179)			Extreme Scores Excluded (α = .743; *n*=7902)		
	Item Mean	Correlation with Total Score	α with Item Deleted	Item Mean	Correlation with Total Score	α with Item Deleted
28	.034	.434	.858	.080	.429	.727
38	.040	.459	.855	.092	.451	.723
35	.072	.594	.842	.167	.582	.695
32	.149	.701	.827	.343	.595	.686
24	.157	.682	.829	.362	.545	.697
55	.257	.678	.830	.591	.373	.736
54	.276	.725	.823	.635	.439	.721
53	.349	.646	.837	.803	.206	.760

Exhibit 3-4

RASCH RELIABILITY ESTIMATES FOR THE 12-MONTH AND 30-DAY SCALES

Scale	Model Type	Including Households with Extreme Scores	Excluding Households with Extreme Scores
12-month scale	Dichotomous	.63	.74
	Trichotomous	.58	.70
30-day scale	Dichotomous	.00	.57
	Trichotomous	.00	.44

Treating such question sets as trichotomous items reduces the number of items in the scale, and hence reduces the estimated reliability.

With extreme scores excluded, the reliability estimates for the 12-month scale are .74 (dichotomous) and .70 (trichotomous). The reliability estimates for the 30-day scale are .57 and .44.

Unlike the previous reliability indicators, the Rasch reliability estimate decreases when extreme scores are included. Thus, the reliability estimates for the 12-month scale are .63 and

.58 with the extreme scores included. For the 30-day scale, because 88 percent of the households that passed the screener responded negatively to all questions, the reliability estimate falls to zero when cases with extreme scores are included.

3.5 RELIABILITY IN IDENTIFYING CASES WITH NO FOOD INSECURITY CONDITIONS

As noted earlier, none of the reliability statistics deal adequately with situations in which a large percentage of cases have extreme scores. For present purposes, then, the statistics are primarily useful in indicating the scales' reliability in distinguishing the level of food insecurity among households that experience at least one of the conditions measured by items included in the scales. The statistics provide little information about the scales' reliability in distinguishing between households that experience none of the food insecurity conditions measured and households that experience one or more of the conditions.

To provide additional insight on this point, a further analysis was conducted. The analysis follows the split-half procedure: for each scale, we separate the items into two groups to constitute two new scales; we then examine the relationship between the two new scales. The scales are split as described earlier, but each of the new scales is then collapsed into a dichotomous variable. The two categories on the dichotomous variable are (1) "answered all questions negatively," and (2) "answered one or more questions positively." The agreement between the new dichotomous items is then assessed.

A simple test of correspondence is the percentage of cases classified similarly by the two variables. When the population is unevenly divided between the two categories of the dichotomous variables, however, a high rate of agreement can occur by chance. The more appropriate test is therefore the Kappa statistic. The Kappa statistic is a measure of the extent to which there is agreement above and beyond what would be expected by chance. Kappa (κ) is computed as:

$$\frac{(\textit{percent observed agreement}) - (\textit{percent agreement expected by chance alone})}{100\% - (\textit{percent agreement expected by chance alone})}$$

To test the hypothesis H_0: $\kappa = 0$ *vs*. H_1: $\kappa > 0$, we can use the lambda statistic $\lambda = \frac{\kappa}{se(\kappa)}$. A formula for the estimation of the standard error of κ can be found in Rosner (1986).

Landis and Koch (1977) suggested that a κ below 0.4 represents poor agreement, between 0.4 and 0.75 represents good agreement, and greater than 0.75 represents excellent agreement.

The percent agreement between paired subscales and the Kappa statistics are shown in Exhibit 3-5. As expected, the two scales in each pair are in agreement in a high percentage of cases—around 85 percent for the 12-month scale, and around 95 percent for the 30-day scale. More importantly, the κ values are all close to .70, which is toward the high end of the range representing "good" agreement.[4]

Exhibit 3-5

LEVEL OF AGREEMENT BETWEEN DICHOTOMIZED SPLIT-HALF SCALES

	Households with Children		Households without Children	
	Percent Agreement	κ	**Percent Agreement**	κ
12-month scale	84.8%	.70	85.8%	.69
30-day scale	94.5%	.68	95.1%	.67

This suggests that the scales provide a reasonable level of reliability in distinguishing between households that have experienced any of the measured facets of food insecurity and households that have not experienced any of these conditions. It is particularly worth noting that the κ statistics for the 30-day scale are quite similar to those for the 12-month scale, even though the 30-day subscales have very few items and a very high percentage of respondents answering all questions negatively. These factors appear to reduce the 30-day scale's reliability in discriminating among households that have experienced one or more of the measured conditions, but the scale remains reasonably strong at distinguishing those that have experienced any of the conditions from those that have not.

[4] In all of the comparisons, the λ statistic indicates that the level of agreement is significantly greater than would be expected by chance ($p < .001$).

3.6 SUMMARY

Although there is no absolute rule regarding minimum acceptable levels of reliability, the literature provides at least some rough guidelines. Nunnally (1978), writing in the context of the more traditional measures of reliability, suggests that reliabilities of about .70 can be sufficient to suggest general reliability, particularly in the early stages of measurement development. Nunnally suggests that for basic research, requiring a very high reliability (e.g., above .80) can be counterproductive, as resources are devoted to improving the scale instead of learning about the underlying phenomenon. He also argues, however, that scales used to support decisions regarding the treatment of specific individuals should have reliabilities exceeding .90.

Using the three traditional measures and following the conventional practice of including households with extreme scores, both the 12-month scale and the 30-day scale would be judged quite reliable. Estimated reliability values range from .86 to .93 for the 12-month scale, and from .79 to .89 for the 30-day scale.

As noted previously, however, this conventional approach yields statistics that can be influenced by the type of highly-skewed distributions that characterize the food insecurity scales. A more conservative approach is to separate two types of reliability. The first considers the scale's reliability in describing the level of food insecurity among households that experience one or more of the food insecurity or hunger conditions measured by items in the scale. The second asks about the scale's reliability in distinguishing between households that have *vs.* have not experienced any of the measured food insecurity or hunger conditions.

The 12-month scale fares quite well on both dimensions of reliability. When households that answered all questions negatively are excluded from the analysis, the Rasch reliability estimate ranges from .70 to .74, and the more traditional indices range from .74 to .88. Using the dichotomous split-half test, the κ statistics are .69 to .70. Although this approach is novel, and no established benchmarks provide standards for "good" reliability, all of these scores are in the acceptable range for other uses of the statistics.

The 30-day scale is equally reliable at distinguishing households that have *vs.* have not experienced any of the measured food insecurity and hunger conditions. The κ statistics of .67 to .68 are nearly the same as those for the 12-month scale. The 30-day scale, however, seems less reliable at distinguishing among levels of food insecurity for households that experience one

or more of the measured conditions. When we consider only the households that answered at least one question affirmatively, reliability estimates range from .36 to .65.

Two factors reduce the 30-day scale's estimated reliability in distinguishing levels of food insecurity and hunger among households that experience one or more of the measured conditions. First, the number of independent items on the 30-day scale is small. Second, the 30-day scale measures a narrower range of food insecurity, because some of the less severe questions were not asked in the 30-day time frame. To increase the reliability of the 30-day scale to be more comparable to the 12-month scale, it would probably be necessary to add more 30-day items to the Food Security Survey, and in particular to add items measuring less severe conditions of food insecurity than those currently included in the scale.

CHAPTER FOUR
DEFINING RANGES OF THE FOOD SECURITY SCALE

The analyses discussed in earlier chapters provide the basis for concluding that food security can be reliably measured as a unidimensional phenomenon. Households can be ranked on the basis of scale values across a continuous range indicating the severity of food insecurity experienced within the household. The full range of severity measured extends from no measurable food insecurity at all, through increasing levels of severity characterized by reduced food intake and hunger for household members, to some maximum measured level. Although the phenomenon of food insecurity can be viewed as unidimensional and continuous, several distinct ranges of severity are of interest. Identifying these ranges of severity enables one to supplement the continuous food security scale, subdividing it to create a categorical variable providing a comparatively simple measure of food security status in terms of several broad ranges of severity.

In this chapter we describe the conceptual and empirical bases for *a priori* expectations regarding the structure of a categorical food security status variable, and the process leading to definition of categorical ranges within the continuous food security scale. Several specific issues related to selection of threshold levels or scale dividing lines are summarized, and the final categorical food security status variable is described.

4.1 CONCEPTUAL BASIS FOR A CATEGORICAL FOOD SECURITY STATUS VARIABLE

The first threshold level of severity, or dividing line, to be identified on the unidimensional food security scale is the point of transition from food secure status to food insecure status. In addition to this threshold, two other cutpoints, deriving from the LSRO/AIN conceptual definitions of food security, food insecurity, and hunger, are of interest.[1] As noted in the main

[1] The conceptual rationale underlying the measurement of food insecurity and hunger developed in the present study is described in Bickel, Andrews and Klein (1996). The research background leading to this measurement approach is documented in the U.S. Department of Agriculture report, *Food Security Measurement and Research Conference: Papers and Proceedings*, Alexandria, VA: USDA Food and Consumer Service, Office of Analysis and Evaluation, June 1995.

report of this study,[2] the LSRO/AIN conceptual clarification provides a working definition of hunger as "the uneasy or painful sensation caused by a lack of food" and identifies hunger as "a potential but not necessary consequence of food insecurity" (Anderson/LSRO, 1990). Previous studies examined by the AIN expert group had led to a consensus view of hunger as "nested" within the broader phenomenon of food insecurity, and occurring at the more severe levels of food insecurity as experienced in U.S. households.

Moreover, empirical evidence supports the conceptual view of household-level food insecurity as a managed process involving identifiable patterns or stages of behavioral responses to food insufficiency as the degree of such insufficiency increases (Radimer, Olson and Campbell, 1990; Basiotis, 1992; Cristofar and Basiotis, 1992; Radimer *et al*., 1992; Wehler, Scott and Anderson, 1992; Burt, 1993; Cohen, Burt and Schulte, 1993). Within this framework, food insecurity in the household begins with an initial stage characterized by adult household members' experiences of food insufficiency, anxiety about their food situation, and adjustments in their budget and food management patterns. These latter behavioral "coping strategies" may involve efforts to augment the household's food supply from emergency or other non-normal sources, and may involve modifications to the variety and quality of food available to household members, but normally do not include reduction in overall quantity of food intake. In this initial stage there is little or no evidence that household members experience actual hunger — "the uneasy or painful sensation caused by a lack of food" — as a result of their household's level of food insecurity.

The second stage involves intensification of food economizing behaviors, some of which lead to patterns of reduced food intake among one or more of the adults in the household. When children are present in a household, efforts are made to spare them from food intake reduction through various rationing strategies. If the household's food insecurity persists or worsens, however, a third stage appears in which adult hunger is manifested in more severe forms (e.g., going whole days with no food) and, in households with children, the children experience actual hunger, revealed in patterns of reduced food intake.

[2] Hamilton *et al.* (1997), *Household Food Security in the United States in 1995: Summary Report of the Food Security Measurement Project*, Alexandria, VA: U.S. Department of Agriculture, Food and Consumer Service, June 1997, Chapters One and Two.

This conceptual framework suggests four potentially identifiable stages or levels of severity within the continuous food security variable. Those severity-level categories are: (1) Food Secure; (2) Food Insecure with No Hunger Evident; (3) Food Insecure with Moderate (adult) Hunger Evident; and (4) Food Insecure with Severe Hunger (child hunger, and severe adult hunger) Evident. Given these conceptual categories, the question is how best to subdivide the 12-month and 30-day scales into ranges of severity that correspond operationally to the designated conceptual categories.

4.2 DEFINING RANGES AND SELECTING SCALE CUTPOINTS

As described in earlier chapters, the Rasch model assigns a scale value to each household based on the number of scale items answered affirmatively relative to the total number of items answered.[3] As an interdependent part of its estimation from the data, the model also ranks scale items according to their level of severity on the basis of the actual response patterns of all households in the data. The 18 items in the final 12-month scale are shown in Exhibit 4-1, with items listed by increasing order of severity from top to bottom in the table. If all responses were perfectly ordered, an affirmative response to any scale item would occur only in conjunction with affirmative responses to all prior, or less severe, scale items. Therefore, as perfect scale ordering is approached among the actual sample households, any number "*n*" of affirmative responses approaches exact correspondence to the first *n* items in the scale. Although the data are not perfectly ordered for all households, in fact the most common pattern of household responses (the mode) does follow the sequential order of severity.[4] That is, the modal household

[3] For ease of explication this discussion is presented without addressing separately the cases of households with and without children. Readers should note that these two types of households were presented different numbers of items, because questions addressing conditions of children in the household were not presented to households without children. The form of the Rasch measurement model and the BIGSTEPS software that implements the model take these differences into account in calculating household scale scores.

[4] For example, among households with no children, 82 percent followed the modal pattern on the 12-month items. Households answering "no" to all questions, however, amount to 65 percent of the total. Among households answering "yes" to at least one question, 49 percent followed the modal pattern. For the non-modal households, responses deviate from the pattern that would be observed under perfect ordering. Some households answer "yes" to items without answering "yes" to all prior items. A non-modal household with *n* affirmatives has answered negatively one or more of the *n* less-severe questions, instead affirming one or more of the more severe questions. The Rasch model implicitly considers them equivalent, in effect treating all households as modal and assigning both households the same scale value.

that answers *n* items affirmatively gives "yes" responses to the *n* least severe items in the scale sequence.

Defining ranges on the continuous scale is the operational means of assigning values to the categorical variable measuring households' food security status. This categorical measure identifies the particular range of severity of food insecurity that a given sample household has experienced in the prior 12-month or 30-day period. Defining the appropriate scale ranges for classifying households according to food security status involves identifying subsets of the sequential indicator items that best correspond to the conceptual categories described above. After a subset is identified in general terms, it is necessary to identify the appropriate classification boundaries, or points of transition from one severity range to the next. Each such boundary is marked by a particular "threshold item." The threshold items and their classification boundaries developed in the present study for the purpose of giving operational definition to the categorical food security status variable are depicted by the shaded rows in Exhibit 4-2.[5]

Thus, the scale itself, with items ranked from least to most severe, provides a meaningful framework within which to identify operationally the designated ranges of behaviors and conditions corresponding to the conceptual construct summarized above. The scale, whose values range from 0 to 10, must be subdivided in terms of numeric values so that a household with a particular scale value can be assigned to a particular food security status category. This subdivision, however, can be accomplished by considering the behaviors and conditions represented by values at each point on the scale.

The procedure for subdividing the scale rests on two features of the scaling methodology described above. First, household values on the food security scale are based fundamentally on a simple count of the number of questions to which they respond affirmatively. Second, most households' responses follow the sequential logic of item severity: a household that says "yes" to a particular question typically says "yes" to all less severe questions as well.

In general, then, one can characterize households that have a particular scale value as having responded affirmatively to a particular group of questions. Exhibit 4-2, which is organized in terms of increasing severity of the questions, illustrates the point. A household that

[5] Exhibits 4-1 and 4-2 in the main report of this study (Hamilton *et al.*, 1997), also illustrate this division of the scaled indicator items into the respective severity-level classes of the categorical food security measure.

Exhibit 4-1

ITEMS IN THE FINAL 12-MONTH SCALE LISTED BY INCREASING SEVERITY LEVEL

Item Label	Item Content (All questions refer to the last 12 months)
Q53	Household members worried whether food would run out before they got money to buy more (sometimes or often).
Q54	Respondent reports that the food they bought just didn't last, and they didn't have money to get more (sometimes or often).
Q55[a]	**Household members couldn't afford to eat balanced meals (sometimes or often).**
Q58	Household relied on a few kinds of low-cost foods to feed children because they were running out of money to buy food (sometimes or often).
Q24	Adults in the household cut the size of meals or skipped meals because there wasn't enough money for food.
Q56	Household couldn't afford to feed children a balanced meal, because they couldn't afford that (sometimes or often).
Q32	Respondent ate less than he/she felt they should because there wasn't enough money to buy food.
Q25[a]	**Adults in the household cut the size of meals or skipped meals because there wasn't enough money for food in at least 3 of the last 12 months.**
Q57	Children were not eating enough because household couldn't afford enough food (sometimes or often).
Q35	Respondent was hungry but didn't eat because couldn't afford enough food.
Q38	Respondent lost weight because there wasn't enough food.
Q40	Adults cut the size of children's meals because there wasn't enough money for food.
Q28[a]	**Adults in household did not eat for a whole day.**
Q47	Children were hungry but household couldn't afford more food.
Q29	Adults in household did not eat for a whole day in at least 3 of the last 12 mos.
Q43	Children skipped meals because there wasn't enough money for food.
Q44	Children skipped meals because there wasn't enough money for food in at least 3 of the last 12 mos.
Q50	Children did not eat for a whole day because there wasn't enough money for food.

[a] Indicates threshold items in the scale. For each designated range of severity comprising the categorical food-security variable, the subset of indicators beginning with the threshold item and continuing through the successively more severe indicators, up to the next identified threshold, serve operationally to define and characterize that designated range.

Exhibit 4-2

THRESHOLD ITEMS DEFINING RANGES OF THE FOOD SECURITY SCALE

Questions (in order of increasing severity)		Households with Children		Households without Children	
		Number of Affirmatives	Modal Household Value	Number of Affirmatives	Modal Household Value
		0	0.0	0	0.0
Q53	Worried food would run out	1	0.1	1	0.9
Q54	Food bought didn't last	2	1.6	2	2.0
Q55	**Adult not eat balanced meals**	**3**	**2.3**	**3**	**2.8**
Q58	Adult fed child few low-cost foods	4	2.8		
Q24	Adult cut size or skipped meals	5	3.3	4	3.6
Q56	Couldn't feed child balanced meals	6	3.8		
Q32	Adult eat less than felt they should	7	4.3	5	4.3
Q25	**Adult cut size or skipped meals, 3+**	**8**	**4.7**	**6**	**5.0**
Q57	Child not eating enough	9	5.2		
Q35	Adult hungry but didn't eat	10	5.6	7	5.8
Q38	Adult lost weight	11	6.0	8	6.5
Q40	Cut size of child's meals	12	6.4		
Q28	**Adult not eat whole day**	**13**	**6.8**	**9**	**7.5**
Q47	Child hungry	14	7.3		
Q29	Adult not eat whole day, 3+ mos.	15	7.8	10	10.0
Q43	Child skipped meal	16	8.4		
Q44	Child skipped meal, 3+ mos.	17	9.3		
Q50	Child not eat for whole day	18	10.0		

gives one affirmative answer most often answers Q53 affirmatively, a household with two affirmatives most often affirms Q53 and Q54, and so on.

For each question, the exhibit shows the number of affirmative responses and the associated scale value for households whose responses follow the sequential logic of item severity. For example, if the most severe question affirmed by a household with children is Q24, that household has also responded affirmatively to the four less severe questions (Q53, Q54, Q55, and Q58) and has a total of five affirmative responses. Its corresponding scale score is 3.3. The

exhibit also shows parallel, but slightly different, values for a similar household without children. Q58 is not applicable to that household. Thus, if the most severe question it affirms is Q24, it will have a total of just four affirmative responses. Because the Rasch model, however, computes a scale value that takes into account the number and severity of the questions the household was asked, the scale value for the household without children (3.6) is quite close to the value for the household with children (3.3).

It is possible to describe any point on the scale in terms of the questions that the "modal" or typical household with that scale value has answered affirmatively. Similarly, one can say that all modal households with values at or above a specified point on the scale have responded affirmatively to ***at least*** the group of questions corresponding to the specified point. For example, all modal households with values at or above 2.3 have responded affirmatively to at least the three least severe questions in the scale (Q53, Q54, Q55). All modal households with values of 4.7 or higher have responded affirmatively at least to Q24 and to all applicable less severe questions.[6]

Thus, although the scale itself is a continuous measure of a single dimension (i.e., the severity level of food insecurity), it can be subdivided by considering the collection of conditions and behaviors associated with particular ranges of scale values. In this manner, the scale and the severity rankings provided by the Rasch model yield a statistical framework for defining conceptually meaningful categories for the food security status variable. Within this statistical framework, however, the exact location of the category boundaries or scale thresholds depends upon informed judgment about how best to interpret the conceptual constructs based upon the LSRO/AIN definitions and the previous empirical research findings on food security and hunger. The next section reviews those judgments and the reasoning behind them.

[6] Non-modal households with a given scale value have, by definition, not responded affirmatively to all of the applicable less severe questions, but instead have responded affirmatively to more severe questions. For example, a non-modal household (with children) with a scale value of 2.3 must have answered three questions affirmatively. Instead of Q53, Q54, and Q55, however — the three least severe questions — the household might have said "yes" to Q53, Q54, and Q58, although saying "no" to Q55.

4.3 EVIDENCE OF FOOD INSECURITY

The LSRO/AIN definitions of food security and food insecurity are:

- ***Food security:*** "Access by all people at all times to enough food for an active healthy life. Food security includes at a minimum: (1) the ready availability of nutritionally adequate and safe foods, and (2) an assured ability to acquire acceptable foods in socially acceptable ways (e.g., without resorting to emergency food supplies, scavenging, stealing, or other coping strategies)" (Anderson/LSRO, 1990, p. 1598).

- ***Food insecurity:*** "Limited or uncertain availability of nutritionally adequate and safe foods or limited or uncertain ability to acquire acceptable foods in socially acceptable ways" (*ibid.*).

Several dimensions or aspects of food security are apparent in these definitions, of which the most central and fundamental is described as "enough food for an active, healthy life" — i.e., a sufficient quantity of acceptable foods to meet the household's basic needs. A number of additional dimensions are also apparent, including the nutritional quality and safety of available foods, the social acceptability of the means of obtaining food, and the household's assurance or certainty of its ability to obtain needed food. These additional dimensions of the broad conceptual definition of food security, however, are not directly captured in the questions incorporated in the food security scale. Rather, the measure focuses on the simple quantitative dimension of "enough" food. The food quality dimension is represented only to the extent that some particular quality of food (in both nutritional and conventional senses) is perceived and understood by households members to be necessary. The scale consists entirely of items indicating either this quantitative or qualitative aspect of food sufficiency, as experienced and understood by the household respondent, in relation to his or her self-perception of basic needs.

Several of the questions included in the CPS Food Security Supplement were intended to capture those aspects of households' food coping behaviors that seek to augment insufficient household food supply through emergency or other non-normal means. These extraordinary coping methods, such as obtaining food from food banks or pantries, borrowing money for food, taking children to others' homes for meals, or getting meals at soup kitchens, have been regarded as good behavioral indicators of a condition of food insecurity or insufficiency within the household, and they may be presumed to reflect the concept of acceptability of sources or means of food-acquisition within U.S. social norms. These food-augmenting coping behavior items in

the CPS data, however, do not factor together with the indicators that are included in the measurement scale. Thus, they represent a dimension of the conceptual definition of food security — the assurance of access to food through socially-acceptable means — that is not represented within the unidimensional measure of severity of food insecurity.[7]

Examining the items in the 12-month scale, shown in severity-ranked order in Exhibits 4-1 and 4-2, the basic question is how many items must be answered affirmatively in order to provide clear evidence of food insecurity as defined above. Item Q53 could be interpreted as indicating uncertainty about the household's access to adequate acceptable food, or the ability to acquire it in socially acceptable ways. By itself, however, this subjective item may be considered to lack face validity as a sufficient indicator of food insecurity. An affirmative response to only this one item was therefore judged by the technical analysis team as insufficient to indicate the threshold level of food insecurity.

Giving affirmative responses to two items (in the modal case, items Q53 and Q54) indicates worry or anxiety about the household's food position, and also initial perceptions of insufficiency of the household's food supply (food bought just didn't last). Although these two items together provide stronger evidence of household food insecurity, they were still judged insufficient to establish unequivocally that severity has reached the threshold level required for the categorical measure of food insecurity. Including item Q55, however, captures not only reports that the household food supply is substandard, but also efforts to cope with this insufficient food supply in ways that, although they may maintain the quantity of food intake, reduce the perceived quality of diets below the level the respondent understands to be needed to maintain "balanced meals."

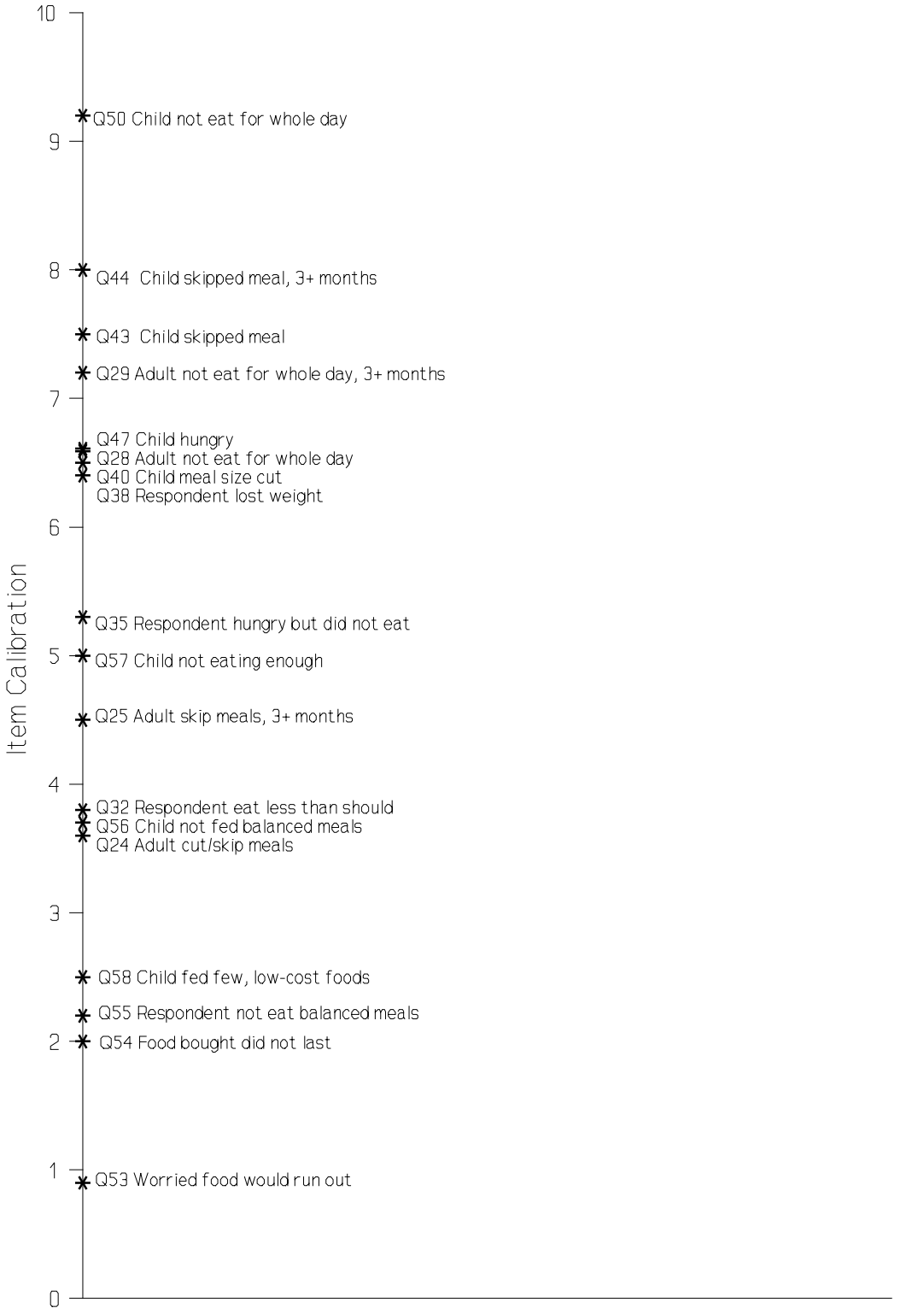
It is useful to consider the relative severity of items as well as the simple rankings shown in prior exhibits. Exhibit 4-3 therefore maps the relative severities, using the item calibrations presented in Chapter Two.

The three least-severe items in the scale (Q53, Q54, and Q55) appear just prior to a substantial gap in the spacing of item calibrations, indicating a large difference in severity between these items and the group comprised by items Q24, Q56, and Q32. Although item Q58 (child fed few low-cost foods) is very close in severity to the item Q55 and consistent in

[7] See Chapter Five for further discussion of these indicators of coping behaviors.

Exhibit 4-3

SEVERITY RANKING OF QUESTIONS IN FOOD SECURITY SCALE



Note: Item calibrations show relative severity of questions from 0.9 (least severe) to 9.2 (most severe).

conceptual content, selection of the threshold or cutpoint item aims at identifying the point of transition from food security into food insecurity. Thus, the first item completing a group that is conceptually and statistically consistent with food insecurity was judged most appropriate for identifying the threshold. Item Q55 meets this criterion, and the set of three household- or adult-level items answered affirmatively by modal households responding "yes" to item Q55, taken together, was judged to provide sufficient evidence that the household has experienced food insecurity, although at a level not yet showing evidence of actual hunger among household members.

4.4 SUBJECTIVE REPORTING OF HUNGER

As summarized above, this research has aimed to develop both a continuous measure of severity and a broad categorical measure of resource-constrained food insecurity that can differentiate three broad ranges of severity, the two most severe of which involve actual hunger for household members. This measurement task is guided by the LSRO/AIN conceptual definitions of food insecurity and hunger, where hunger is nested as "a potential but not necessary consequence" of food insecurity, and is defined as "the uneasy or painful sensation caused by a lack of food." Therefore, an essential measurement task is to identify households whose members have experienced actual hunger — the "uneasy or painful sensation caused by a lack of food" — as a result of constrained or insufficient household financial resources. Food insecurity or hunger resulting from eating disorders, dieting, or causes other than household resource constraints are not being measured.

Three related factors enter into the conceptual consideration of what constitutes the specific phenomenon being measured. These are access to adequate food, the physiological sensation of hunger, and potential malnutrition. The relationships between the first two of these — the basic dimension of food insecurity and hunger as experienced within households — constitute the focus of the present research. The relationship of this basic experiential dimension to malnutrition (which is also defined as nested — a "potential but not necessary consequence" — within food insecurity) is not addressed in this research.

All items in the CPS Food Security Supplement addressing aspects of food insecurity or hunger contain explicit language making it clear to respondents that the condition being asked about is specifically caused by constrained household financial resources. For example, item Q53

states "I/We worried whether (my/our) food would run out before (I/we) got money to buy more." Item Q54 states "The food (I/we) bought just didn't last, and (I/we) didn't have money to get more," whereas item Q55 states "(I/We) couldn't afford to eat balanced meals." Such qualifying language is included consistently in all food insecurity and hunger items in the CPS instrument, including all those appearing in the food security scales. As a result, within the limits of unidentifiable measurement error, affirmative responses to scale items can be expected to reflect clear understanding by respondents that such answers are identifying resource-constrained conditions. Although the possibility of respondents' intentional misreporting exists, as in every survey, the history and nature of the CPS, the high degree of preparedness of CPS interviewers, and the careful design and testing of the Food Security Supplement items all tend to reduce this and other types of measurement error. This point is important because identifying the second classification boundary — the transition from food insecurity with no hunger evident into food insecurity with moderate hunger (adult hunger) evident — relies primarily on evidence that reduced food intake consistent with hunger has occurred within the referenced time period among adults in the household, and that this hunger has resulted specifically from the resource-constrained food insecurity of the household.

The task faced by the analysis team of determining the most appropriate severity level of the initial boundary for the severity range of food insecurity with hunger present involved two kinds of judgment. First, it was necessary to decide which specific items available in the scale should be taken to indicate actual hunger for one or more adults in the household attributable to resource constraint. These potentially include measures of reduced quantities of food intake for adult household members (e.g., Q24, Q25), respondents' subjective assessment of intake adequacy (Q32), or direct perception and report of personal hunger (Q35). Second, given the scale items available, a judgment is required as to how many such items are needed to provide sufficient evidence that household members have experienced actual hunger due to resource constraint. As explained below, the threshold ultimately chosen relies on evidence of a repeated pattern of reductions in food intake by adults over the referenced time period.

The physiological sensation of hunger is experienced universally by all humans, and a large research literature exists examining the nature of the experience in the context of basic

human physiology and clinical nutrition.[8] Several articles from this research literature are summarized in Appendix A of the present volume. The studies described in this literature provide strong support for the validity of subjective reporting of the sensation of hunger (see, for example, Mattes and Friedman, 1993), although they find considerable variation in how the sensation is experienced and described.

These studies seem to provide clear evidence that when usual patterns of eating are interrupted by reducing food intake through actions such as cutting the size of meals or skipping meals, the "uneasy or painful sensation caused by a lack of food" is the natural result. The intensity of the sensations experienced is positively associated with the length of the period of abstinence, although they diminish and may disappear altogether after an extended period of fasting (usually several days). The results reported in this literature are thus consistent with the use of items indicating that reduced food intakes below usual or normal meal patterns, due to resource stringency, are evidence that hunger has been experienced.

Referring to Exhibit 4-3 above, after Q55 the next most severe item to indicate reduction of food intake among adults is item Q24 (Adults cut/skip meals). Note that this item appears in Exhibit 4-3 at virtually the same level as child item Q56 (Child not fed balanced meals), which indicates reduction in the quality of diets provided to children in the household at this level of severity of food insecurity. The next item (Q32, Respondent eat less than should) indicates that food intake has fallen below the respondent's own normative standard for the amount of food he or she should be eating.

An affirmative response to item Q25 indicates that, in addition to all of the foregoing conditions, adults in the household cut the size of or skipped meals in three or more of the previous twelve months due to constrained resources, indicating a pattern of repetition of reduced food intakes among adult household members. This item was judged to provide sufficient additional evidence for the presence of adult hunger in the household, and was chosen, therefore, as the item indicating the point of transition from the category of food insecurity with hunger not evident to the category of food insecurity with adult hunger evident. Households in which the respondent answered affirmatively to item Q25 will, in the modal case, also have answered

[8] See Mattes and Friedman (1993) and Read, French and Cunningham (1994) for two general reviews covering much of this research (see References, Appendix A).

affirmatively to all previous items, indicating the household has experienced a comparatively severe level of food insecurity. The affirmative answer to item Q25 indicates that adults in the household have experienced, in addition, a pattern of repeated reductions in food intakes of a type that the physiological research literature indicates is normally accompanied by the "uneasy or painful sensation caused by a lack of food," or hunger.

When considering the selection or identification of cutpoint items, and when deciding whether affirmative responses to items or sets of items yielded sufficiently clear evidence of a particular condition (e.g., resource-constrained adult hunger), the study team employed a general principle of requiring a pattern of repetition of either behaviors or items, or both. Thus, in considering items indicating reduced food intake among adults, Q25 was viewed as providing sufficient evidence because it involved occurrence of the behavior "cutting or skipping meals" in a recurring pattern over the previous twelve months. Similarly, when considering items indicating the existence of food insecurity with no hunger evident, a pattern of affirmative responses to a sequential series of items was considered stronger evidence than affirmation of only one or two pertinent items. This principle was employed to provide additional assurance against response error.[9]

4.5 EVIDENCE OF CHILD HUNGER AND SEVERE ADULT HUNGER

Exhibit 4-3 shows items Q38, Q40, Q28, and Q47 all grouped at nearly the same level of severity and located at a considerably increased level of severity beyond items Q25, Q57, and Q35. The logic described above for selection of item Q25 as the threshold item for food insecurity with adult hunger evident might suggest item Q40 (size of children's meals cut) as a likely candidate for the best item indicating the transition into food insecurity with severe hunger, because children's hunger is conceptually the most salient aspect of severe hunger in the household. For reasons similar to those outlined above, however, a more severe item was chosen. The wording of item Q40 allows the respondent to answer affirmatively if children in the household had their meal size cut due to resource constraint only once or a small number of times within the previous twelve months. Here again, sufficient evidence of hunger among

[9] Issues of response error are discussed further in Chapter Eight.

children was thought to require either a repetitive pattern of reduced food intake or a multiple series of responses indicating such a condition.

Note that the child items indicating meals being cut and skipping meals occur as two separate items, unlike the adult version, in which these two conditions are combined as one item. The item addressing children skipping meals appears in Exhibit 4-3 at a much higher level of severity than the item regarding size of children's meals being cut. Skipping meals, as would be expected, reflects a more severe condition than cutting the size of meals. In addition, adult items Q38, Q28, and Q29, all of which indicate comparatively severe levels of adult hunger, appear prior to child item Q44, which indicates a pattern of repeatedly skipped meals among children.

These circumstances led team members initially to choose item Q47 (child hungry but couldn't afford more food) as the cutpoint indicating the beginning of food insecurity with child or severe adult hunger evident. Assignment of household food security status using item Q47 as this cutpoint, however, led to anomalous results due to the different numbers of items presented to households with and without children. This anomaly was avoided by choosing item Q28, which appears at virtually the same severity level as item Q47 in Exhibit 4-3, as the cutpoint item indicating the transition from food insecurity with adult hunger evident into food insecurity with child and severe adult hunger evident.

In modal households with children responding affirmatively to item Q28, two items related to reduction of food intake among children receive "yes" answers: item Q57 (children were not eating enough) and item Q40 (children had meal size cut). Moreover, respondents in all household types respond affirmatively to Q35, Q38, and Q28, indicating that adults in the households "were hungry but did not eat because they couldn't afford food," "lost weight because there wasn't enough food," and did "not eat for a whole day because there wasn't enough money for food." Affirmative responses to these items, taken together with affirmative responses to all less severe items, appear to provide clear and strong evidence of child hunger and severe adult hunger.

4.6 SUMMARY

The primary task of the food security measurement study was to identify, test, and develop a unidimensional measure of food insecurity and hunger based on the CPS food security data, if a statistically strong and sound measure of this kind could be found. The Rasch measurement method was successful in producing a unidimensional, continuous-variable measure of severity of food insecurity and hunger from the CPS data that met these requirements. The second task of the project, which was dependent upon the success of the underlying continuous measure, was to develop a categorical-variable measure of several designated ranges of severity of food insecurity, and the classification of households into these designated severity ranges or categories, as follows:

- food secure
- food insecure with hunger not evident
- food insecure with moderate hunger
- food insecure with severe hunger

The conceptual construct for these designated ranges of severity was drawn from the AIN/LSRO conceptual definitions of food insecurity and hunger, from other prior research on food security measurement, and from limiting the measurement effort to one of the central elements of the broad food security concept that is amenable to direct measurement, the direct household experience of insufficient food to meet basic needs. Other elements of the broad conceptual definition, such as safety of food, actual nutritional adequacy of diets, and social acceptability of food acquisition, are not encompassed in the present measure of severity of food insecurity.

The categorical measure of food security status depends on classifying households into identifiable ranges of severity on the underlying continuous severity measure. The aim in identifying or selecting the appropriate ranges of severity on the continuous measure was to achieve acceptably close correspondence to the conceptual bases of the designated broad food security status categories described above. The operational means of establishing the several severity ranges was to select the most appropriate indicator items from among those available in the continuous measurement scale to identify, or define operationally, the classification boundaries, or thresholds, separating each designated severity range category from the next. This

task involved judgment as to which items best reflect the transition from one broad range or category of severity to the next.

Identification of the threshold items and their associated scale cutpoint scores for each level of the categorical food security status variable involved use of statistical results from the Rasch model, guided by the LSRO/AIN conceptual definitions of hunger and the results of previous research in the areas of physiology, clinical nutrition, and food security measurement. Team members combined these factors to select thresholds or cutpoint items that are most consistent with the statistical results, empirical evidence, and the conceptual framework representing the predominant understanding of food insecurity and hunger within the nutrition science community.

CHAPTER FIVE
THE RESOURCE AUGMENTATION QUESTIONS

In fitting the model for the 12-month food security scale, one group of questions was conspicuously not included because they did not meet the statistical criteria for inclusion in the scale. These questions involve actions that households might take to deal with a problem of constrained food resources, and specifically actions other than reducing food intake or otherwise modifying the internal household management of food resources. The questions refer to actions such as putting off other bills in order to buy food, or obtaining meals from soup kitchens. The class of actions has variously been termed “coping” or “resource augmentation” behaviors.

Because resource augmentation behaviors are pertinent to one dimension of the LSRO/AIN definition of food insecurity — the ability to acquire food in “socially acceptable ways” — the research team considered it important to explore the possibility of supplementing the primary food security scale with some composite based on the resource augmentation questions. For example, the food security status variable, rather than simply being based on a subdivision of the primary scale, might also take into account the household's value on the resource augmentation composite. Ultimately it was concluded that, although such a composite might be useful for some researchers in particular situations, it does not add significant value to the food security status variable.

This chapter reviews both the conceptual underpinnings of the effort to construct a composite, the procedures that were implemented, and the likely effect of using a composite such as that described.

5.1 TWO DIMENSIONS OF FOOD INSECURITY

The LSRO/AIN conceptual definition of food insecurity includes several diverse aspects or dimensions of households’ food situations, of which only one central element — the direct experience of insufficient food to meet basic needs — is captured in the measure developed from the CPS food security data.

Households can, however, be food insecure either because they are unable to obtain enough food (for discussion, call this food insecurity "type A"), or because they have to resort

to socially unacceptable ways of obtaining food (call this "type B"). They may also be food insecure for both these reasons. That is, they may resort to socially unacceptable ways of obtaining food and still not obtain access to sufficient food (call this "type A&B").

Because resource-constrained hunger is understood to be nested within food insecurity, it will not occur in a household unless that household is food insecure. If a household is food insecure type A (unable to obtain enough food) at a sufficient level of severity, then hunger may result. Likewise, if a household is food insecure type A&B, hunger may still emerge, despite the household's efforts to augment its available food through various coping measures. If a household's food insecurity is limited to type B only, however, the presence of basic food insufficiency and hunger within the household cannot be inferred from this information. This relationship is illustrated in Exhibit 5-1.

Exhibit 5-1

ILLUSTRATION OF ROLE OF RESOURCE AUGMENTATION BEHAVIORS

Food Availability		Mode of Acquisition		Food Security Status
Sufficient food available	**AND**	Socially acceptable acquisition	☞	Food secure
Limited or uncertain availability (anxiety, adjustments to budget management, adjustments to food quality)	**OR**	Resource augmentation via socially unacceptable means	☞	Food insecure with hunger not evident
Severely limited availability (reduced food intake and other indicators)	. . . . .		☞	Food insecure with evidence of hunger

The availability of sufficient foods to meet basic needs (food insecurity type A). This dimension is well represented in the final unidimensional 12-month scale. As described in the previous chapter, scale development activities demonstrated that it is possible to define a range of values on this scale that can be used to classify households as "food insecure" on the basis of

limited availability of foods relative to household need, operationally indicated by a pattern of anxiety about the adequacy of the household's food supply, and deterioration in the quality and quantity of food available in the household.

The ability to acquire foods in socially acceptable ways, or via normal channels (food insecurity type B). The scale development models employed do not capture this dimension. Using the final 12-month scale to classify households as food insecure leaves open the possibility that some households relying on extraordinary coping methods to acquire food in socially unacceptable ways will be classified as food secure.

This situation emerges because the items in the CPS Food Security data that address this latter dimension of food insecurity do not fit the measurement models leading to the final 12-month scale. Two sets of items ask questions that provide indications of whether households obtained food in ways that might be considered socially unacceptable. One set of items asks whether households undertook actions to augment their food supply or other household resources within the previous 12 months. These items are summarized in Exhibit 5-2.

Exhibit 5-2

RESOURCE AUGMENTATION ITEMS IN THE FOOD SECURITY SURVEY INSTRUMENT

Item Label	Item Summary/Description
Q18	"get food or borrow money for food from family or friends?"
Q19	"send or take children to the homes of friends or relatives for a meal?"
Q21	"put off paying a bill so you would have money to buy food?"
Q22	"get emergency food from a church, food pantry, or food bank?"
Q23	"eat meals at a soup kitchen?"

A second set of items asks whether members of the household obtained food through federal food assistance programs. These programs include food stamps, elderly feeding programs, the child and adult care feeding program, school feeding programs, and WIC. There are two strong arguments, however, for not using these items to classify households as food insecure.

First, participation in such programs may not be considered "socially unacceptable" by many of the participants. There is some evidence to that effect, although this point has not been adequately researched (Trippe and Beebout, 1988; Fraker, 1990; Radimer, Olson and Campbell, 1990; Trippe, Doyle and Asher, 1992; Olson, Frongillo and Kendall, 1995).

Second, there is a problem of logical circularity that could diminish the usefulness of the food insecurity measures for policy considerations. The food insecurity measures are potentially useful in helping policy makers assess the need for government food assistance programs. Including program participation in the food insecurity measures, however, permits the following potentially perverse result: If the government makes programs more available (for example, by increasing the income eligibility threshold for free school lunches, or food stamps), more people will participate and the experienced level of food insecurity would be expected to decline. The measured level, however, may either decline or increase, depending on how the participation indicator interacts with other indicators of the condition. Conversely, if the government cuts back on programs, participation will decline and the effect of the participation indicator may cause the measured level of food insecurity to go down (i.e., the food insecurity problem can be "solved" by taking away the programs). Because of this situation, participation in government food assistance programs was not included in the candidate pool of items for a resource-augmentation index.

For the classification of households as food insecure to be more fully consistent with the LSRO/AIN definitions, there would need to be a way to include information on food acquisition through ways that are not socially acceptable (non-normal channels). An important part of the indicator items used in earlier efforts to develop measures of food insecurity and hunger reflect actions or behaviors undertaken by household food managers to avoid or ameliorate hunger when food or financial resources become scarce. Sometimes referred to as "coping behaviors" or "coping strategies," these behaviors include actions aimed at augmenting the amount of food available to the household, or its financial resources for food, and they can include actions to acquire food in ways that may be considered socially unacceptable, such as those actions shown in Exhibit 5-2.

5.2 THE COMPOSITE RESOURCE AUGMENTATION INDEX

The feasibility of creating a separate resource augmentation scale was first examined by fitting the five items in Exhibit 5-2, along with others, in Rasch models. Reliability estimates for the resulting scales were below acceptable levels, however. This was due partially to the small number of items available for inclusion in the models. Efforts to increase scale reliability by adding items to the model were not successful.

One possible reason for the lack of fit has to do with the widely uneven opportunity for the resource augmentation actions, or coping behavior, across households. For example, in order to obtain food from a church, food pantry, or food bank, households must have access to these facilities. That is, they must live in an area where such services are provided, and be able to get to them. Similarly, with borrowing money or food from family or friends, such social-support relationships must be available to the household before they can employ this coping strategy. With the possible exception of item Q21 ("put off paying a bill so that you would have money to buy food"), none of the resource augmentation items listed in Exhibit 5-2 are necessarily available to all households who might use them if the opportunity were present.

Because the attempt to construct a Rasch scale was not successful, the research team considered instead the creation of a simple composite or index based on the number of resource augmentation questions a household answered affirmatively. The index is derived using the five resource augmentation items in Exhibit 5-2. The proportions of each type of household answering affirmatively to the resource augmentation items, weighted to represent the true population proportions, are shown in Exhibit 5-3.

The items with the largest proportions of affirmative responses are Q18 (get food or borrow money from friends or relatives) and Q21 (put off paying bills to have more money to buy food), items which could be interpreted by some respondents as not indicating behaviors that are socially unacceptable. The research team therefore felt that, if the index were to be used in classifying households as food insecure, such classification should be based on a pattern of at least three affirmative responses. This conforms with the general principle of redundancy (either in items or behaviors), employed in Chapter Four in making decisions about items yielding evidence of food insecurity or hunger.

This principle requires clear evidence of a pattern of repetition of an action (e.g., involuntary reduction of food intake in the case of food insecurity with moderate hunger), or a

Exhibit 5-3

POPULATION WEIGHTED PROPORTIONS OF AFFIRMATIVE RESPONSES TO THE RESOURCE AUGMENTATION QUESTIONS BY HOUSEHOLD TYPE

Item Label	Item Description	Households with Children	Households with Elderly but No Children	Households without Children or Elderly	All Households
Q21	Put off paying bills to have money to buy food	20.8%	4.5%	9.9%	12.5%
Q18	Get food or borrow money from friends or relatives	12.6%	3.1%	8.1%	8.4%
Q22	Get emergency food from church, food pantry or food bank	4.7%	1.6%	2.1%	3.0%
Q19	Send or take children to friends or relatives for a meal	3.3%	N/A	N/A	1.3%
Q23	Eat meals at a soup kitchen	0.4%	0.2%	0.7%	0.5%

pattern of repetition of affirmative responses to different items indicating an action or state (e.g., three out of five resource augmentation items). Although application of this principle requires an exercise of judgment by the study team, it provides an element of assurance against error that might otherwise arise.

5.3 EFFECTS OF USING THE COMPOSITE RESOURCE AUGMENTATION INDEX

A resource augmentation index as described above, and estimates of "type B" food insecurity derived from the index, provide a potential means of broadening the basic categorical measure of food insecurity prevalence to include the dimension of food insecurity involving reliance on non-normal, "emergency," or "socially unacceptable" forms of food acquisition. The practical effect of broadening the reach of the categorical food security measure in this way, however, turns out to be slight. This is because a very large proportion of the households that would be classified as food insecure on the basis of the resource augmentation index are already classified as food insecure by the underlying measurement scale and the classification criteria for the food security status indicator. The number of ***additional*** households that would be classified as food insecure ***solely*** on the basis of "type B" food insecurity, as measured by the resource augmentation index, is quite small.

The resource augmentation index would be used only to classify households as food insecure with no hunger evident, because it only addresses food insecurity type B (described above). By the logic outlined earlier, type B food insecurity alone cannot be taken as evidence of the presence of hunger, and therefore cannot be used to classify a household as food insecure with hunger evident.

Approximately 0.3 percent (rounded) of all households in the (weighted) sample would be classified as Food Secure on the basis of the scale-based measure, but would be classified as Food Insecure with Hunger not Evident on the basis of the resource augmentation composite. This would raise the proportion classified as Food Insecure with Hunger not Evident from 7.8 percent to 8.0 percent of the population, as shown in Exhibit 5-4. In terms of population weighted values, just under one quarter million additional households would be classified as Food Insecure with Hunger not Evident if the resource augmentation index were used in this way.

Exhibit 5-4

EFFECTS OF THE COMPOSITE RESOURCE AUGMENTATION INDEX ON THE NUMBER AND PERCENT OF HOUSEHOLDS CLASSIFIED AS FOOD SECURE AND FOOD INSECURE WITH NO HUNGER EVIDENT

Measure	Food Secure Households		Food Insecure Households with Hunger not Evident	
Twelve-month scale	**Number (thousands)**	**Percent**	**Number (thousands)**	**Percent**
Without the resource augmentation index	88,270	88.1	7,783	7.8
With the resource augmentation index	88,020	87.8	8,029	8.0

5.4 SUMMARY

Food insecurity can occur as a result of households experiencing either, or both, of two dimensions included in the LSRO/AIN conceptual definition of food insecurity. The measurement scale developed in the study addresses one of these dimensions (limited or uncertain availability of enough food to meet basic needs), but does not capture the second (limited or uncertain ability to acquire acceptable foods in socially acceptable ways).

In principle, the resource augmentation questions in the CPS Food Security Supplement might be used to create a composite measure that could take this second dimension into account. The research team concluded, however, that it would not be advisable at this time to incorporate such a measure into the definition of households' food security status. This conclusion was based on two considerations.

First, it is not clear that the existing items in the CPS Supplement constitute a sufficiently strong representation of the construct of food acquisition in ways that are not socially acceptable. The fact that the items did not meet the criteria for construction into a scale suggests that they do not make up a coherent and complete picture of the dimension of concern. Incorporating the resource augmentation index into the food security status indicator could therefore create a source of variability in the status indicator that might reduce the validity of comparisons across groups or over time.

Second, incorporating the resource augmentation index into the food security status indicator would make only a small difference in prevalence estimates, a difference that could be considered within the "noise level" of the estimates. If the effect were large — if it suggested that a very large number of food insecure households were being ignored by omission of the resource augmentation index — it might be worth accepting the consequences of potentially increased variability in the prevalence estimates. With only a small effect, however, the costs of including the index appear to outweigh its benefits.

Resource augmentation or coping behaviors therefore constitute an important area for future research. Better understanding is needed of the array of such behaviors that actually exists, the conditions in which they are taken, and their relationship to the dimension of food security captured in the primary food security scale. With improved understanding, it should be possible to refine and improve the current approach to measuring food security.

CHAPTER SIX

EXTERNAL CONSTRUCT VALIDATION OF THE FOOD SECURITY MEASURES

External construct validation is the process of formally examining the relationships of a scale or construct to other similar or related measures of the construct that were not used in developing the scale being tested. In the present context, this involves comparing estimates of the households' food security scale scores and food security status to other household measures that are believed to be related to food security and that were not used in construction of the food security measures.

There is no definitive measure of food security — no "gold standard" against which these food security measures can be tested. Absent such a gold standard, the best approach is to examine the relationships between the food security measures and other measures that are understood to have a bearing on food security, provided that one can define an unambiguous hypothesis about the bivariate relationship between the validation measure and the food security measures. Following these criteria, the analysis compares the food security measures to household food expenditures, income, income relative to the poverty line, and the household respondents' report of the sufficiency of food eaten in the household.

For the food security scale, a continuous variable, we calculate the coefficient of correlation between the validation items and the households' measure on the food security scale. For the categorical measure of food security status, we examine the percent of households at each level of severity of food insecurity within groups defined by the validation items. Because the food security scales were developed using unweighted data, the construct validation results reported here are also unweighted.

6.1 RELATIONSHIP OF CONSTRUCT VALIDATION ITEMS TO FOOD SECURITY

The LSRO expert panel identified four dimensions of food security that need to be addressed at the household and individual level when measuring food security. Those four dimensions are: the quantity of food intake, the quality of food intake, anxiety about the adequacy of food supply, and social acceptability of the source of food (Anderson/LSRO, 1990).

The CPS Food Security Supplement included questions intended to capture each of these four dimensions, and all dimensions are included in some form in the food security measures developed for this study.[1] The discussion below considers each of the four chosen construct validation items, focusing on the theoretically expected relationship between the construct validation item and food sufficiency, the limitations of the validation item as a measure of food security, and the results of the construct validation effort.

6.2 WEEKLY FOOD EXPENDITURES PER HOUSEHOLD MEMBER

The predominant way for households to acquire food is to purchase it through normal commercial channels with the financial resources available to the household. One would therefore expect to see a relationship between the level of food expenditures and the quantity and quality of food intake, two dimensions of food security. Hence, one item we chose for comparison with the food security measures is food expenditures: specifically, weekly food expenditures per household member.[2] The hypothesis is that weekly food expenditures per household member will be negatively correlated with the level of food insecurity.

The obvious connection between food expenditures and food insecurity might not be as strong as it first appears because there are several conceptual weaknesses of weekly food expenditures as a measure of food sufficiency and some limitations in using the available expenditure data for validation purposes. Conceptually, food expenditures are not an ideal measure of food sufficiency because expenditures do not include food from most in-kind programs (although food stamp purchases are included) or home-grown food, do not reflect differences in costs for food across localities, and even in per-capita form do not perfectly adjust for the specific food needs of a household. Moreover, households that have, on average, sufficient weekly food expenditures may still experience weeks where they do not have financial

[1] Both the 12-month and 30-day scales include items related to the quantity of food intake (e.g., cutting or skipping meals) and a smaller number of items related to quality (e.g., not being able to serve balanced meals). Anxiety is explicitly represented only in the 12-month scale (e.g., being worried that food would run out before more money was available). Social acceptability is not explicitly addressed by any of the items in the continuous scales, but is measured by the resource augmentation items used in constructing the categorical food status variable for the 12-month period.

[2] Respondents were asked a battery of questions regarding food expenditures in the past week as well as "usual" food expenditures. Respondents were instructed to include purchases made with food stamps in their report of expenditures on food.

resources to purchase sufficient amounts of food. One data limitation of the measure is that the food expenditure questions on the survey ask about usual weekly and monthly spending patterns on food, but do not give a specific time period for which usual weekly or monthly spending is reported, and hence may not match the time period over which food security is measured. Also, food expenditures for the entire household are reported by one respondent, but this respondent may not have accurate information on the food expenditures of other household members. Finally, much of the variation in per capita food expenditures, especially at higher expenditure levels, presumably reflects variations in convenience or luxury that would not be relevant to or captured by the food security measures, because the food security measures give all food secure households the same score.

Exhibit 6-1 shows that the correlation coefficients between the food security scales and weekly food expenditures per household are negative, as expected. The values of the coefficients, however, are quite small: −.12 for the 12-month scale and −.07 for the 30-day scale. The cross tabulation of food expenditures with the categorical food security status variable reported in Exhibit 6-2 also shows the expected pattern: the lower the level of food expenditures, the more likely the household is to be in each of the food insecure categories. For the 12-month scale, 1.4 percent of households that report weekly food expenditures less than $20 per person are in the most severe category of food insecurity, whereas only 0.5 percent of households reporting expenditures of more than $40 per person are in this category. More generally, about 21 percent of the low food spending households are in one of the three food insecure categories, whereas only 6.5 percent of the high food spending households are in one of the food insecure categories.

The same pattern is evident for the 30-day scale, although the overall percentage of households classified as food insecure is much smaller than in the 12-month scale. For example, 4.4 percent of the households reporting they spend less than $20 per household member are classified as showing evidence of hunger in the household, whereas a much smaller 1.2 percent of the households reporting spending $40 or more per household member show evidence of hunger.

Exhibit 6-1

CORRELATION COEFFICIENTS FOR FOOD SECURITY SCALE SCORES AND OTHER VARIABLES RELATED TO FOOD SECURITY

	Food Security Measure	Weekly food Expenditures per Household Member	Annual Household Income	Income Relative to the Poverty Line
12-Month Variable				
Food security measure	1.00	−.12	−.32	−.33
Weekly food expenditures per household member	−.12	1.00	.23	.36
Annual household income	−.32	.23	1.00	.89
Income relative to the poverty line	−.33	.36	.89	1.00
30-Day Variable				
Food security measure	1.00	−.07	−.16	−.16
Weekly food expenditures per household member	−.07	1.00	.23	.36
Annual household income	−.16	.23	1.00	.89
Income relative to the poverty line	−.16	.36	.89	1.00

6.3 HOUSEHOLD INCOME

The financial resources of a household are a primary determinant of the level of household food expenditures. This means that the financial resources of a household should be related to the quantity and quality of food intake. Moreover, households' anxieties about the adequacy of their food resources and their likelihood of resorting to non-normal modes of food acquisition are presumed to increase as their financial resources diminish. Hence, household income is related to all four dimensions of food insecurity mentioned earlier. Household income is expected to be negatively correlated with the level of food insecurity. Here, it is measured both as total annual income and as income relative to the federal poverty line for the given household composition.

Despite the logical connection between household income and food insecurity, one would not expect the correlation to be perfect for several reasons. In particular, food assistance programs, which are designed to ameliorate food insecurity, are specifically targeted to

households with low income, which reduces the relationship between food insecurity and income. Additional reasons one would not expect income to be perfectly correlated with food insecurity are that household income does not include all the assets of a household (e.g., savings), is not adjusted for the food and nutritional needs of a household, and does not reflect the competing demands for financial resources within a household. Also, total income for a year may be substantial even though the year includes periods of time where financial resources are tight, such as a period when the primary earner is unemployed. Finally, much of the variation in income occurs at higher income levels where there is little or no corresponding variation in the food security measures.

The income measure on the April 1995 data file also has several shortcomings that make it an imperfect measure of household income. Chief among them are: income is a categorical variable, analytically converted to a pseudo-continuous variable by taking the midpoints of the categories; household income is derived from a question that asks about family rather than household income, although in most cases these units are the same; and the 12-month period over which income is measured does not exactly match the time period for which food security is measured for most of the sample households.[3] A further potential limitation is that the income measured is cash income, and does not capture the value of food stamp benefits or other in-kind food assistance. One might expect that food security would be more closely related to a measure of income incorporating such transfers than to the cash income measure.

As expected, Exhibit 6-1 shows that food insecurity is clearly negatively related to both annual household income and poverty-scaled income. Annual household income and the 12-month food security scale have a correlation coefficient of −.32, whereas income and the 30-day food security scale have a smaller correlation coefficient of −.16. The correlations with poverty-scaled income are almost identical. All of these correlations are stronger than the correlation of the food security scales with food expenditures.

[3] Each CPS sample is divided into eight approximately equal rotation groups, with each group interviewed four consecutive months, dropped out for eight consecutive months, then brought back in for four more consecutive months before being retired. The household income measure is usually from the first month of each four consecutive month spell in the CPS survey. Thus, although food security is measured for the 12 months preceding the April 1995 survey, income is measured for the 12 months preceding: January 1995 for rotation groups four and eight; February 1995 for rotation groups three and seven; March 1995 for rotation groups two and six; and April 1995 for rotation groups one and five.

Exhibit 6-2

RELATIONSHIP OF THE FOOD SECURITY STATUS VARIABLE TO OTHER VARIABLES

	Food Secure	Food Insecure, Hunger not Evident	Food Insecure, Moderate Hunger Evident[a]	Food Insecure, Severe Hunger Evident[b]
12-Month Variable				
Weekly food expenditures per household member				
<$20	79.2%	13.9%	5.6%	1.4%
$20-29	88.0	8.6	2.9	0.6
$30-39	92.4	5.3	2.0	0.3
$40 or more	93.7	4.0	1.9	0.5
Income relative to poverty line				
<50%	59.5	24.2	11.4	4.9
50-100%	69.6	20.1	8.2	1.9
101-185%	82.6	11.9	4.6	0.9
>185%	96.2	2.6	1.0	0.2
Food sufficiency variable (from one-part version of question)				
Often not enough to eat	15.8	29.0	18.4	36.8
Sometimes not enough to eat	21.8	31.5	36.3	10.5
Enough but not always the kinds of food we want do eat	63.6	25.9	9.4	1.0
Enough of the kinds of food we want to eat	95.9	3.4	0.6	0.1
Food sufficiency variable (from two-part version of question)				
Often not enough to eat	14.4	23.2	33.4	29.0
Sometimes not enough to eat	24.0	38.0	31.3	6.6
Enough but not always the kinds of food we want do eat	67.1	25.7	6.3	0.9
Enough of the kinds of food we want to eat	96.7	2.7	0.5	0.1

[a] See notes at end of exhibit.

Exhibit 6-2 *(continued)*

RELATIONSHIP OF THE FOOD SECURITY STATUS VARIABLE TO OTHER VARIABLES

	No Hunger Evident	Food Insecure, Moderate Hunger Evident[a]	Food Insecure, Severe Hunger Evident[b]
30-Day Variable			
Weekly food expenditures per household member			
<$20	95.6%	3.6%	0.8%
$20-29	98.0	1.7	0.3
$30-39	98.7	1.1	0.2
$40 or more	98.8	1.0	0.2
Income relative to poverty line			
<50%	90.1	7.3	2.6
50-100%	94.4	4.3	1.3
101-185%	96.7	2.8	0.4
>185%	99.3	0.6	0.1
Food sufficiency variable (from one-part version of question)			
Often not enough to eat	52.6	18.4	29.0
Sometimes not enough to eat	65.3	28.2	6.5
Enough but not always the kinds of food we want to eat	94.4	5.0	0.6
Enough of the kinds of food we want to eat	99.7	0.3	0.0
Food sufficiency variable (from two-part version of question)			
Often not enough to eat	48.4	30.5	21.1
Sometimes not enough to eat	76.3	19.9	3.8
Enough but not always the kinds of food we want to eat	96.6	3.2	0.3
Enough of the kinds of food we want to eat	99.8	0.2	0.0

[a] Limited to adult hunger at identifiable but moderate levels of severity.

[b] Evidence of children's hunger and severe adult hunger.

Exhibit 6-2 shows that among households whose income is less than half of the federal poverty level, more than 40 percent are classified as having experienced some kind of food insecurity in the past 12 months, and 5 percent fall into the most severe category of food insecurity. In contrast, only 4 percent of the households with annual income above 185 percent of the poverty level are classified as food insecure, and only 0.2 percent are in the most severe category of food insecurity. The patterns for the 30-day food security scale are similar: 10 percent of households with income below 50 percent of the poverty level have experienced hunger in the past 30 days, whereas less than 1 percent of households with income more than 185 percent of the poverty level have such experiences.

6.4 FOOD SUFFICIENCY

The final construct validation item is a single-item household food sufficiency measure that has been used in previous research, appearing in a substantial number of national food use and other types of surveys. Specifically, the respondent was asked which best describes the food eaten in their household: enough of the kinds of food we want to eat; enough, but not always the kinds of food we want to eat; sometimes not enough to eat; or often not enough to eat. This measure focuses directly on two of the four dimensions of food insecurity: the quantity and quality of food intake. Hence, we expect households that report a more severe food insufficiency experience to be classified in a more severe food insecurity category.

The single-item food sufficiency measure does not have a clearly defined time reference, simply asking respondents to characterize the "food eaten in your household." It does not explicitly address two of the dimensions of food insecurity (anxiety and socially unacceptable modes of food acquisition). It has four categories, which could permit a category-by-category comparison with the four-category 12-month measure of food insecurity, but the conceptual underpinnings of the two categorization schemes are not identical.[4] A technical complication with the food insufficiency measure is that it has been applied in the CPS Supplement in two

[4] For example, a respondent answering "not always the kinds of food we want" can be expressing food preferences unrelated to food insufficiency due to inadequate resources. At the more severe levels, the food sufficiency measure distinguishes households on the basis of the frequency with which the situation occurs ("sometimes" *vs.* "often" not enough to eat). The primary basis for distinguishing between the two most severe categories of the food security variable, on the other hand, is whether the experience of hunger is limited to adults in the household *vs.* adults and children both experiencing hunger.

formats: a one-question version and a two-question version.[5] Each version was applied to a different portion of the CPS sample. Both versions were compared to the categorical food security status variables, and the results are reported in Exhibit 6-2.

Both versions of the food sufficiency measure have the expected strong positive relationship with food security status: households that report more severe food insufficiency experiences tend to be classified in a more severe food insecurity category. For the one-part version of the food sufficiency measure, 84 percent of households reporting "often not enough to eat" are classified as food insecure by the 12-month scale, including 37 percent in the most severe category of food insecurity. In contrast, only 4 percent of households reporting "enough of the kinds of food we want to eat" are in any of the food insecure categories, including only 0.1 percent in the most severe category of food insecurity.

The 30-day food security scale shows a similarly strong relationship with the one-part and two-part versions of the food sufficiency question, although the overall prevalence of food insecurity is smaller for the 30-day scale. Very few of the households reporting "enough of the kinds of food we want to eat" are classified in either of the categories evidencing hunger: only 0.2 to 0.3 percent. In contrast, around half of the households reporting "often not enough to eat" are classified as food insecure with evidence of hunger.

6.5 SUMMARY

The results of these analyses are consistent with the view that the food security measures presented in this study constitute valid measures of the underlying constructs of food insecurity and hunger. This provides as much assurance as one can expect at this stage that the measures provide the desired information. By their nature, however, the tests conducted here cannot be conclusive. All of the items used as points of comparison were designed to measure something other than food security. A perfect measure of food security would therefore not be exactly correlated with any of them, but there is no basis for knowing exactly how close the correlation should be. Further validation will be desirable, including additional comparisons of the food security variables to potentially related measures. In particular, the relationship of the present

[5] See question 11a for the single-question version of the food sufficiency item, and questions 11 and 12 for the two-question version in the CPS Food Security Supplement instrument.

measure of the central dimension of food insecurity and hunger as experienced in U.S. households and established measures of the nutritional quality of diets and their health consequences will be an important area of further research.

CHAPTER SEVEN

PROCEDURES FOR CALCULATING STANDARD ERRORS FOR FOOD SECURITY PREVALENCE ESTIMATES

The *Summary Report* presents standard errors for our food security prevalence estimates as an indicator of the degree of uncertainty surrounding reported point estimates due to relying on a sample from the population rather than an entire census of the population.[1] The conventional procedures for estimating the standard errors of estimates, as incorporated in most statistical software packages, are appropriate only for simple random samples from the population. For efficiency reasons, the Current Population Survey (CPS) relies on a complex sampling design that does not result in a simple random sample of households in the U.S. population. Accordingly, it is necessary in the food security analysis to estimate variances by other means. This chapter briefly explains the CPS sampling design and the method used for calculating standard errors of estimates of the prevalence of food insecurity from April 1995 CPS data.

7.1 CPS SAMPLE DESIGN

The sampling design used by the CPS is essentially a two-stage sampling procedure. In the first stage, CPS stratifies groups of counties (Primary Sampling Units, or PSUs) and Metropolitan Statistical Areas (MSAs) according to geographic location, and basic labor force and demographic characteristics of the geographic area (from the most recent decennial census). One PSU from each stratum is selected to represent the stratum in the sample. Because only households in the selected PSUs are eligible to be in the sample, the usual variance estimation formulas developed under the assumption of simple random sampling will underestimate the between-PSU portion of variance to the extent that the chosen PSUs do not capture the variability among all PSUs.

In the second stage of sampling, clusters of households within PSUs are selected to be in the sample. In this case, the usual variance estimation formulas will underestimate the within-PSU portion of variance to the extent that there is homogeneity within households in a cluster.

[1] See Appendix E of the *Summary Report* (Hamilton *et al.*, 1997).

Variance estimates provided by all-purpose statistical packages, such as SAS, assume simple random sampling from the population of interest and equal weighting of each household; hence, they are not appropriate for use with the complex CPS sampling procedure. Below, we describe the variance estimation procedure used for taking into account the between-PSU variance (sampling of MSAs and county groups) and within-PSU variance (sampling of households within PSUs). This procedure was developed in consultation with statisticians from the Census Bureau's CPS Division.

7.2 ADJUSTMENT FACTOR FOR BETWEEN-PSU VARIANCE

We are unable to estimate between-PSU variance directly because information on a household's PSU is withheld from the data files to protect confidentiality of respondents. Instead, we have calculated an adjustment factor to reflect this stage of sampling in our variance estimates.

To estimate the between-PSU variance indirectly, we referred to unpublished components of variance for several available CPS labor force estimates from November 1995, and components of variance for a more extensive set of labor force estimates reported in Train and Cahoon (1978).[2] Because we are primarily interested in household estimates, we considered those characteristics from the Train and Cahoon paper that are usually based on one person per household.[3] The between-PSU variance as a proportion of total variance was about 5 percent for two of the characteristics, and about 9 percent for the third characteristic. The between-PSU variance tended to comprise a somewhat smaller proportion of total variance in the November 1995 CPS estimates than in the 1978 study.[4] Averaging these proportions, we estimated that the between-PSU variance for April supplement estimates is about 6 percent of total variance. This

[2] Train, G. and L. Cahoon, "The Current Population Survey Variances, Inter-Relationships, and Design Effects," *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 1978, p. 443-448.

[3] These characteristics are the number of self-employed in each household, teenage labor force participation, and teenage unemployment.

[4] The only characteristics available from November 1995 are the total number of employed persons, unemployed persons, and labor force participants. The between-PSU variance for these characteristics were compared to the Train and Cahoon (1978) estimates for the same characteristics.

translates into applying a factor of 1.06 to our direct estimates of the within-PSU variance to obtain estimates of total variance.

7.3 ESTIMATION OF WITHIN-PSU VARIANCE

We used the "random groups" method to estimate the within-PSU component of variance.[5] This procedure measures the sensitivity of an estimate to the particular sample drawn from within selected PSUs. That is, the within-PSU component of variance is calculated by finding the variance of estimates obtained with different samples drawn from the same PSUs.

The eight CPS rotation groups provide an ideal mechanism for dividing the CPS sample to estimate the variance among different samples taken from the same PSUs.[6] Each rotation group is an independent sample of households from all the sample PSUs. Thus, we can divide the CPS sample into half-samples with four rotation groups in each. Each rotation group also independently has the full CPS ratio estimation procedure applied, in which the sample weights are adjusted to independent estimates of the civilian non-institutional population of the U.S. Thus, use of rotation groups allows us to reflect the reduction in variance due to application of the ratio estimation procedure to population controls used in the CPS.[7] Exhibit 7-1 shows the allocation of the eight rotation groups for each of the 30 definitions of half-samples used for our calculations.

To estimate the within-PSU component of the variance for an estimate of the total number of households in food security category *i* (e.g., the total number of households in the U.S. that are ***food insecure with moderate (adult) hunger evident***), we used the random groups variance formula (see Hansen, Hurwitz, and Madow, 1953, p. 440), adjusted for half-samples that

[5] For a more detailed description of this method, see Hansen, Hurwitz and Madow, 1953.

[6] Each month a new rotation group is added to the CPS sample and an old rotation group is dropped from the sample. Sample households in a rotation group are surveyed for four consecutive months, then take eight months off, before being surveyed for four more consecutive months and then dropped from the sample; e.g., a household that enters the survey in January 1995 will be interviewed in January, February, March and April 1995, and again in the same months in 1996.

[7] The SUDAAN program is often used to estimate variances in weighted samples. To use the estimation procedure in the SUDAAN variance program, however, it is necessary to know the weights before application of the ratio estimation procedure, as well as the final weights. Because these pre-ratio estimation weights are not available, the SUDAAN program is not applicable in the present instance.

Exhibit 7-1

ROTATION GROUPS IN EACH DEFINITION OF HALF-SAMPLE

Definition of Half-Samples	CPS Rotation Groups in Half-Sample 1	CPS Rotation Groups in Half-Sample 2
1	6, 1, 2, 3	4, 5, 7, 8
2	6, 1, 2, 4	3, 5, 7, 8
3	6, 1, 2, 5	3, 4, 7, 8
4	6, 1, 3, 4	2, 5, 7, 8
5	6, 1, 3, 5	2, 4, 7, 8
6	6, 1, 4, 5	2, 3, 7, 8
7	6, 2, 3, 4	1, 5, 7, 8
8	6, 2, 3, 5	1, 4, 7, 8
9	6, 2, 4, 5	1, 3, 7, 8
10	6, 3, 4, 5	1, 2, 7, 8
11	7, 1, 2, 3	4, 5, 6, 8
12	7, 1, 2, 4	3, 5, 6, 8
13	7, 1, 2, 5	3, 4, 6, 8
14	7, 1, 3, 4	2, 5, 6, 8
15	7, 1, 3, 5	2, 4, 6, 8
16	7, 1, 4, 5	2, 3, 6, 8
17	7, 2, 3, 4	1, 5, 6, 8
18	7, 2, 3, 5	1, 4, 6, 8
19	7, 2, 4, 5	1, 3, 6, 8
20	7, 3, 4, 5	1, 2, 6, 8
21	8, 1, 2, 3	4, 5, 6, 7
22	8, 1, 2, 4	3, 5, 6, 7
23	8, 1, 2, 5	3, 4, 6, 7
24	8, 1, 3, 4	2, 5, 6, 7
25	8, 1, 3, 5	2, 4, 6, 7
26	8, 1, 4, 5	2, 3, 6, 7
27	8, 2, 3, 4	1, 5, 6, 7
28	8, 2, 3, 5	1, 4, 6, 7
29	8, 2, 4, 5	1, 3, 6, 7
30	8, 3, 4, 5	1, 2, 6, 7

are not exactly the same size (see Cochran, 1977, p.139). The variance formula for totals is listed below:

$$2 * \left[\left(X_{i1} - \frac{N_1 * X_i}{N} \right)^2 + \left(X_{i2} - \frac{N_2 * X_i}{N} \right)^2 \right] \qquad (1)$$

where

X_{i1} is the weighted number of households in the first half-sample that are in food security category *i*;

X_{i2} is the weighted number of households in the second half-sample that are in food security category *i*;

X_i is the weighted number of households in the full sample that are in food security category *i*.

N_1 is the weighted number of households in the first half-sample;

N_2 is the weighted number of households in the second half-sample; and

N is the weighted number of households in the full sample.

This yields the estimated within-PSU variance for estimated totals from a single definition of half-samples. We form 30 definitions of half-samples from the eight rotation groups, resulting in 30 estimates of variance. The final within-PSU variance estimate is an average of these 30 estimates.[8]

For estimated proportions, such as the proportion of all households experiencing food insecurity with moderate hunger evident, we use the variance formula for ratios in Cochran (1977, p. 155). The variance of the proportion, X_i/Y, where X_i is the estimated number of households in food security category *i*, and Y is the estimated number of households in the population, is:

$$var\left(\frac{X_i}{Y_i}\right) = \frac{var(X_i)}{Y^2} + \frac{X^2 * var(Y)}{Y^4} - \left(\frac{2 * X_i}{Y}\right) * cov\ (X_i, Y)] \qquad (2)$$

[8] The CPS modified its sample design in the spring of 1994; hence, three rotation groups (6,7,8) were chosen with the old design and five were chosen with the new design. We have chosen our half-samples such that all three rotation groups under the old sample design are never in the same half-sample.

where:

$var(X_i) = (X_1 - X_2)^2$;

$var(Y) = (Y_1 - Y_2)^2$; and

$cov(X_i,Y) = (X_1 - X_2) * (Y_1 - Y_2)$.

This yields the estimated within-PSU variance for proportions from a single definition of half-samples.[9] The average from the 30 definitions of half-samples provides the final estimate of the within-PSU component of variance.[10]

7.4 CALCULATION OF THE STANDARD ERRORS

Finally, the estimated variance is calculated by multiplying the estimated within-PSU variance by the 1.06 between-PSU variance adjustment factor. The standard errors reported in the exhibits in this report are simply the square roots of the variances estimated using the above procedures.

Our calculations indicate that for the entire population of U.S. households (sample size 44,730), the standard errors for households in a particular food security category range from 0.07 percentage points (food insecure with severe hunger) to 0.36 percentage points (food secure). For smaller subgroups of the population, the standard errors tend to be larger. For example, for subgroups with sample sizes between 1,000 and 2,000, the standard errors range from 0.28 and 0.58 percentage points for the food insecure with severe hunger; and the standard errors range from 1.29 to 1.87 percentage points for the food secure estimates.

[9] When calculating the variance of ratios ***where the denominator is the population of interest***, the adjustment for different sized half-samples made for the calculation of the variance of totals is unnecessary, because the differences in sample sizes are already taken into account with the variance of the denominator term and the covariance between the numerator and denominator.

[10] To convert the variances calculated for proportions experiencing food security status *i* into variances for the percent of the population experiencing food security status *i*, multiply the variance for proportions by 10,000 (i.e., 100 squared).

CHAPTER EIGHT

POTENTIAL SOURCES OF BIAS IN PREVALENCE ESTIMATES

One of the main purposes of collecting the CPS food security data and developing food security measures from the data is to estimate the prevalence in the United States of food insecurity and hunger on a consistent basis over time and across population groups. To that end, the continuous-measure food security scales were further developed into the food security status variable, as described in Chapter Four. This chapter addresses the question of whether, assuming that the conceptual and operational definitions of the status variable are acceptable, prevalence estimates based on that categorical variable can be considered unbiased.

Three potential sources of bias are examined here:

- Screening bias, which might result from the fact that the full battery of food security questions was asked of all lower-income, but only some higher-income, households;
- Response bias, which occurs if households systematically paint a too-rosy or too-bleak picture of their circumstances; and
- Random error bias, which can occur when the true prevalences in the population are highly skewed.

It is impossible to present definitive estimates of the bias resulting from any of these potential sources. Such an analysis would require the household classifications produced in this study to be compared to classifications using a separate, authoritative measure of food security, but no such measure exists. The discussions below are therefore largely theoretical and speculative, attempting to provide a perspective on the likelihood of each of the possible types of bias.

The general sense is that each of these three sources may contribute some bias, but that the magnitudes of bias are likely to be small and the biases probably move in counterbalancing directions. Screening can lead only to a downward bias in the estimated prevalence of food insecurity and hunger. Response bias also seems likely to move prevalence estimates downwards. Random error, on the other hand, would probably yield upward bias in prevalence

estimates. The net effect of these countervailing forces cannot be determined with the available data.

8.1 SCREENING BIAS

In order to reduce respondent burden, the full battery of food security questions was applied to higher-income households only if they passed through a set of screening questions. The screen consisted of two main elements. Households were screened out if they had annual household incomes above 185 percent of the federal poverty line ***and*** if they gave no indication of food insufficiency in response to the single-item food sufficiency question and two other screening questions (Q15 and Q16). Households that had incomes below 185 percent of poverty, and higher-income households that gave some indication of food insufficiency on the screening questions, were asked the full battery of questions.[1,2]

In total, about 26,000 higher-income households did not pass through the screen and were not asked the battery of food security questions. This amounts to nearly 60 percent of the full sample of around 45,000.

It is possible that some respondents who were screened out would have responded affirmatively to some of the questions used in the food security scales. To the extent that this occurred, the estimated prevalence of food insecurity and hunger is biased downwards.

Although the data do not offer a direct way to assess the bias, some insight is possible through looking at the two main screening criteria separately. Households with incomes above 185 percent of poverty passed through the screen if they answered "sometimes [or] often not enough to eat" on the food sufficiency indicator. Similarly, households indicating no food insufficiency passed through the screen if they had incomes below 185 percent of poverty. The food insecurity prevalence estimates for these households are shown in Exhibit 8-1.

[1] The food sufficiency question was asked in two forms: a single question with four possible responses (Q11A), and two questions with three and two response categories, respectively (Q11 and Q12). In either formulation, all households that answered "sometimes [or] often not enough to eat" passed through the screen and were asked the main battery of questions.

[2] The complete screener included two additional paths through besides income and the food sufficiency response, but those two had the predominant impact on screening decisions.

Exhibit 8-1

PERCENT OF SAMPLE ULTIMATELY CLASSIFIED AS FOOD INSECURE, BY INCOME AND FOOD SUFFICIENCY MEASURE

	Above 185% of Poverty	Below 185% of Poverty
No food insufficiency indicated on key screening question	Unknown[a]	16.3%[b]
Food insufficiency indicated on key screening question	59.2%[b]	83.8%[b]

[a] The proportion is necessarily close to zero: most households in this category were screened out and not asked the full battery of questions. The status variable classifies all of these households as Food Secure. A few higher-income households passed the screener based on responses to two other questions, indicating potential food insecurity (Q15, Q16).

[b] Cell percentages represent the food security classification of those households described by the row and column headings. The upper-right cell, for instance, shows that among households that indicated no food insufficiency on the key screening question but had incomes below 185 percent of the poverty line, 16.3 percent are classified as Food Insecure. This includes households classified into any of the three food insecure categories on the 12-month status variable. The three categories are: Food Insecure with Hunger not Evident, Food Insecure with Moderate (Adult) Hunger, and Food Insecure with Severe (Child and Severe Adult) Hunger.

As the exhibit indicates, the food sufficiency item is a fairly powerful screen for food insecurity. A majority of households that indicated food insufficiency on this key screening question were subsequently classified as food insecure on the basis of the full battery of questions. Even among those who reported incomes above 185 percent of poverty, nearly 60 percent of those who indicated food insufficiency on this screening question are classified as food insecure on the scale. On the other hand, the food insecurity rate is only about 16 percent for those households who indicated no food insufficiency on the screening question but were given the full battery of questions because their incomes were below 185 percent of the poverty line.

These figures imply that the percentage of screened-out households who would have been classified as food insecure had they received the full battery of questions is probably very low, but probably not zero. It is therefore likely that the screening procedure imparts a small downward bias to the estimated prevalence of food insecurity.

8.2 RESPONSE BIAS

In assessing response bias, the concern is that survey respondents may, on average, respond to some or all questions in the instrument in a way that systematically portrays the household as more food insecure or less food insecure than its true condition.

Response bias can be accurately judged only by comparing survey responses or scale values to a more definitive measure of the condition. No such definitive measure is available in the present study, however.

The only available indications of bias come from researchers who have been involved in previous efforts to develop measures of food insecurity and hunger. This evidence, largely impressionistic and anecdotal, suggests that survey respondents have some tendency to portray their condition as ***less severe*** than reality, but little tendency to exaggerate their problem. The principal motivations are perceived to be:

- ***Pride/shame*** — a desire not to reveal to an outsider (the interviewer) a condition that the respondent believes could reduce the dignity of his or her image;

- ***Fear of government intervention*** — particularly among low-income households with children, a fear that children might be removed from the household;

- ***Things could be worse*** — households who have experienced more severe conditions than at present may consider their current situation not to be a problem. This is believed to be especially likely for elderly persons who recall the depression of the 1930s;

- ***Reduced standards*** — persons living at a particular level of food insecurity may perceive it to be normal, and may not answer affirmatively to questions about cutting meal size or eating less than they feel they should. At consistently low levels of food intake, individuals may not experience the physical sensation of hunger that a food secure person would experience with the same level of intake. Elderly persons with diminished appetite may not perceive low food intake levels to be problematic.

The response patterns for the CPS food security data provide very little basis for assessing the likelihood or magnitude of any of these potential sources of downward response bias. One point worth noting, however, is that households with elderly members are estimated to have somewhat lower prevalences of food insecurity and hunger than other types of households. This would be consistent with the possible under-reporting biases mentioned for the

elderly, but this analysis cannot distinguish between this and alternative possible reasons for low prevalence of food insecurity in elderly households.

8.3 RANDOM ERROR IN SURVEY RESPONSES

If all respondents have no inclination to portray their situation as better or worse than their true condition, some questions will still be answered inappropriately by some respondents. This may occur because the respondent is confused or not paying attention. It may occur because the question has a legitimate interpretation in addition to the predominant interpretation intended by the survey designers. It may occur because the interviewer hears or records the answer incorrectly. Such errors have no inherent bias: a respondent who should say "yes" is as likely to say "no" as conversely. For present purposes, we consider these errors random.

If the population is roughly evenly divided with respect to the phenomenon being measured, random response error does not lead to bias in prevalence estimates. That is, the number of households who are truly food secure but erroneously classified as food insecure would be offset by an equal number of truly food insecure households who are erroneously classified as food secure.

If the population distribution is highly skewed, however, random error can result in biased prevalence estimates.[3] An example helps to illustrate the issue. Suppose that 90 percent of the population is truly food secure, and 10 percent is truly food insecure. Suppose further that each group has a 10 percent probability of giving responses that cause households to be misclassified. It follows that 9 percent of the population (.9 x .1) is truly food secure but misclassified as food insecure, whereas just 1 percent of the population (.1 x .1) is truly food insecure but misclassified as food secure. In this example, the estimated prevalence of food insecurity would be 18 percent, which means that it has a substantial upward bias relative to the true prevalence of 10 percent.

The best way to assess the extent of this bias would be to compare the classifications used in the study to an independent and definitive measure of food insecurity. Such an assessment would determine both the sensitivity of the classification (the probability that a truly food insecure household would be correctly identified) and its specificity (the probability that a

[3] See Habicht and Meyers (1991) for a detailed discussion of this issue.

truly food secure household would be correctly classified). In the absence of an independent and definitive measure, any analysis is essentially hypothetical.

Some perspective on the issue can be obtained by considering the nature of the food security scale. Fundamentally, a household's value on the scale of severity of food insecurity, and hence its classification into one of the food security status categories, is determined by the number of questions it answers affirmatively. Exhibit 8-2 shows the distribution of households without children, for whom ten questions in the food security battery are applicable and used in the 12-month scale.[4] Among these households, 63 percent were screened out and an additional 24 percent gave negative answers to all ten questions (first row). The percentages in the remaining categories range from 3.3 percent to 0.3 percent, generally declining as the number of affirmatives increases.

Exhibit 8-2

PERCENT OF SAMPLE HOUSEHOLDS WITHOUT CHILDREN BY NUMBER OF AFFIRMATIVE RESPONSES

Screened out or no affirmatives	87.0%
One	3.3%
Two	2.2%
Three (food insecure with hunger not evident)[a]	2.4%
Four	1.1%
Five	1.0%
Six (food insecure with evidence of moderate (adult) hunger[a]	1.1%
Seven	0.8%
Eight	0.5%
Nine (food insecure with evidence severe (child and severe adult) hunger)[a]	0.3%
Ten	0.4%

[a] Minimum number of affirmatives to be classified into the specified category. Percentages based on unweighted counts.

The exhibit illustrates two important points. First, the response distribution is highly skewed, with the vast bulk of the respondents either screened out or giving no affirmative

[4] For households with children, 18 questions were applicable and used in the 12-month scale.

responses. It is reasonable to assume that the true distribution of food insecurity in the population is similarly skewed.

Second, the greatest potential source of upward bias in the prevalence estimates lies in that portion of the population that either should have been screened out or should have given no positive responses. Assume that the true distribution is close to the observed distribution, such that 87 percent or more of all households should be in these first two categories (first row of the table). A relatively low rate of random response error among this group could misclassify a substantial number of households into other categories in the table.

The distribution shown in the table also sets an upper bound on the possible level of random response error leading to false positives. If absolutely no households should have answered any of the questions affirmatively — i.e., if 100 percent of the households should have been in the first two categories — the implied rate of random response error would be 13 percent. It is obviously unreasonable to believe that no one ever gave an affirmative response correctly. Thus, the true rate of random response error must be well under this upper bound.

A key question in assessing the likely level of random response error is how the error might be distributed across the possible levels of the scale. It seems reasonable to believe that small errors would be more common than large errors. That is, if a household should not have answered any questions affirmatively, it would be more likely to give one erroneous affirmative than two; two erroneous affirmatives would be more likely than three; and so on.

If all households responding in error made errors in just one response, prevalence estimation bias would necessarily be extremely small. Households that should give zero affirmative responses (the largest group of households) would give no more than one. Households that give just one affirmative response are classified as food secure. Thus, no one in the largest group of households would be misclassified into the wrong food security status. The only groups that could bias the prevalence estimates in this situation would be those adjacent to the dividing lines between food security status categories. For example, the first dividing line comes between households with two affirmative responses (classified as food secure) and those with three affirmatives (classified as food insecure without hunger). Some households that should give two affirmative responses might give three, and thus be misclassified as food insecure, and some who should give three affirmatives might give two, and thus be misclassified as food

secure. Because the two adjacent groups are quite similar in size, however, no substantial bias would occur even if the probability of error were extremely high.[5]

Putting these various considerations together, the key question is, what percent of the households who should give zero affirmative responses (the dominant population group) actually give three or more affirmatives, thereby contributing bias to the prevalence estimates? It seems unlikely that this percentage is large, which would mean that the prevalence estimates are not strongly biased. As noted at the outset, however, it is not possible to go beyond this kind of speculative analysis without a separate and more definitive measure of food insecurity and hunger.

8.4 SUMMARY

This analysis considered three possible sources of bias in prevalence estimates. Two of the three factors (screening bias and response bias) seem likely to mean that the estimates understate the extent of food insecurity in the population. One factor (random error) seems likely to work in the opposite direction. Of these three factors, only one — the possible tendency among respondents to underreport the condition being measured — seems capable of producing more than a small bias, and this possible downward response bias is speculative only. Thus, although the probable direction of any net or overall bias may be downward, its actual direction is indeterminant and its magnitude is most likely to be small.

[5] None of the pairs of adjacent groups differ in size by more than 0.2 percentage points. Thus, a 100 percent error rate would yield a bias in the prevalence estimate of just that amount: 0.2 percentage points.

SELECTED BIBLIOGRAPHY

Anderson, S.A. (ed.), "Core Indicators of Nutritional State for Difficult-to-Sample Populations." *Journal of Nutrition* 120 (11S): 1557-1600, 1990. A report prepared by the Life Sciences Research Office, Federation of American Societies for Experimental Biology for the American Institute of Nutrition under Cooperative Agreement No. HPU 880004-01-0, Nutritional Status Indicators of Low-income Populations with the Office of Disease Prevention and Health Promotion Department of Health and Human Services.

Andrich, D., "Scaling Attitude Items Constructed and Scored in the Likert Tradition." *Educational and Psychological Measurement* 38: 655-680, 1978.

Andrich, D., "A Model for Contingency Tables Having an Ordered Response Classification." *Biometrics* 35: 403-415, 1979.

Basiotis, P., "Validity of the Self-Reported Food Sufficiency Status Item in the U.S. Department of Agriculture's Food Consumption Surveys." In *American Council on Consumer Interests 38th Annual Conference: The Proceedings*, ed. V.A. Haldeman. Columbia, MO: [publisher], 1992.

Bickel, G.W., M.S. Andrews, and B.W. Klein, "Measuring Food Security in the United States: A Supplement to the CPS." In D. Hall and M. Stavrianos, eds., *Nutrition and Food Security in the Food Stamp Program*. Alexandria, VA: USDA Food and Consumer Service, January 1996.

Burt, M.R., "Hunger Among the Elderly: Final Report of a National Study on the Extent and Nature of Food Insecurity Among American Seniors." Washington, DC: The Urban Institute, November 1993.

Cohen, B.E., M.R. Burt, and M.M. Schulte, "Hunger and Food Insecurity Among the Elderly." Washington, DC: The Urban Institute, February 1993.

Cristofar, S.P. and P.P. Basiotis, "Dietary Intakes and Selected Characteristics of Women Ages 19-50 Years and Their Children Ages 1-5 Years by Reported Perception of Food Sufficiency." *Journal of Nutrition Education* 24(2), 1992.

Fraker, T.M., *The Effects of Food Stamps on Food Consumption: A Review of the Literature.* Washington, DC: Mathematica Policy Research, October 1990.

Habicht, J.-P. and L. Meyers, "Principles of Effective Surveys of Hunger and Malnutrition in the United States." *Journal of Nutrition* 121, 1991.

Hambleton, R.K. (ed.), *Application of Item Response Theory*. Vancouver: Educational Research Institute of British Columbia, 1983.

Hamilton, W.L. *et al.*, "Food Security in the United States in 1995: Report of the Food Security Measurement Study." Cambridge, MA: Abt Associates Inc., report prepared for USDA Food and Consumer Service, June 1997.

Hansen, M.H., W.N. Hurwitz, and W.G. Madow, *Sample Survey Methods and Theory.* New York: Wiley and Sons, 1953

Landis, J.R. and G.C. Koch, "The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33: 159-174, 1977.

Mattes, R.D. and M.I. Friedman, "Hunger." *Digestive Diseases* 11: 65-77, 1993.

MacDonald, R.P., *Factor Analysis and Related Methods.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.., 1985.

Nunnally, J.C., *Psychometric Theory*, 2nd ed. New York: McGraw-Hill, 1978.

Olson, C.M., E.A. Frongillo, and A. Kendall, "Validation of Measures for Estimating the Prevalence of Hunger and Food Insecurity in the Current Population Survey Module: A Combination of Cornell and CCHIP Items." In *Conference on Food Security Measurement and Research, Papers and Proceedings, January 21-22, 1994.* Washington, DC: USDA/FCS and DHHS/NCHS, July 1995.

Radimer, K.L., C.M. Olson, J.C. Greene, C.C. Campbell, and J.P. Habicht, "Understanding Hunger and Developing Indicators to Assess it in Women and Children." *Journal of Nutrition Education* 24 (1): 36S-44S, 1992.

Radimer, K.L., C.M. Olson, and C.C. Campbell, "Development of Indicators to Assess Hunger." *Journal of Nutrition* 120 (Suppl 11): 1544-1548, 1990.

Read, M.A., S. French, and K. Cunningham, "The Role of the Gut in Regulating Food Intake in Man." *Nutrition Reviews* 52: 1-10, 1994.

Rosner, B.A., *Fundamentals of Biostatistics*, 2nd ed. Boston, MA: PWS Publishers, 1986.

Rulon, P.J., "A Simplified Procedure for Determining the Reliability of a Test by Split Halves." *Harvard Educational Review* 9: 99-103, 1939.

Scott, R.I., C. Wehler, and J. Anderson, "Measurement of Coping Behaviors as an Aspect of Food Security." In *Food Security Measurement and Research Conference: Papers and Proceedings,* Appendix A. Alexandria, VA: USDA Food and Consumer Service, June 1995.

Train, G. and L. Cahoon. "The Current Population Survey Variances, Inter-Relationships, and Design Effects." *Proceedings of the Section on Survey Research Methods of the American Statistical Association,* 1978, p. 443-448.

Trippe, C. and H. Beebout, *Food Stamp Program Participation Rates Among the Poverty Population, 1980-1987.* Washington, DC: Mathematica Policy Research, November 1988.

Trippe, C., P. Doyle, and A. Asher, *Trends in Food Stamp Program Participation Rates: 1976 to 1990.* Washington, DC: Mathematica Policy Research, July 1992.

U.S. Department of Agriculture, Food and Consumer Service, Office of Analysis and Evaluation, *Food Security Measurement and Research Conference: Papers and Proceedings.* Alexandria, VA: USDA Food and Consumer Service, June 1995.

Wehler, C.A., R.I. Scott, and J.J. Anderson, "The Community Childhood Hunger Identification Project: A Model of Domestic Hunger — Demonstration Project in Seattle, Washington." *Journal of Nutrition Education* 24 (1): 29S-35S, 1992.

Wright, B.D., and G.N. Masters, *Rating Scale Analysis: Rasch Measurement*. Chicago, IL: Mesa Press, 1982.

Appendix A

REVIEW OF LITERATURE FROM PHYSIOLOGY AND CLINICAL NUTRITION RESEARCH ADDRESSING THE NATURE OF HUNGER

APPENDIX A

REVIEW OF LITERATURE FROM PHYSIOLOGY AND CLINICAL NUTRITION RESEARCH ADDRESSING THE NATURE OF HUNGER

The literature summarized below was reviewed to answer two questions that are central to identifying conditions of actual hunger, defined as "the uneasy or painful sensation caused by a lack of food," as experienced within resource-constrained, food-insecure households. These questions are:

1. Are subjective reports of hunger reliable as indicators of a measurable level of food deprivation?

2. Is reduction in food intake a reasonable indicator or precursor of "the uneasy or painful sensation caused by a lack of food," or hunger?

The literature does not yield an unequivocal answer to the first question, though in sum it supports the validity of subjective reports of hunger. There is, however, considerable heterogeneity both in the experience of sensations reported as hunger (or associated with hunger) and in the extent to which these sensations are predictive of food intake. Moreover, there is not a one-to-one correspondence between the length of food deprivation and the intensity of sensations reported as hunger. A number of factors appear to condition the relation between food intake and physical sensations of hunger.

Regarding the second question, there is fairly strong support in the literature for the view that reduced food intake does lead to physical sensations of hunger. Up to a point, the intensity of such sensations are positively associated with the extent of food deprivation. Under conditions of prolonged fasting, however, the physical sensations associated with hunger diminish in intensity, and are generally extinguished altogether.

A reasonable conclusion from these studies is that people who experience patterns of undesired reductions in food intake below usual levels, such as may occur with severe limitation of household resources, do experience physical sensations of hunger.

Summary of Physiological Literature

Vanderweele and Geiselman (1986) review the proceedings of a symposium addressing relationships between appetite and need states in animals (including humans) that was part of the 1985 American Psychology Association Annual Meetings. Though most of the research presented in this symposium involved animal experiments, several themes of relevance to the current discussion emerged.

Geiselman presents results that relate to the distinction between "hunger" (defined as "food drive that occurs in response to post-absorptive stimulation") and "appetite" (defined as "food incentive occurring in response to pre-absorptive stimulation such as the anticipation, sight, smell, or taste of palatable food"). Results observed following infusion of hexoses into different locations within the gastric systems of laboratory animals lead Geiselman to conclude that both hunger and appetite involve common underlying physiologic mechanisms.

Vanderweele and Geiselman characterize the research presented at this symposium as generally supporting the hypothesis that both hedonic and physiological factors operate in the regulation of food intake behaviors in humans and other animals, and that it is thus unlikely that appetite and hunger can be completely separated. These findings provide some support for the experience of an "uneasy or painful sensation caused by a lack of food" when meal size is reduced below usual levels, or when meals are skipped, in that they indicate such sensations can arise both from physiological factors (e.g., rates of gastric emptying, intestinal absorption of nutrients, lipogenesis, and changes in blood glucose levels), and learned or conditioned associations (e.g., related to social and cultural factors, food preferences, palatability of foods, and time schedules regarding meals or other instances of food intake).

DeCastro and Elmore (1987) investigated the relationship between the subjective state of hunger and objective food intake among 31 free-living adults (9 male and 22 female) using a diary self-report method in which subjects recorded everything they ate or drank, the time of occurrence, and their degree of subjective hunger at the beginning of each eating occurrence over seven consecutive days. Food intake records were used to estimate stomach contents over time, based on a previously established formula.

DeCastro and Elmore conclude that the results of this study indicate that the intensity of self-reported hunger is dependent primarily on the contents of the stomach. This conclusion is based primarily on the finding of significant positive correlations between self-rated subjective

hunger at the beginning of meals and the duration of the pre-meal interval, significant negative correlations between self-rated subjective hunger and estimated pre-meal contents of the stomach, and significant multiple regression coefficients for estimated components of the contents of the stomach when regressed against the intensity of self-rated hunger. Total protein and food energy in the stomach contents were found to be the most salient factors in determining subjective hunger.

These researchers conclude that their results indicate that, as the stomach empties (a process occurring over a few hours), especially of protein, the internal subjective state of hunger increases. This leads them to observe that "the longer its [sic] been since the last time the subject ate, the hungrier they report themselves to be."

This study indicates that the contents of the stomach and stomach emptying are primary factors influencing the intensity of the "uneasy or painful sensations" reported as hunger. It also supports the likelihood that persons who reduce the size of their meals, or who skip meals, do experience hunger, because these behaviors make it more likely that the stomach contains little protein and few calories of food energy equivalent.

Wardle (1987) used three separate methods to track hunger among healthy women over a two-week period (including a behavioral method — amount of food intake; a physiological method — stimulated salivation prior to meals; and subjective reporting — paper and pencil recordings of the intensity of global hunger sensations, selection of experienced hunger symptoms, feeling of fullness, and type of foods selected). Results showed that ratings of hunger symptoms, preferred foods, global hunger, and satiety were sensitive to the length of time subjects were deprived of food and the content of the previous intake.

The behavioral, physiological, and subjective measures of hunger employed were consistent, reliable, and stable over time. Thus, this study indicates that the intensity of self-reported hunger increases with the length of the deprivation period. Moreover, because the usual intervals between intake episodes in the study were comparable to normal intervals between meals, these results support the likelihood that persons in food insecure households who cut the size of their meals, or who skip meals, experience "uneasy or painful sensations" that they report as hunger.

Harris and Wardle (1987) used a modified version of the 36-item "Monello and Mayer Hunger-Satiety Questionnaire" to assess hunger symptoms among two groups of subjects in pre-

meal and post-meal periods (group 1 — 274 female and 97 male undergraduate and adult education students; group 2 — 73 female and 84 male medical students). These researchers found a significant relationship between hours of deprivation and level of hunger reported. An unexpected finding from this study is the wide heterogeneity in hunger symptoms reported by subjects, with unexpectedly small numbers (e.g., 4-7) of symptoms reported in common by most subjects. The most commonly reported "uneasy or painful" symptoms were: emptiness, ache, urge to eat, rumbling, and hollowness. The relatively small numbers of commonly-endorsed symptoms lead these researchers to the conclusion that their results indicate that neither food deprivation nor reported hunger necessarily imply the perception of any particular bodily sensations. They further conclude that "it proved impossible to identify a specific subset or constellation of hunger symptoms which were characteristically experienced by hungry people."

This study is important for two reasons. First, it supports the contention that persons who cut meal size, or who skip meals, experience hunger, by showing that the longer the interval between episodes of food intake, the hungrier subjects report themselves to be. In addition, this study indicates that there is considerable heterogeneity in the way people experience the "uneasy or painful sensation caused by a lack of food." As Mattes and Friedman (1993) point out, the sensation of hunger is reported to occur in a variety of ways in many parts of the body. People experience a wide variety of sensations that they report as hunger, and the intensity of these sensations clearly increases (up to a point) as the period of deprivation increases.

Sepple and Read (1989) carried out an experiment in which ten normal healthy male volunteers were intubated with instrumentation to enable precise measurement of blood glucose levels, gastric emptying, and intestinal motor activity. Subjects who had fasted overnight were monitored for six hours, with measurements taken every 20-30 minutes before and after eating a meal. Subjects also completed a short questionnaire at 30-minute intervals to assess subjective ratings of hunger, fullness, anxiety, nausea, desire to eat, and other sensations. Intensity of each sensation was indicated by marking line analogues.

Sepple and Read found that eating the meal reduced the intensity of hunger ratings in all subjects, and totally abolished hunger in seven of the ten. The time for hunger to recur varied from 90-360 minutes, but was less than two hours for seven of the ten subjects. Once hunger ratings began to increase, they rose steadily in all subjects.

These researchers' results do not support the hypothesis that hunger is related to a decline in blood glucose concentration, but do strongly support the role of gastric emptying. A large and significant correlation was found between the time taken for 90 percent of the meal to empty and the time hunger started to increase ($r = 0.75$, $p < 0.02$). Maximum hunger ratings were obtained in all subjects when less than 10 percent of the meal remained in the stomach.

Sepple and Read conclude that their results are most compatible with the hypothesis that the sensation of hunger is induced by a reduction in stimulation of receptors in the upper intestine by nutrients present in food. Further support for this hypothesis was provided by observed appearance of fasting motor patterns in the upper intestine (also thought to reflect declining levels of nutrients) only at peak hunger levels in most subjects, and always after hunger had begun to increase.

The importance of the results of this clinical experiment is that it provides strong objective physiological evidence that emergence of the sensation of hunger accompanies, and is strongly related to, the emptying of food from the gastrointestinal tract, and that this occurs within a relatively short time period (about 2-4 hours). These results support the expectation that persons who reduce their food intake below usual levels, or who skip meals altogether, do experience sensations of hunger.

Mattes (1990) examined the relationship between self-reported hunger ratings and food intake in a study involving twelve normal male and twelve normal female subjects whose food intake and hunger ratings were recorded every waking hour for seven consecutive days. Although Mattes did not find a significant correlation between hunger ratings and food intake during the previous hour, he did find moderately large and significant correlations between hunger ratings and intake during the following hour.

Mattes observed two clear peaks in both the intensity of hunger ratings and food intake during each day, occurring at approximately 1200 and 1800 hours, or corresponding to the mid-day and evening mealtimes. Comparing difference scores for hunger ratings at the beginning and end of two-hour periods during which food was eaten in the intervening period, or not, Mattes found declines in hunger ratings over periods when eating occurred, and increases in hunger ratings over periods during which eating did not occur. The changes in ratings over the eating versus the no-eating periods were significant.

Curiously, Mattes found markedly different relationships between self-reported hunger and food intake on weekdays compared to weekend days. He posits that this suggests that eating in response to increased hunger sensations may involve an "entrained," or learned component, perhaps related to weekday work schedules. The differences observed over weekdays and weekend days are not consistent, Mattes suggests, with a strict energy depletion-hunger association, suggesting instead the influence of multiple factors.

Mattes' results indicate that the intensity of sensations of hunger among humans has two observable peaks during each day, corresponding closely to the timing of the mid-day and evening meals. More importantly, analysis of changes in the intensity of hunger ratings over time periods approaching usual inter-meal intervals (two hours) shows significant increases in hunger ratings over intervals in which eating does not occur, and significant declines in hunger ratings over intervals in which eating does occur.

Finally, the variations observed in the relationship between hunger ratings and food intake on weekdays versus weekend days supports the operation of a learned component in the relationship. Associations based on entrained or learned responses to hunger are consistent with social and cultural factors influencing the relationship between hunger and food intake, and with the experience of hunger if usual patterns of food intake are interrupted.

Ogden and Wardle (1990) examined the relationship between cognitive restraint of food intake and internal cues provided by caloric content of a pre-load, or pre-meal, intake. These authors found a significant effect of time since previous intake on subjective ratings of hunger. Subjects responded to an increased period of food deprivation (between morning intake and lunch) with an increase in subjective hunger ratings. Moreover, subjects with higher-calorie pre-loads at previous intake rated their subjective sensation of hunger significantly lower than subjects with low-calorie intakes. This indicates a significant effect of internal cues related to caloric content of previous intake on the level of subjective hunger sensation.

This study supports the general finding from other research that longer intervals between meals, or occurrences of food intake, lead to higher ratings of the sensation of hunger. It also indicates that persons who cut the size of their meals, or who skip meals, are more likely to experience the "uneasy or painful sensation caused by a lack of food" more intensely as the immediate post-meal time interval increases.

In one of the more unusual studies reviewed, ***DeCastro (1991)*** used seven-day food intake diaries with 121 male and 194 female adult subjects, together with pre- and post-meal implementations of a seven-point "full-hungry" scale, to assess food intake and hunger across seasons of the year. DeCastro not only found that subjects self-rated themselves as significantly hungrier (less sated) before meals than after meals, but also that, overall, subjects reported themselves significantly less hungry before meals eaten in the summer than in the winter or spring. Moreover, subjects reported themselves significantly hungrier (less sated) after meals eaten in the summer than in the winter or spring, and in the fall than in the winter.

DeCastro concludes that these results indicate that ingestion of a meal during the summer and fall leaves people hungrier than in the winter and spring. De Castro also finds that food and nutrient intakes are significantly greater in the summer and fall, with a mean increase of 14 percent in the fall. This study not only indicates that people who cut meal size or skip meals are more likely to report experiencing hunger than those who eat meals on their accustomed schedule, but that there is also seasonal variation in the intensity of hunger under these circumstances.

Lappalainen et al. (1990) examine hunger and food craving among two groups of obese patients. One group was fed a protein-sparing, well-balanced low-calorie diet for three weeks (1200-1600 Kcal per day), and the other provided "fasting therapy" in which food intake was reduced first to 800 Kcal per day, then to 200 Kcal per day over three days, and maintained at 200 Kcal per day for 19 additional days.

These researchers found that both frequency of hunger/craving responses and reactivity to food stimuli (reported changes in hunger state when shown pictures of food) decreased among the fasting group, but not among the group fed the protein-rich low-calorie diet. During the last (third) week of fasting, reactivity to food stimuli was completely abolished, and frequency of hunger/craving responses was reduced nearly to zero.

These results indicate that persons who reduce their food intake below its usual level (e.g., by cutting meal size or skipping meals) are likely to experience hunger and food-craving sensations, but if food intake is reduced dramatically and maintained at a very low level (as in prolonged fasting, or not having anything to eat for several days), sensations of hunger or food craving actually decline, and may disappear altogether.

Mattes and Friedman (1993) review several studies measuring factors associated with variability in subjective reporting of hunger. These authors report four definitions of hunger appearing in the literature reviewed:

1. Hunger is frequently operationally defined in terms of experimental or external conditions, such as the number of hours of food deprivation, or the size of the last meal. Emphasis is on attempting to make the term more objective or operational.

2. Hunger is used to refer to an intervening motivational state, or drive, that links experimental treatments or antecedent conditions (e.g., food deprivation) with behaviors aimed at obtaining food. Intensity of hunger is then inferred from either verbal reports of the desire for food, or willingness to perform a task for food.

3. Hunger is (most commonly) used to describe the subjective sensations associated with the need for food. The focus here is on the experience of various bodily states, sensations, or feelings, not on their causes.

4. Hunger is viewed as a physiological or metabolic state that results from a lack of energy or nutrients. This deficit state, which is detected by the nervous system, can modify eating behavior and food intake, and produce various subjective sensations collectively referred to as hunger. The emphasis with this usage is on physiological cause(s), not on the somatic manifestations of the need for food.

Mattes and Friedman focus their review primarily on the latter two of these definitions (i.e., hunger as a subjective experience, and hunger as a physiologically- or metabolically-based state resulting from the lack of energy or nutrients), expressing the view that "defining hunger in terms of a subjective or physiological state deals more directly with the experience and mechanisms of hunger, and therefore appears more relevant to basic research and clinical practice concerns." These researchers report findings of moderate and statistically significant correlations between reduced food intake and subjective reports of hunger, though the focus of their review is more on self-reported hunger as a predictor of food intake, rather than on whether reduction in food intake leads to the sensation of hunger.

Mattes and Friedman report findings from a study involving 800 individuals' responses to hypothetical fasts of varying durations wherein, under conditions described as extreme hunger, more than 90 percent of subjects indicated they experienced gastric sensations. This percentage declined to 50 percent of subjects reporting gastric sensations two hours prior to a typical meal. Similar patterns were noted for the mouth, throat, head, and general bodily sensations, although the proportion of subjects reporting these sites was smaller. Differences were observed by gender

and age, with mouth sensations more common among males than females, and head sensations more frequently reported by adults than by adolescents.

Mattes and Friedman also report results from their own research using open-ended questionnaires administered to 83 university students (45 male and 38 female), asking them to describe sensations associated with hunger of varying intensity (from slightly hungry to extremely hungry). Gastric sensations were also the most commonly reported symptom among subjects in this study, with 55 percent reporting stomach growls and 34 percent reporting stomach aches. When asked to report the body sites where sensations associated with different levels of intensity of hunger were experienced, 70 percent of subjects reported sensations in their stomachs when they were "slightly" hungry, 82 percent when they were "moderately" hungry, 92 percent when they were "very" hungry, and 88 percent when they were "extremely" hungry. Fewer subjects reported sensations in their heads, with 10 percent experiencing hunger-related sensations when "slightly" hungry, increasing to 43 percent when they were "extremely" hungry.

Mattes and Friedman identify a number of physiological factors related to the sensation of hunger, including reduction in gastric distention (reduced feeling of pressure caused by emptying of food from the stomach), metabolic signals transmitted from the liver and small intestine, and sensory input from the oral cavity. All of these factors are conceptually consistent with experience of an "uneasy or painful sensation caused by a lack of food" when meals are cut or skipped.

Read et al. (1994) review research on the role of gastrointestinal processes in regulation of food intake in humans. These authors emphasize the role of factors associated with emptying of nutrients from the upper small intestine (which occurs within a relatively short period after ingestion of food) in signaling the sensation of hunger. They also discuss the phenomenon of nutrient adaptation, wherein humans (and other animals) adapt to a particular pattern of availability of energy and other nutrients, leading to moderation of the sensations of hunger and satiety under persistent conditions of reduced intake. This adaptive process (also noted by Mattes and Friedman (1993) and Lappalainen *et al.* (1993), and discussed above) leads to a reduction in the intensity of hunger sensation, and its eventual extinction after prolonged fasting, and implies that deviation from normal eating patterns (such as would occur if meals are cut or skipped) can lead to more intense subjective sensations of hunger than would occur under more prolonged intake reduction.

Read *et al.* (1994) conclude that the human gastrointestinal tract is compatible with the observed tendency of humans to eat three or four meals a day, and that the gastrointestinal mechanisms that trigger both initiation and termination of eating behavior operate within a relatively short time period (e.g., a few hours). This implies a very high likelihood that recurrent reduction of food intake by cutting meal size or skipping meals because of insufficient money to buy food will lead to the experience of an "uneasy or painful sensation caused by a lack of food." Adaptation of the gastrointestinal system to prolonged changes in food or nutrient availability may moderate the intensity of this sensation if intake reduction is prolonged or stabilized at reduced levels.

Rolls (1993) examines appetite, hunger, and satiety among the elderly population, reviewing a number of studies addressing causes and consequences of reduced food intake, decline in appetite and olfactory and gustatory sensory acuity, and reduction of sensory-specific satiety among elderly persons. Rolls also reports results of her own research on decline in sensory-specific satiety among older persons.

Satiety is generally the converse of hunger; the hungrier one is, the less sated they are. Satiety is both the complex of sensations that reduce the motivation to eat as more food is ingested, and the declining palatability of specific foods as they are ingested. Sensory-specific satiety is the decline in pleasantness of a particular food following consumption of that food. Sensory-specific satiety is associated with decreased consumption of the previously-eaten food and a shift in consumption to other food choices. It appears to decrease, and even disappear, among persons over age 65 years.

This study is of interest not because it provides evidence that persons who cut meal size or skip meals experience hunger, but because it suggests that the nature of the "uneasy or painful sensation caused by a lack of food," or the sensation of hunger, changes in ways that may reduce or mask its intensity among elderly persons. To the extent this occurs, elderly persons may actually under-report their experience of hunger.

References

DeCastro, J. "Seasonal Rhythms of Human Nutrient Intake and Meal Pattern." *Physiology and Behavior* 50: 243-248, 1991.

De Castro, J., and D.K. Elmore. "Subjective Hunger Relationships with Meal Patterns in the Spontaneous Feeding Behavior of Humans: Evidence for a Causal Connection." *Physiology and Behavior* 43: 159-165, 1987.

Harris, A., and J. Wardle. "The Feeling of Hunger." *British Journal of Clinical Psychology* 26: 153-154, 1987.

Lappalainen, R., P. Sjoden, T. Hursti, and V. Vesa. "Hunger/Craving Responses and Reactivity to Food Stimuli During Fasting and Dieting." *International Journal of Obesity* 14: 679-688, 1990.

Mattes, R.D. "Hunger Ratings Are Not a Valid Proxy Measure of Reported Food Intake in Humans." *Appetite* 15: 103-113, 1990.

Mattes, R.D., and M.I. Friedman. "Hunger." *Digestive Diseases* 11: 65-77, 1993.

Ogden, J., and J. Wardle. "Cognitive Restraint and Sensitivity to Cues for Hunger and Satiety." *Physiology and Behavior* 47(3): 477-81, 1990.

Read, M.A., S. French, and K. Cunningham. "The Role of the Gut in Regulating Food Intake in Man." *Nutrition Reviews* 52: 1-10, 1994.

Rolls, B.J. "Appetite, Hunger, and Satiety in the Elderly." *Reviews of Food Science and Nutrition* 33(1): 39-44, 1993.

Sepple, C.P., and N.W. Read. "Gastrointestinal Correlates of the Development of Hunger in Man." *Appetite* 13: 183-191, 1989.

Vanderweele, D.A., and P.J. Geiselman (eds.). "Can Appetite and Need State Be Separated?" Abstracts from presentations at a satellite symposium of the American Psychological Association meetings at Occidental College, Los Angeles, CA, August 1985. *Appetite* 7: 187-194, 1986.

Wardle, J. "Hunger and Satiety: A Multidimensional Assessment of Responses to Caloric Loads." *Physiology and Behavior* 240: 577-582, 1987.

Appendix B

PREVALENCE OF HOUSEHOLD FOOD SECURITY STATUS (30-DAY SCALE)

Exhibit B-1

PREVALENCE OF HOUSEHOLD FOOD SECURITY STATUS BY SELECTED CHARACTERISTICS OF HOUSEHOLDS: 30-DAY SCALE

[Numbers in thousands. Poverty status refers to household status in the preceding year.] Characteristic	Total	No Food Insecurity with Hunger Evident[a]		Food Insecure with Moderate Hunger Evident[b]		Food Insecure with Severe Hunger Evident[c]	
		Number	Percent of Total	Number	Percent of Total	Number	Percent of Total
Household Composition							
All races:							
With children under 18 yrs	38,113	36,877	96.8	1031.9	2.7	203.6	0.5
With children under 6 yrs	18,282	17,661	96.6	519.8	2.8	101.0	0.6
With elderly;[d] no children	27,805	27,567	99.1	201.9	0.7	36.6	0.1
With no elderly or child	34,291	33,359	97.3	725.0	2.1	207.3	0.6
All household types	100,210	97,803	97.6	1958.8	2.0	447.5	0.4
White:							
All households							
With children under 18 yrs	30,438	29,622	97.3	686.3	2.2	130.1	0.4
With children under 6 yrs	14,467	14,048	97.1	358.3	2.5	60.4	0.4
With elderly; no children	25,012	24,838	99.3	150.7	0.6	23.8	0.1
With no elderly or child	29,163	28,517	97.8	507.0	1.7	139.1	0.5
Black:							
All households							
With children under 18 yrs	5,841	5,485	93.9	298.4	5.1	57.6	1.5
With children under 6 yrs	2,826	2,669	94.4	130.0	4.6	27.2	1.0
With elderly; no children	2,321	2,265	97.6	43.4	1.9	12.0	0.5
With no elderly or child	3,852	3,598	93.4	197.1	5.1	56.4	1.5
Other:							
All households							
With children under 18 yrs	1,833	1,770	96.6	47.2	2.6	16.0	0.9
With children under 6 yrs	989	944	95.5	32.0	3.2	13.4	1.4
With elderly; no children	472	464	98.2	7.9	1.7	0.9	0.2
With no elderly or child	1,276	1,244	97.4	20.9	1.6	1.2	0.9
Hispanic:[e]							
All households							
With children under 18 yrs	4,475	4,274	95.5	160.6	3.6	41.3	0.9
With children under 6 yrs	2,539	2,379	93.7	122.2	4.8	6.9	1.5
With elderly; no children	1,151	1,115	96.9	27.3	2.4	8.4	0.7
With no elderly or child	2,075	1,991	95.9	47.2	2.3	37.3	1.8

Notes at end of exhibit

Exhibit B-1 *(continued)*

PREVALENCE OF HOUSEHOLD FOOD SECURITY STATUS BY SELECTED CHARACTERISTICS OF HOUSEHOLDS: 30-DAY SCALE

[Numbers in thousands. Poverty status refers to household status in the preceding year.] Characteristic	Total	No Food Insecurity with Hunger Evident[a]		Food Insecure with Moderate Hunger Evident[b]		Food Insecure with Severe Hunger Evident[c]	
		Number	Percent of Total	Number	Percent of Total	Number	Percent of Total
Household Income Category[f]							
(All races and household types)							
Below $10,000	14,977	13,893	92.8	822.0	5.5	262.4	1.8
$10,000 - $19,999	16,717	16,043	96.0	568.4	3.4	105.2	0.6
$20,000 - $29,999	15,625	15,319	98.0	268.2	1.7	37.7	0.2
$30,000 to $39,999	12,149	12,046	99.2	87.2	0.7	15.9	0.1
$40,000 - $49,999	8,539	8,488	99.4	43.2	0.5	8.3	0.1
Above $50,000	22,370	22,319	99.8	50.3	0.2	0.8	**
Household Income-to-Poverty Ratio[f]							
(All races and household types)							
Under 0.50	5,545	4,987	89.9	415.7	7.5	142.2	2.6
Under 1.00	15,808	14,617	92.5	904.2	5.7	287.2	1.8
Under 1.30	21,810	20,304	93.1	1178.6	5.4	326.9	1.5
Under 1.85	35,115	33,239	94.7	1503.7	4.3	372.0	1.1
1.85 and over	65,094	64,564	99.2	455.1	0.7	75.5	0.1
Household Relationship							
(All races)							
Households with children under 18	38,113	36,877	96.8	1031.9	2.7	203.6	0.5
Married couple families	26,841	26,347	98.2	445.2	1.7	48.9	0.2
Female head, no spouse	8,941	8,290	92.7	509.6	5.7	140.9	1.6
Male head, no spouse	2,332	2,241	96.1	77.1	3.3	13.8	0.6
Households with no children or	34,291	33,359	97.3	725.0	2.1	207.3	0.6
Living alone	13,724	13,151	95.8	434.5	3.2	138.6	1
Households with elderly but no	27,805	27,567	99.1	201.9	0.7	36.6	0.1
Living alone	11,699	11,544	98.7	131.7	1.1	23.0	0.2
Area of Residence							
(All races and household types)							
Inside Metropolitan areas	60,657	59,155	97.5	1215.0	2.0	287.1	0.4
In central city	24,055	23,266	96.7	671.2	2.8	117.7	0.5
Not in central city	36,602	35,889	98.0	543.8	1.5	169.4	0.5
Outside Metropolitan areas	23,877	23,298	97.6	478.0	2.0	100.8	0.4

Notes on next page

Exhibit B-1 *(continued)*

PREVALENCE OF HOUSEHOLD FOOD SECURITY STATUS BY SELECTED CHARACTERISTICS OF HOUSEHOLDS: 30-DAY SCALE

NOTES:

[a] No or minimal indicators of resource-constrained hunger evident for household members (corresponds to the combined categories of "food secure" and "food secure, hunger not evident" in the 12-month scale).

[b] Multiple indicators of resource-constrained hunger evident for adult household members.

[c] Multiple indicators of resource-constrained hunger evident for children in household and/or indicators of severe adult hunger.

[d] Elderly persons are defined as persons aged 60 years and older in this report.

[e] Persons of Hispanic ethnicity can be of any race.

[f] Income and poverty status refer to household income in a recent 12-month period, varying among rotation groups in the CPS sample. Income is missing for 9.8 percent of households but their income-to-poverty ratio category was imputed by the Census Bureau.

[g] For confidentiality reasons the CPS did not report the area of residence for 15.6 percent of households. The estimates shown are for households with area of residence identified.

Exhibit B-2

STANDARD ERRORS FOR PREVALENCE OF HOUSEHOLD FOOD SECURITY STATUS BY SELECTED CHARACTERISTICS OF HOUSEHOLDS: 30-DAY SCALE

[Numbers in thousands. Poverty status refers to household status in the preceding year.] Characteristic	Sample Size (in ones)	No Food Insecurity with Hunger Evident[a]		Food Insecure with Moderate Hunger Evident[b]		Food Insecure with Severe Hunger Evident[c]	
		Number	Percent of Total	Number	Percent of Total	Number	Percent of Total
Household Composition							
All races:							
With children under 18 yrs	16,914	61	0.16	67	0.18	24	0.06
With children under 6 yrs	7,934	41	0.23	41	0.22	14	0.08
With elderly;[d] no children	12,485	40	0.14	35	0.13	8	0.03
With no elderly or child	15,248	85	0.25	55	0.16	38	0.11
All household types	44,647	110	0.11	97	0.10	42	0.04
White:							
All households							
With children under 18 yrs	13,808	51	0.17	52	0.17	18	0.06
With children under 6 yrs	6,391	37	0.25	32	0.22	14	0.10
With elderly; no children	11,283	29	0.12	26	0.11	4	0.02
With no elderly or child	13,137	55	0.19	39	0.13	25	0.09
Black:							
All households							
With children under 18 yrs	2,023	18	0.31	22	0.38	13	0.21
With children under 6 yrs	959	26	0.91	23	0.82	7	0.26
With elderly; no children	926	16	0.68	13	0.57	5	0.21
With no elderly or child	1,370	40	1.03	27	0.69	16	0.43
Other:							
All households							
With children under 18 yrs	1,083	14	0.75	13	0.70	8	0.45
With children under 6 yrs	584	8	0.84	7	0.73	7	0.75
With elderly; no children	276	5	0.98	4	0.94	1	0.14
With no elderly or child	741	8	0.62	7	0.57	7	0.58
Hispanic:[e]							
All households							
With children under 18 yrs	1,529	27	0.61	21	0.48	12	0.28
With children under 6 yrs	857	23	0.90	18	0.69	13	0.49
With elderly; no children	406	10	0.88	7	0.63	5	0.39
With no elderly or child	695	15	0.74	12	0.59	13	0.63

See notes to Exhibit B-1

Exhibit B-2 *(continued)*

STANDARD ERRORS FOR PREVALENCE OF HOUSEHOLD FOOD SECURITY STATUS BY SELECTED CHARACTERISTICS OF HOUSEHOLDS: 30-DAY SCALE

[Numbers in thousands. Poverty status refers to household status in the preceding year.] Characteristic	Sample Size (in ones)	No Food Insecurity with Hunger Evident[a]		Food Insecure with Moderate Hunger Evident[b]		Food Insecure with Severe Hunger Evident[c]	
		Number	Percent of Total	Number	Percent of Total	Number	Percent of Total
Household Income Category[f]							
(All races and household types)							
Below $10,000	6,368	42	0.28	47	0.31	45	0.30
$10,000 - $19,999	7,651	69	0.41	59	0.35	12	0.07
$20,000 - $29,999	7,202	40	0.26	33	0.21	13	0.09
$30,000 to $39,999	5,808	19	0.15	18	0.15	6	0.05
$40,000 - $49,999	4,037	15	0.18	15	0.18	5	0.05
Above $50,000	10,690	16	0.07	16	0.07	1	0.01
Household Income-to-Poverty[f]							
(All races and household types)							
Under 0.50	2,219	25	0.45	25	0.45	30	0.54
Under 1.00	6,650	51	0.32	61	0.38	42	0.26
Under 1.30	9,384	85	0.39	83	0.38	50	0.23
Under 1.85	15,594	103	0.29	92	0.26	50	0.14
1.85 and over	29,053	25	0.04	18	0.03	17	0.03
Household Relationship							
(All races)							
Households with children under 18	16,914	61	0.16	67	0.18	24	0.06
Married couple families	12,295	39	0.15	40	0.15	7	0.03
**Female head, no spouse	3,677	67	0.75	66	0.73	16	0.18
**Male head, no spouse	942	15	0.63	12	0.53	8	0.33
Households with no children or	15,248	85	0.25	55	0.16	38	0.11
Living alone	5,941	60	0.44	38	0.28	32	0.23
Households with elderly but no	12,485	40	0.14	35	0.13	8	0.03
Living alone	5,222	31	0.26	23	0.20	9	0.07
Area of Residence:							
(All races and household types)							
Inside Metropolitan areas	24,214	91	0.15	71	0.12	32	0.05
In central city	9,606	74	0.30	51	0.21	28	0.11
Not in central city	14,608	34	0.09	29	0.08	15	0.04
Outside Metropolitan areas	12,532	40	0.17	29	0.12	22	0.09

See notes to Exhibit B-1

APPENDIX C

PARTICIPANTS IN FEDERAL INTERAGENCY WORKING GROUP FOR FOOD SECURITY MEASUREMENT

USDA/FCS

Steven Carlson
Gary Bickel
Margaret Andrews
Sharron Cristofar
Dawn Aldridge

USDA/CNPP

Peter Basiotis
Bruce Klein (formerly FCS)

USDA/ERS

David Smallwood
James Blaylock
Mark Nord
Donald Rose
Victor Oliveira

USDA/ARS

Mary Hama

USDA/CSREES

Nancy Leidenfrost

DHHS/CDC/NCHS

Ronnette Briefel
Katherine Alaimo

DHHS/CDC&P

BettyLou Sherry

DHHS/AOA

Jean Lloyd

Census Bureau/CPS

Julie Feliciano
Maria Reed
Kathleen Stoner

Census Bureau/CSMR

Jennifer Hess