August 2013

Feinstein International Center

Strengthening the humanity and dignity of people in crisis through knowledge and practice

How Do Different Indicators of Household Food Security Compare? Empirical Evidence from Tigray

Daniel Maxwell, Jennifer Coates and Bapu Vaitla

Tufts University
Gerald J. and Dorothy R. Friedman School of Nutrition Science and Policy

Feinstein International Center
Tufts University
114 Curtis Street
Somerville, MA 02144
USA
tel: +1 617.627.3423
fax: +1 617.627.3428
fic.tufts.edu

Suggested Citation

Maxwell, Daniel, Jennifer Coates, and Bapu Vaitla (2013). *How Do Different Indicators of Household Food Security Compare? Empirical Evidence from Tigray.* Feinstein International Center, Tufts University: Medford, USA.

Acknowledgements

This paper is based on data collected by the "Livelihoods Change over Time" (LCOT) study in two districts of Tigray State in northern Ethiopia. We would like to thank Dr. Girmay Tesfay and Dr. Nigussie Abadi who led the field team that collected the data for the LCOT study, and their colleagues from the College of Dryland Agriculture and Natural Resources at Mekelle University.

At the Feinstein International Center, we thank Peter Walker, Rosa Pendenza, Ann O'Brien, Beth O'Leary, and Anita Robbins, for their support to the study. We are very grateful to Joyce Maxwell for editing and formatting this lengthy report.

And of course we would like to thank the many people in Tsaeda Amba and Seharati Samra woredas in Tigray who generously gave their time and knowledge in four rounds of survey interviewing.

Funding for this research came from the Swedish International Development Agency (SIDA). We gratefully acknowledge SIDA's support throughout the research and publication process.

All analysis is by the authors, who are responsible for any errors.

The Authors
August 2013

Cover Photo Credit

James P. Wirth, GroundWork

Table of Contents

I. Introduction

With recent food crises at both regional and global levels, and renewed commitments from major donor countries to address chronic hunger, food security is more prominent on the policy agenda today than it has been in the past. Food security hardly needs to be defined again, and this paper follows others in using the UN Food and Agriculture Organization (FAO) definition.[1] The renewed emphasis on addressing constraints to food security has intensified the search for accurate, rapid, and consistent indicators of food security. Barrett (2010) notes that approaches to measurement follow the four major "pillars" of food security—availability, access, utilization, and risk (sometimes alternatively labeled stability or vulnerability) —which in turn tend to follow different strands of analysis. Measures of food access are important for many reasons but, practically speaking, they are most urgently required for purposes of early warning, for assessment of current and prospective status of at-risk populations, and for monitoring and evaluating specific programs and policies. More recently, given innovations such as the Integrated Phase Classification (IPC) tool, they are important for establishing the comparability of food security status in dissimilar contexts—a task that is critical for targeting resources on any sort of rational or impartial basis.

Different measures of the access dimension of food insecurity are used interchangeably, without a good idea of which food-security dimensions are captured by which measures. The associated risk is that the number of food-insecure individuals is underestimated when single measures are applied that are incongruent with a more holistic food-security definition (Coates, 2013). This paper seeks to address this important gap in the literature through a comparative empirical analysis of common food-security indicators. Specifically, the paper addresses the following questions:

1) How do the seven measures listed below compare—do they tell the same "story" about household food insecurity and classify households similarly?
2) Which elements of food insecurity does each of these measures capture?
3) How can metrics be combined or used in complementary ways to yield a more multidimensional picture of a household's food insecurity situation?

The paper draws on four rounds of data from a panel survey of 300 rural households in northern Ethiopia to compare seven different measures collected across all rounds: (1) Coping Strategies Index (CSI); (2) Reduced Coping Strategies Index (rCSI); (3) Household Food Insecurity and Access Scale (HFIAS); (4) The Household Hunger Scale (HHS); (5) Food Consumption Score (FCS); (6) Household Dietary Diversity Scale (HDDS); and (7) a self-assessed measure of food security (SAFS).[2]

We assess inter-correlations among the seven indicators and then analyze whether the different measures detect the same or different dimensions of the complex phenomenon of food insecurity. Finally, we use these results to illustrate other means of capturing food insecurity, including combinations of stand-alone indicators as well as the development of an operationally simple but dimensionally rich food security tool. The answers to the questions posed in this study have implications for (1) which measure, or which combinations of measures, is more appropriate for a given purpose and (2) the costs of relying on single measures or indicators, in terms of potential misclassification of the food-insecure.

II. Measuring Food Security

Given the multidimensional nature of food security, practitioners and policy makers have long recognized the need for a variety of means of measurement (Kennedy 2002, FAO 2013). The "holy grail" of food security measurement would be a single measure that is valid and reliable, comparable over time and space, and which captures different elements of food security. In spite of the development of many different indicators in the past decade, no single one meets these criteria (Coates and Maxwell 2012). For many years, age-adjusted per-capita caloric intake was considered the "gold standard" for access to food at the household level, and anthropometric measures of nutritional status were the gold standard at the individual level (Hoddinott and Yohannes 2002, Weismann et al. 2006; Coates et al. 2007). But nutritional status is determined by many factors, food security being only one (Young and Jaspars 2006, 2009). And while per-capita caloric intake reflects current consumption—the question of quantity—it does not address many other elements of the complicated notion of "food security," such as quality (dietary diversity and micronutrient sufficiency), vulnerability and risks, and fluctu¬ations and trends in consumption over time. Caloric intake can also be very time consuming and expensive to measure and so is rarely used as a measure of food security in any context other than basic research.

Progress on developing indicators has been mixed. The FAO measure for prevalence of undernourishment—most prominently used to track progress against

1 "Food security exists when all people, at all times, have physical and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life" (World Food Summit, 1996).

2 The actual wording of the questions used in the household survey to derive these indicators is reprinted in Annex A.

Millennium Development Goals—has recently been completely revamped, but it is not intended for detecting short-term changes or acute emergencies (Cafiero 2012). Significant progress has been made on utilizing Household Consumption and Expenditure Surveys (HCES) for tracking food insecurity (Smith et al. 2006, Fiedler 2013), but this kind of measure is not typically used for rapid field assessments in dynamic contexts as data are only updated every few years. Looking specifically at rapid field measures, several recent studies have confirmed a significant correlation between the Food Consumption Score (FCS)—a food frequency measure developed by the World Food Programme (WFP)—and caloric consumption, but degrees of correlation across contexts vary, and it often underestimates the prevalence of food insecurity (IFPRI 2008, Coates et al. 2007). One recent field validation test conducted in Latin America found that the proposed "universal" thresholds for the FCS were badly misclassifying food insecurity (defined in that study as caloric adequacy)—but also found that conducting field validation tests everywhere would be prohibitively expensive, putting into question whether thresholds could actually be considered universal, even while validating the FCS measure in terms of correlation with other indicators (WFP 2010). A series of articles outlined the development of a Coping Strategies Index that correlates with both caloric intake and other measures of food access (Maxwell 1996; Maxwell et al. 1999; Maxwell, Caldwell, and Langworthy 2008). A different strand of research outlines the development of the Household Food Insecurity Access Scale (Coates, Swindale, and Bilinsky 2006; Webb et al. 2006). Recently, there has also been a rise in self-assessments of food security status, as noted by Headey (2011, 2013).

Rapid, accurate, cross-contextual indicators of food security have been developed over the past decade or so. These fall into several recognizable categories:

- Dietary diversity and food frequency. This type of metric captures the number of different kinds of food or food groups that people eat and the frequency with which they eat them, and sometimes involves weighting these groups. The result is a score that represents the diversity of intake, but not necessarily the quantity, though such scores have been shown to be significantly correlated with caloric adequacy measures (IFPRI 2006, Coates et al. 2007). The Food Consumption Score (FCS) is a specific type of dietary diversity index used primarily by the World Food Programme (WFP 2009). The Household Dietary Diversity Score (HDDS)—similar to the FCS, but with a 24-hour recall period without frequency information or weighted categorical cut-offs—has been widely promoted by the UN Food and Agriculture Organization and USAID (FANTA 2006, FAO 2010).
- Spending on food. Given the propensity of people closer to the edge of poverty to spend a greater and greater proportion of their income on food, estimating the proportion of expenditure on food has become an important measure (Smith et al. 2006).
- Consumption behaviors. These measures capture food security indirectly, by measuring behaviors related to food consumption. Perhaps the best known example is the Coping Strategies Index or CSI (Maxwell and Caldwell 2008), which counts the frequency and severity of behaviors in which people engage when they do not have enough food or enough money to buy food. Recent work on the CSI has identified a more "universal" sub-set of coping behaviors found to be relevant in 14 different context-specific CSI instruments (Maxwell, Caldwell, and Langworthy 2008). This "reduced CSI" (rCSI) is probably more widely used now than the original form, but tends to measure only the less-severe coping behaviors. Versions of the CSI have been widely adopted by WFP/VAM (World Food Programme/Vulnerability Analysis Mapping unit), FAO/FSNAU (UN Food and Agriculture Organization/Food Security and Nutrition Analysis Unit for Somalia), and the Global IPC (Integrated Phase Classification) team, among others. The Household Hunger Scale (HHS—see below) is also essentially a behavioral measure. It tends to capture more-severe behaviors.
- Experiential measures. Some indicators combine behavioral with psychological measures. The Latin America and Caribbean Food Security Scale (ELCSA) and the Household Food Insecurity Access Scale are the best known and most widely used of these measures in international contexts (ELCSA Scientific Committee 2012; Swindale and Bilinsky 2006; Coates, Swindale, and Bilinsky 2006). The HFIAS was designed to capture household behaviors signifying insufficient quality and quantity, as well as anxiety over insecure access. The Household Hunger Score (HHS) was derived from the HFIAS as a culturally-invariant subset of questions, and includes three specific questions, none of which are psychological in nature (Deitchler et al. 2010). USAID, FAO, and others have adopted and promoted the HFIAS and HHS.
- Self-assessment measures. Though highly subjective in nature and perhaps too easy to manipulate in programmatic contexts, self-assessment measures have been introduced in recent years. These include self-assessments of current food security status in a recent recall period and the change in livelihood status over a longer period of time. Self-assessed food security was collected through the Gallup poll to examine the food security effects of the global food price crisis (Headey 2011, 2013).

> Recently, Headey and Ecker (2013) critique self-assessment measures in a review of food security metrics. In the present study, the self-assessed food security measure was constructed by asking households to characterize the state of food security as compared to a "normal" (not good, not bad) year on a five-category scale.[3]

There is strong evidence that all these measures are capturing something about the multidimensional nature of food security. However, there have been few assessments of which dimensions of food security are captured by each measure and few direct empirical comparisons among them. The rest of this paper will compare these measures directly to answer the research questions noted earlier.

III. Empirical Context and Methodology

The data for this paper come from the Livelihoods Change Over Time (LCOT) four-round panel survey conducted in two districts (woredas) of northern Ethiopia between August 2011 and February 2013. The overall objective of the LCOT panel survey is to assess household resilience in the face of an annually recurring shock: the "hunger season." This time of year is characterized by price inflation as the previous year's harvest stocks diminish and grain prices increase in local markets. The higher prices during the hunger season are coupled with increased illness, especially malaria and acute respiratory infections, during the months immediately preceding the harvest. Rates of acute undernutrition and morbidity increase, and households are often forced to sell key assets, especially livestock, to meet basic needs. Families also engage in a wide range of behaviors—some harmful or unsustainable—to cope with hunger season difficulties[4].

To capture within-year as well as across-year livelihood dynamics, we chose to collect panel data on our sample two times a year: at the height of the hunger season in August and in the middle of the postharvest season in February (two-to-three months after harvest). We expected the former time to be when households have the least amount of available income and food stocks, and the latter time to represent the period when households have the greatest amount of income and food. Each survey round not only gathers information on the situation prevailing at the time, but also asks retrospective questions about household decisions and experiences over the six-month period prior to the survey (i.e., since the last survey round) and, in the case of the food security measures discussed earlier, over the one month preceding the survey. All four rounds of the survey are now complete.

The survey sample was stratified to represent the livelihood- and food-security related variables of two woredas: Tsaeda Amba (Eastern Tigray) and Seharti Samre (Southern Tigray). In each woreda, 150 households were selected, 75 from each of two kebeles (sub-district units). The sub-kebele (i.e., village- or kushet-level) sampling units were obtained by systematic selection with a random start. The probability of each sampling unit being selected was proportional to the village's size. Within the village, sampling of households was done by random selection of transects.

IV. Analytical Results

A. How do these seven measures compare?

1. How well do the different measures correlate?

To examine this question, the strength of the correlation among indicators was examined using Spearman's r, which is similar to Pearson's r but is used to examine non-parametric relationships. The HFIAS, HHS, CSI, rCSI, and SAFS scales were adjusted so that a higher score indicates greater food insecurity, whereas higher HDDS and FCS scores indicate greater dietary diversity and food frequency and, thus, less food insecurity. Thus, inverse correlations among some of these indicators were expected. The Spearman's r correlations among the seven measures are generally quite strong, are associated in the expected direction, and are all significant at the $p<0.01$ level, as shown in the table of pooled correlations (all rounds) (Table 1). The CSI, rCSI, and HFIAS are particularly well correlated. As would be expected, the FCS—measuring dietary diversity—is much more strongly correlated with the HDDS than with either the HFIAS or CSI. The food security self-assessment measure is moderately well correlated with HFIAS, CSI, and rCSI, but more weakly with the other measures (although even with these latter variables, the associations are statistically significant). Of the alternative measures, both rCSI and HDDS maintain the same strength of correlation with other measures as their expanded or weighted counterparts, CSI and FCS, do. However,

[3] Food secure; slightly food insecure; moderately food insecure; very food insecure; extremely food insecure.

[4] For interim results on this broader research, see an earlier paper by Vaitla et al. (2012). The current paper is restricted to the empirical comparison of food security indicators and does not address the central research questions of the main study.

Table 1. Spearman's rho correlations between food security measures, all rounds pooled							
	CSI	rCSI	HFIAS	HHS	FCS	HDDS	SAFS
CSI	1	0.95	0.85	0.44	-0.51	-0.56	0.45
rCSI	0.95	1	0.84	0.42	-0.48	-0.53	0.46
HFIAS	0.85	0.84	1	0.48	-0.57	-0.63	0.46
HHS	0.44	0.42	0.48	1	-0.34	-0.34	0.23
FCS	-0.51	-0.48	-0.57	-0.34	1	0.92	-0.24
HDDS	-0.56	-0.53	-0.63	-0.34	0.92	1	-0.29
SAFS	0.45	0.46	0.46	0.23	-0.24	-0.29	1
* All correlations significant at the p<0.01 level							

Table 2. Classification systems of food security measures			
Indicator	Original Category	Original qualitative label	Converted binary classification
CSI/ rCSI*/HFIAS	1	Food secure	Food secure
	2	Mildly food insecure	
	3	Moderately food insecure	Food insecure
	4	Severely food insecure	
HHS	1	Little to no hunger	Food secure
	2	Moderate hunger	Food insecure
	3	Severe hunger	
FCS	1	Acceptable	Food secure
	2	Borderline	Food insecure
	3	Poor	
SAFS	1	Food secure	Food secure
	2	Slightly food insecure	
	3	Moderately food insecure	Food insecure
	4	Very food insecure	
	5	Extremely food insecure	
* The guidance manual for CSI and rCSI does not provide thresholds for different categories of food security and insecurity. Thresholds were developed for this study but they are not necessarily intended to be universally applicable.			

HHS does not maintain the same degree of correlation with other measures as its "parent" measure, HFIAS. This is expected, as the HHS measures the most extreme consequences of food insecurity, while the HFIAS captures a greater range of the food security spectrum. By contrast, the rCSI picks up some of the less-severe coping behaviors; thus, we see a relatively weaker correlation between HHS and rCSI than between the other measures—which may be explained in part by the fact that the "hunger season" in the years of the survey was milder than in other recent years.

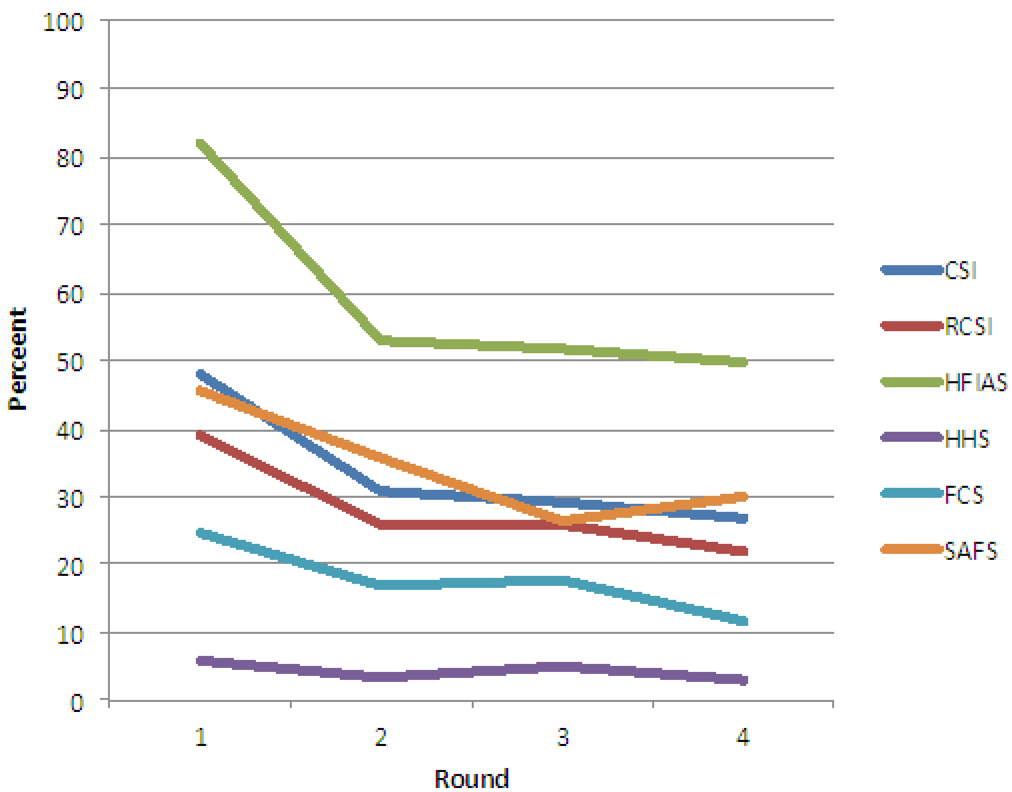
The strength of correlation between the various food security measures does not change in any consistent pattern when disaggregated by rounds or by a hunger season/harvest season split; Annex B presents correlations disaggregated by round.

Overall, the measures correlate quite strongly, and yet there is enough unexplained variance to suggest that each metric might be capturing a different aspect of food security, a topic discussed in greater detail in subsequent sections.

Table 3. Food insecure households as percentage of total, by round and indicator

Round	CSI	RCSI	HFIAS	HHS	FCS	HDDS	SAFS
1	48.0	38.9	81.8	6.0	24.5	46.2	45.7
2	31.0	26.1	52.9	3.4	16.8	43.2	35.9
3	29.1	25.7	51.7	5.2	17.9	36.3	26.4
4	26.8	21.9	49.7	2.9	11.7	34.6	29.8

Figure 1. Food insecure households as percentage of total, by round and indicator



2. *What do the different measures tell us about changes in food insecurity across rounds?*

Most of these measures can be used in a continuous form (as a scale or index) or as a categorical indicator (for estimating prevalence). Different assumptions and methods were used to derive cut-points for the creation of each of these categorical indicators. These indicators also classify households into different numbers of categories, and in some cases the qualitative labels attached to these categories differ as well. This information is summarized in Table 2. (Note that HDDS does not have established categorical cut-offs and is analyzed only as a continuous measure; for this reason, it is excluded from Table 2). In some of the analysis below, we convert the classification systems into binary categories of "food secure" and "food insecure," for simplicity of presentation purposes. The final column in Table 2 presents these binary classifications.[5]

We now compare trends in food security across rounds, as shown by the different measures. Annex C shows more detailed results.

Table 3 and Figure 1 suggest that, as measured by any of the indicators, food security improved across the timeframe of the survey; the differences in prevalence between rounds are statistically significant across all

[5] Sensitivity analysis of the conclusions in the following pages, using different binary classification systems, was also performed. The results were substantively similar.

Table 4. Coefficient of variation of all measures, by round							
ROUND	CSI	rCSI	HFIAS	HHS	FCS	HDDS	SAFS
1	0.87	0.83	0.66	2.50	0.37	0.33	0.46
2	1.33	1.35	1.16	3.32	0.36	0.35	0.52
3	1.31	1.28	1.11	3.54	0.38	0.35	0.50
4	1.41	1.40	1.14	4.21	0.35	0.35	0.43
Pooled	1.20	1.18	0.99	3.23	0.37	0.35	0.49

Table 5. Percentage of households with highest possible food security score, by indicator and round							
ROUND	CSI	rCSI	HFIAS	HHS	FCS	HDDS	SAFS
1	15.3	17.9	12.0	81.5	0.0	0.0	30.8
2	46.9	50.2	39.4	89.0	0.0	0.0	38.6
3	38.5	44.1	38.9	90.3	0.0	0.0	44.7
4	47.0	49.1	41.5	92.6	0.0	0.0	33.7
Pooled	36.7	40.1	32.7	88.2	0.0	0.0	37.4

measures and all rounds at the $p<0.01$ level (with the exception of the HHS, where differences are significant at the $p<0.05$ level). Not only was each hunger and harvest season better than the previous one, but the second hunger season (Round 3) was only slightly different than the harvest season just before (Round 2), indicating that the positive across-year trend is dampening the potentially negative within-year seasonal effects.

The HFIAS measure shows the highest prevalence of food insecurity in the survey areas across all four rounds. The CSI and rCSI showed similar dynamics, suggesting that the questions in the rCSI are adequately picking up the same kinds of change in insecurity in this context that is detected by the CSI. The SAFS closely tracks the coping strategies measures. The household hunger scale HHS—which contains questions that capture the most extreme forms of insecurity—not surprisingly showed the lowest prevalence. The FCS measure showed the next lowest prevalence.

3. *What do the different indicators say about the distribution of food security outcomes within each round?*

The different measures imply different distributions of food security scores among households in a single survey round; Annex D shows the distribution of food insecurity as indicated by each index or scale score within each survey round. What do these measures tell us about inequality in food insecurity in each of the time frames studied? Measuring inequality is problematic when using index or scale scores or categorical measures (Makdissi and Yazbeck 2012); more traditional measures of inequality such as the Gini and Thiel indices cannot be calculated. However, the coefficient of variation (CoV) for each measure by round can give a general relative sense of the shapes of distributions, and from here inferences about inequality can be made[6] (Cowell 2009); the higher the CoV, the “flatter” and more unequal the distribution. Table 4 presents the CoV of each measure by round.

In the absence of an inequality “target,” the numbers themselves cannot suggest absolutely “high” or “low” variances, but we can make comparative statements between measures. The table suggests that the greatest variance exists with the HHS measure. The coping strategies measures and, to a lesser extent HFIAS, also display high variance in this dataset, although much less than HHS. The dietary diversity and self-assessment measures show much less variance, implying that how food security is measured matters greatly when evaluating inequality in a population: in this case, households are classified as having more equal levels of food security when using dietary diversity and self-perception measures than when using measures comprised of the coping strategies, behaviors, and attitudes explored by the CSI and HFIAS and the extreme conditions suggested by HHS. In general, variance in the sample increases as food security outcomes improve across rounds: inequality within the population appears to worsen when times are better for the population as a whole and vice-versa.

6 The coefficient of variation is simply the standard deviation normalized by dividing by the mean.

The distributions in Annex D suggest some reasons for the observed data in Table 4:

- CSI and rCSI. The relatively high variance shown by the CSI and rCSI is largely due to the high number of households reporting "0" values (meaning no coping behaviors reported), especially in the latter three rounds. Table 5 shows the percentage of households reporting the best possible food security state for each measure—which, in the case of the CSI, is a zero value. In the harvest season rounds, nearly half of households had a zero CSI score, quite an unusual distribution in a poor rural area: in comparable studies, zero scores were rare (Maxwell et al. 1999; Maxwell, Caldwell, and Langworthy 2008).[7]
- HFIAS. The HFIAS shows a similar distribution to the CSI in the final three rounds, also due to a large percentage of zero values. It is worth noting that if zero values are excluded, in the hunger seasons (especially Round 1) the HFIAS distributions tend more towards normal than do the CSI measures. (Note that the use of the term "normal" refers to the shape of the distribution, and does not imply "correct" or "true" inequality. Rather the relative distributions display only differences in the degree of inequality suggested by the various measures). The distribution of HFIAS scores may imply that the types of behaviors and attitudes captured by the HFIAS would, in stressful but not catastrophic times, be more normally distributed than the coping strategies in the CSI. Another explanation is that the algorithmic methodology used to determine the HFIAS categories leads to a high estimate of food insecurity, a topic explored later in this paper.
- HHS. An extremely high percentage of households have zero HHS scores in all rounds. This largely explains the high variance seen in the sample. As noted earlier, HHS is concerned with relatively severe conditions—complete lack of food in the household, going to sleep hungry, and going an entire day without eating—which were rare throughout the time period in the population studied.
- FCS and HDDS. The dietary diversity measures have a distribution much closer to normal than any of the others; in fact, the FCS in Round 1 and HDDS in Rounds 1 and 3 are the only measures that pass a formal Shapiro-Wilk test for normality. The shape of the distribution largely explains these measures' low variance relative to the others.
- SAFS. The comparatively low relative standard deviation for the self-assessment measure is largely an artifact of the few discrete values available as responses (1–5 for SAFS).

Table 6. Under- or over-estimation of food insecurity of the row measure relative to the column measure

	CSI	rCSI	HFIAS	HHS	FCS	SAFS
CSI	-	+0.12	-0.49	+0.31	+0.04	-0.11
rCSI	-0.12	-	-0.61	+0.29	+0.12	-0.23
HFIAS	+0.49	+0.61	-	+0.74	+0.57	+0.38
HHS	-0.31	-0.29	-0.74	-	-0.17	-0.38
FCS	-0.04	-0.12	-0.57	+0.17	-	-0.22
SAFS	+0.11	+0.23	-0.38	+0.38	+0.22	-

The strong correlations among the measures as well as the depiction of generally similar food security trends over time suggest that they are indeed all useful measurements of food security. They do also have significant differences, however, reflected both in the correlation coefficients and in the distributions. These differences may reflect different dimensions of food security, a topic explored in more detail in Section B. First, however, we look at the implications of different approaches to the classification of households' food security status.

4. *How differently are households classified by each of the measures?*

a. Static classification

Annex E presents tables of the different food security measures, converted into categorical measures by their severity cut-offs. Annex F provides histograms of every categorical measure, as well as the HDDS continuous measure, by round. Annex G presents cross-tabulations of each indicator (presented in categories) with each other indicators noted in the study. Note the extent to which the cut-offs used to determine household food security status yield dramatically different prevalence estimates, even among indicators that are otherwise reasonably well correlated as continuous measures. summarizes Annex F by presenting the extent (i.e., the number of categories) by which the row food security measure over- or under-calculates food insecurity relative to the column measure[8]; note that this table represents

[7] We are fairly certain that the issue did not stem from faulty data collection; after the high prevalence of zeros in the first round, enumerators in subsequent rounds were retrained and asked to probe carefully and extensively when a household reported not a single manifestation of food insecurity.

[8] The procedure for constructing this value is as follows. First, for measures with unequal numbers of categories (see Table 2), categories are combined following the classification implied by the measure with fewer categories. Then, for each cell in which the row measure and column measure

data pooled across all rounds. For example, in Table 6 we see that CSI under-calculates food insecurity relative to HFIAS by about half a category (-0.49); in other words, for any given household we would expect to see a CSI food security score that indicates it is about half a category less food insecure (i.e., more food secure) than the corresponding HFIAS score. Similarly, HFIAS over-calculates food insecurity relative to HHS by almost three-quarters of a category (+0.74). While the relevance of, for example a score that shows three quarters of a category difference is not operationally clear from this depiction, it does aid in the comparison of the relative classifications of food security status among indicators.

Several conclusions emerge from this table:

- CSI and rCSI. As would be expected, CSI and rCSI classify households similarly; the CSI tends to portray a slightly more food insecure situation relative to the rCSI (+0.12). The CSI and rCSI both tend to show greater food insecurity as compared to the HHS and FCS, but less than the HFIAS, and slightly less than the SAFS measure.
- HFIAS and HHS. HFIAS will give the highest prevalence estimates of food insecurity, as might be expected as it includes less-severe manifestations, including psychological anxiety and food consumption preferences (although this may also be due to the algorithm for classification—a topic we discuss below). In contrast, the HHS depicts the lowest prevalence of food insecurity. Households are only identified as food insecure if their situation is quite severe, as the measure only counts the most extreme behaviors.
- FCS. After HHS, FCS is the least likely to identify a household as food insecure.
- Self-assessment. SAFS, the food security self-assessment measure, was more likely to classify households as food secure than all measures but the HFIAS.

Overall, we see that, while broadly correlated, different measures classify food-insecure households quite differently. A static classification of food insecurity by the different measures implies a hierarchy with respect to the estimation of food insecurity. HFIAS provides the highest estimate of food insecurity prevalence, followed by SAFS, CSI, rCSI, FCS, and HHS (HDDS was not evaluated, as it lacks a "food insecure" cut-off). As discussed in greater detail below, this difference in the degree to which a measure indicates food insecurity can help inform the choice of measure used, depending on the timing (e.g., season and emergency onset) and purpose (e.g., development or relief) of data collection and intervention. At the same time, this presents a puzzle for analysis that relies on more than one indicator: Which one is "correct?"

agree on food security status—that is, for matched pairs—the observations are dropped. For each cell in which the row measure over-calculates the degree of food insecurity (again, relative to the other measure) by one category, the number of such observations are simply summed (i.e., given a weight of 1 and summed); when the over-calculation is two categories, the number of observations are multiplied by 2 and summed; and so on. The same is done for under-calculation, except that the number of such observations is multiplied by -1 and summed, by -2 and summed, etc. The totals for each group are then summed and divided by the total number of observations in question (i.e., the surveys in which food security scores were present for both measures).

b. Dynamic correlation

We can also look at "dynamic correlation": whether different food measures portrayed movement of households in and out of food insecurity similarly across the four rounds. Again, for purposes of simplicity, we reclassified households using the binary categories "food secure" and "food insecure"; see Table 2 for the original categories and how these were converted into binary categories.

Tables 7 and 8 show two ways of looking at this issue. Table 7 shows "net movement into food security": as identified by each measure, the number of households who moved from food insecurity between Round 1 and Round 4 minus the number who were originally food secure but fell into food insecurity, taken as a percentage of the original number of food insecure households. For example, CSI identifies 144 households as food insecure; of these, 87, or 60.4 percent (Column A), were food secure in Round 4: they moved into food security.[9] Conversely, 26 of the 156 households who were originally food secure, or 16.7 percent (Column B), fell into food insecurity by Round 4. Thus the net number of households who moved into food security is 61: 87 - 26 = 61; that number represents 42.4 percent of the original 144 food-insecure households (Column C).

Table 7 shows dynamic differences as well between the measures. There are significant differences between in the indicators in Column A's gross movement into food security: FCS has the lowest rate (i.e., is most likely to depict persistent food insecurity), followed by HFIAS. CSI, rCSI, SAFS, and HHS show relatively high movement into food security. The converse—the measurement of falling into food insecurity (Column B)—follows a different pattern. According to HFIAS, SAFS, and CSI, between one-sixth and nearly one-fourth of food secure households became food insecure (although the number of households originally identified as food insecure was much greater for the latter two indicators). Few households became food insecure according to HHS and FCS. Net totals (Column C) roughly mir-

[9] Note that Round 1 was completed during a hunger season and Round 4 a harvest season, and so some of the movement into food security is likely to be temporary.

Table 7. Net movement into food security from Round 1 to Round 4, by indicator			
	A. Gross movement into food security	B. Gross movement into food insecurity	C. Net as percentage of those originally food insecure and became food secure
CSI	60.4% (87/144)	16.7% (26/156)	42.4% (61/144)
rCSI	60.7% (71/117)	12.0% (22/184)	41.9% (49/117)
HFIAS	38.6% (102/264)	23.7% (9/38)	35.2% (93/264)
HHS	100.0% (18/18)	2.8% (8/284)	55.6% (10/18)
FCS	21.6% (16/74)	1.8% (4/227)	16.2% (12/74)
SAFS	55.1% (76/138)	17.7% (29/164)	34.1% (47/138)

Table 8. Dynamic correlation between indicators of movement into food security between Rounds 1 and 4							
	CSI	rCSI	HFIAS	HHS	FCS	SAFS	Corroboration of row measure by all other measures
CSI	-	65.5%	51.7%	4.6%	5.7%	43.7%	34.3%
rCSI	80.3%	-	52.1%	3.9%	7.0%	43.7%	38.3%
HFIAS	44.1%	36.3%	-	1.1%	2.0%	36.3%	24.3%
HHS	22.2%	27.8%	16.7%	-	22.2%	27.8%	23.3%
FCS	31.3%	37.5%	12.5%	11.5%	-	25.0%	26.3%
SAFS	50.0%	40.8%	48.7%	1.8%	5.3%	-	30.3%
Corroboration of all other measures by column measure	52.7%	45.5%	46.3%	6.0%	5.9%	39.1%	

rored the gross movement into food security of Column A, with the exception of SAFS, for which the amount of households falling into food insecurity reduced net movement considerably.

Table 8 presents a more comparative picture of the relative dynamics implied by each indicator. Of the total households that the row measure reported as having moved into food security between Rounds 1 and 4, each cell represents the percentage of households that the column measure corroborates as having moved into food security. For example, of all the households that the CSI measure reports as having moved into food security (first row), the SAFS measure corroborates the movement into food security rate of 43.7 percent of them (second to last column). Similarly, of all the households that HFIAS reports as having moved into food security, FCS corroborates just 2.0 percent of these. Thus the matrix provides a rough "dynamic correlation matrix" between Rounds 1 and 4 of the binary outcome of food security/insecurity.

Table 8 also provides convenient summary measures in the final column and row. The totals in the final column could be seen as a kind of "movement into food security" implied by a measure: the extent to which all other measures corroborate the claim of movement into food security by a given measure. For example, when rCSI claims that a household moved into food security, in a total of 38.3 percent of possible instances the claim is corroborated, with most of the disagreements coming from HHS and FCS. In contrast, when HHS claims movement into food security, the other measures only corroborate this claim 23.3 percent of the time. Note that CSI and rCSI are most often corroborated by the other measures.

The totals in the final row, meanwhile, capture the "total dynamic correlation" of a given column measure to the full set of other row measures; it is the percentage of times that this given column measure corroborates the instances of households' movement into food security identified by all the row measures combined. Thus we see that CSI, rCSI, and HFIAS tend to corroborate the full set of measures more than the others—corroborating 52.7 percent, 45.5 percent, and 46.3 percent, respectively, of the instances of movement into food security identified by the other measures. HHS and FCS are at the other extreme, corroborating just 6.0 percent and 5.9 percent of the claims of movement into food security by other measures. It is worth noting that the choice of cut-offs used in the development of the categorical indicator for each of these measures was ultimately subjective. Some of the differences observed between among indicators would be increased or decreased by altering the cut-offs. While some sensitivity analysis was done on the cross-corroboration matrices by altering indicator cut-offs, it would be useful to do further work along these lines, including identifying which cut-points on different indicators ultimately yield similar static and dynamic classifications.

B. What dimensions of food insecurity do each of these measures capture?

The World Food Summit definition of food security, adopted and widely promoted by FAO, contains several key elements, or dimensions, each of them difficult enough to measure separately, but nearly impossible to measure in totality using any single indicator. The definition is "Food security exists when all people, at all times, have physical and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life" (World Food Summit 1996). As described by Coates (2013), identifiable "elements" or "dimensions" include:

- Quantity (sufficiency): "…physical and economic access to sufficient…"
- Quality (diversity): "…nutritious food that meets dietary needs…"
- Acceptability: "…preferences…"
- Safety: "safe…food"
- Stability: "…at all times…"

The dimension of "stability" is crosscutting. According to this understanding, "the experience of food insecurity can be isolated from potential causes (i.e., lack of availability, lack of access) and potential consequences, both nutritional and non-nutritional, and can be considered within a framework of risk that could jeopardize the secure achievement of the first four elements" (Coates 2013).

In the sections below we draw on this standard definition of food security and use a combined theoretical and empirical approach to examine the extent to which the indicators examined in this study appear to capture one or more of the five identified dimensions. Each of the component questions of the seven indicators used in this study is listed in Table 9. The last column of the table suggests the dimension each measure or constituent primarily describes, based on the conceptual categorization above.

We then ran correlations between all of the measures and sub-questions (Annex H). To understand whether these measures and sub-questions are grouped into dimensions, we used the correlation coefficients to represent distance on a network diagram. Network diagrams allow correlation matrices to be represented in a visual, intuitive manner. In Figure 2, each node is one of the measures/sub-questions, and the edges (distances) between them represent the correlation coefficients; the higher the correlation coefficient, the shorter the edge. Thus groups of variables that are highly correlated appear as clusters in the diagram, while weakly correlated variables are at the borders of the network. Each dimension hypothesized in Table 9 is coded by color in Figure 2: the blue variable names represent those indicators in the "stability" group, red "acceptability," green "quantity," and yellow "quality/diversity."

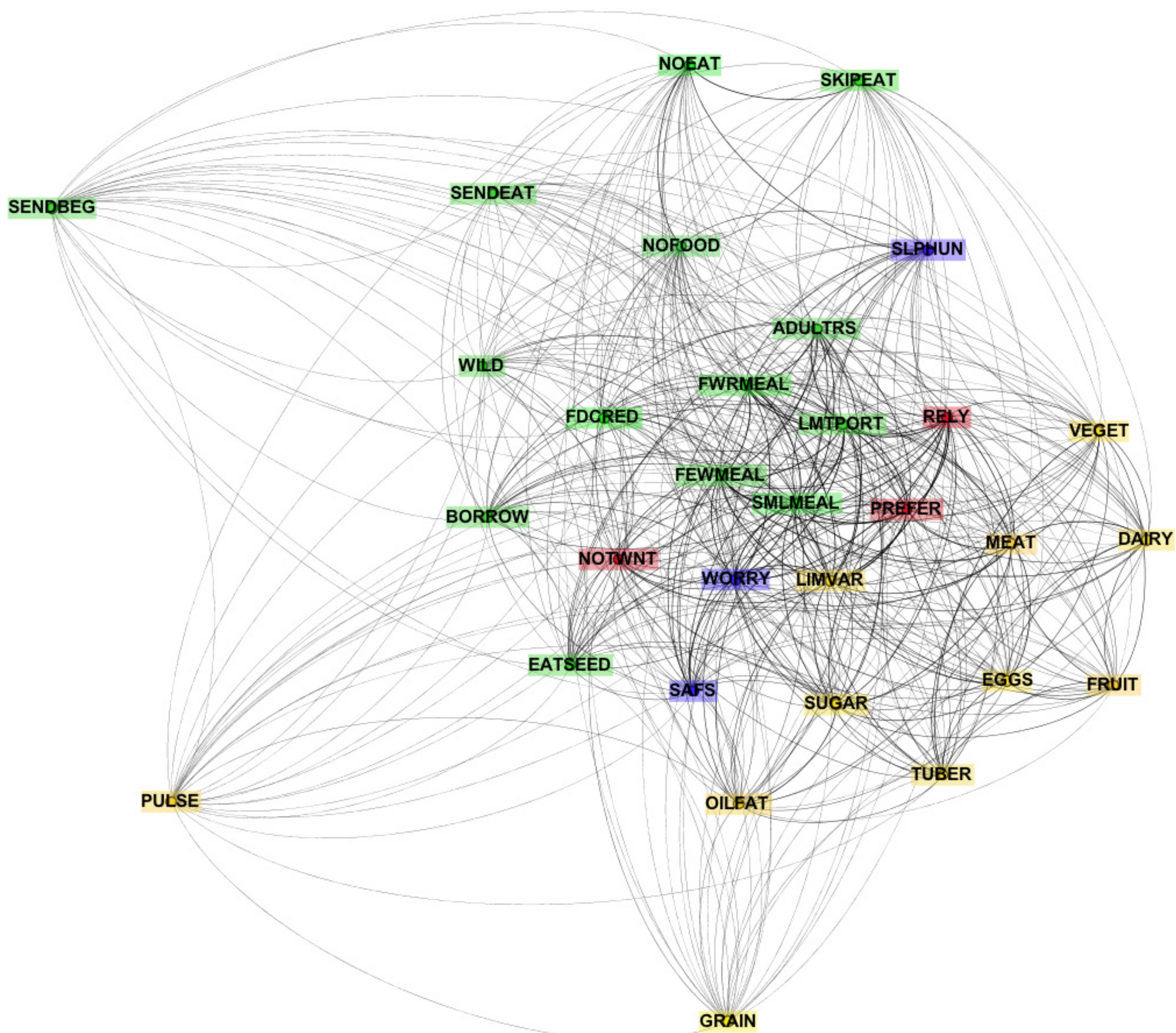
The clustering is evident with respect to the quality/diversity (yellow) questions and the food quantity (green) questions. The remaining stability (red) and acceptability (blue) are less clustered, with some outliers: SLPHUN (going to sleep hungry) from the vulnerability group and NOTWNT (eating foods not wanting to eat) from the preferences group. SENDBEG (sending household members to beg), PULSE, GRAIN, and to a lesser extent SKIPEAT (skipping entire days without eating) are relative outliers, the former two because they are rare extreme behaviors and the latter two because they are very common behaviors. Both characteristics will tend to weaken correlation with other measures.

Part of the reason that the stability measures do not correlate so well might be that the sub-questions selected to represent that dimension only poorly represent the concept. But it may also be the case that stability is not a stand-alone dimension per se in the same way that quantity, quality/diversity or preference is: that is, stability may be an important component of each of those dimensions over time, rather than a stand-alone. Stability (or its inverse, vulnerability) is probably the most difficult element of food insecurity to capture.

Table 9. Indicators and component questions			
Indicator	NAME	Question or specific measure	Dimension
SAFS	SAFS	Self-assessed food security during past 30 days	Stability
CSI/rCSI	RELY	In the past month, how often has the HH had to rely on less preferred or less expensive food?	Acceptability
CSI/rCSI	BORROW	In the past month, how often has the HH had to borrow food, or rely on help from a relative?	Quantity
CSI	FDCRED	In the past month, how often has the HH had to purchase food on credit?	Quantity
CSI	WILD	In the past month, how often has the HH had to gather wild food, hunt, or harvest immature crops?	Quantity
CSI	EATSEED	In the past month, how often has the HH had to consume seed stock held for next season?	Quantity
CSI	SENDEAT	In the past month, how often has the HH had to send HH members to eat elsewhere?	Quantity
CSI	SENDBEG	In the past month, how often has the HH had to send HH members to beg?	Quantity
CSI/rCSI	LMTPORT	In the past month, how often has the HH had to limit portion size at mealtimes?	Quantity
CSI/rCSI	ADULTRS	In the past month, how often has the HH had to restrict consumption by adults in order to allow children to eat?	Quantity
CSI/rCSI	FWRMEAL	In the past month, how often has the HH had to reduce the number of meals eaten in a day?	Quantity
CSI	SKIPEAT	In the past month, how often has the HH had to skip entire days without eating?	Quantity
HFIAS	WORRY	In the past 30 days, how often did you worry that your HH would not have enough food?	Stability
HFIAS	PREFER	In the past 30 days, how often were you or any HH member not able to eat the kinds of foods you preferred?	Acceptability
HFIAS	LIMVAR	In the past 30 days, how often did you or any HH member have to eat a limited variety of foods?	Quality/ diversity
HFIAS	NOTWNT	In the past 30 days, how often did you or any HH member have to eat foods you did not want to eat?	Acceptability
HFIAS	SMLMEAL	In the past 30 days, how often did you or any HH member have to eat a smaller meal than you felt like you needed?	Quantity
HFIAS	FEWMEAL	In the past 30 days, how often did you or any HH member have to eat fewer meals in a day?	Quantity
HFIAS/ HHS	NOFOOD	In the past 30 days, how often was there ever no food in your HH?	Quantity
HFIAS/ HHS	SLPHUN	In the past 30 days, how often did you or any HH member go to sleep at night hungry?	Stability*
HFIAS/ HHS	NOEAT	In the past 30 days, how often did you or any HH member have to go a whole day without eating?	Quantity
FCS/ HDDS	GRAIN	In the past month, how often has the household eaten any food made from grain?	Quality/ diversity
FCS/ HDDS	TUBER	In the past month, how often has the household eaten any tubers?	Quality/ diversity
FCS/ HDDS	PULSES	In the past month, how often has the household eaten any pulses?	Quality/ diversity

FCS/ HDDS	VEGET	In the past month, how often has the household eaten any vegetables?	Quality/ diversity
FCS/ HDDS	FRUIT	In the past month, how often has the household eaten any fruits?	Quality/ diversity
FCS/ HDDS	MEAT	In the past month, how often has the household eaten any meat, fish?	Quality/ diversity
FCS/ HDDS	EGGS	In the past month, how often has the household eaten any eggs?	Quality/ diversity
FCS/ HDDS	DAIRY	In the past month, how often has the household eaten any dairy products?	Quality/ diversity
FCS/ HDDS	SUGAR	In the past month, how often has the household eaten any sugar or honey?	Quality/ diversity
FCS/ HDDS	OILFAT	In the past month, how often has the household eaten any oils, fat, or butter?	Quality/ diversity
* This question could arguably also be classified as a "sufficiency"-related item, and will be examined as such in further analysis.			

Figure 2. Network diagram of correlation coefficients of measures and sub-questions



C. How can measures be combined or used in complementary ways to yield a more multidimensional picture of a household's food insecurity situation?

In the first sub-section below, we use the network diagram from Section B to propose a cross-classification measure. In the second sub-section, we compare our results with another multi-dimensional indicator of food security—the Integrated Phase Classification (IPC Partners 2012). In the third sub-section, we construct a new multi-dimensional indicator of food security based on elements of all the measures we have analyzed in this paper.

1. Cross-classification

Two dimensions in the definition of food security are distinct in both conceptual and empirical terms—the elements of quality (diversity) and quantity (sufficiency). A simple measure to capture more of the dimensions of food security therefore might be proposed simply by combining the two indicators that best capture these two dimensions. CSI and HFIAS are composed mostly of questions that reflect the "quantity" dimension of food security (though both contain elements of acceptability and stability); FCS is composed mostly of questions that reflect the diversity (quality) dimension. Annex J depicts the scatterplots of CSI with HFIAS, CSI with FCS, and HFIAS with FCS, with the cut-offs or thresholds for food secure, borderline, and food insecure in the CSI with FCS plot. It is clear that relying on only CSI, HFIAS or FCS alone would result in the misclassification of a substantial proportion of households as either food secure or food insecure by the measurement of the other indicator. However, cross tabulating them results in a new categorization. In conceptual terms, cross-classification would look like Figure 3. Tables 10 and 11 apply this conceptual approach to the empirical data, yielding different, relevant information from both indicators.

This conceptual cross-tabulation yields results that can be characterized as follows:

- Green: The low CSI/high FCS area is clearly the food-secure group likely to be doing satisfactorily in terms of both quality and quantity.
- Red: The high CSI/low FCS area is clearly the food-insecure group that might require some kind of assistance or intervention.
- Yellow: The intermediate group—those with low CSI but low FCS, those with high CSI but high FCS, or those genuinely intermediate on both—constitutes an in-between category requiring special operational considerations, discussed at greater length below.

Table 10 suggests that nearly 27 percent of households would be misclassified by using only one or the other indicator rather than the combination of both. Table 11 cross-classifies FCS and HFIAS with similar results except that, with HFIAS's higher prevalence estimates of food insecurity, the intermediate category contains a larger proportion of the total sample.

The obvious point arising from this kind of combined indicator is that there is less error (defined as a measure-

Figure 3. Conceptual rendering of cross-classification between CSI/HFIAS and FCS

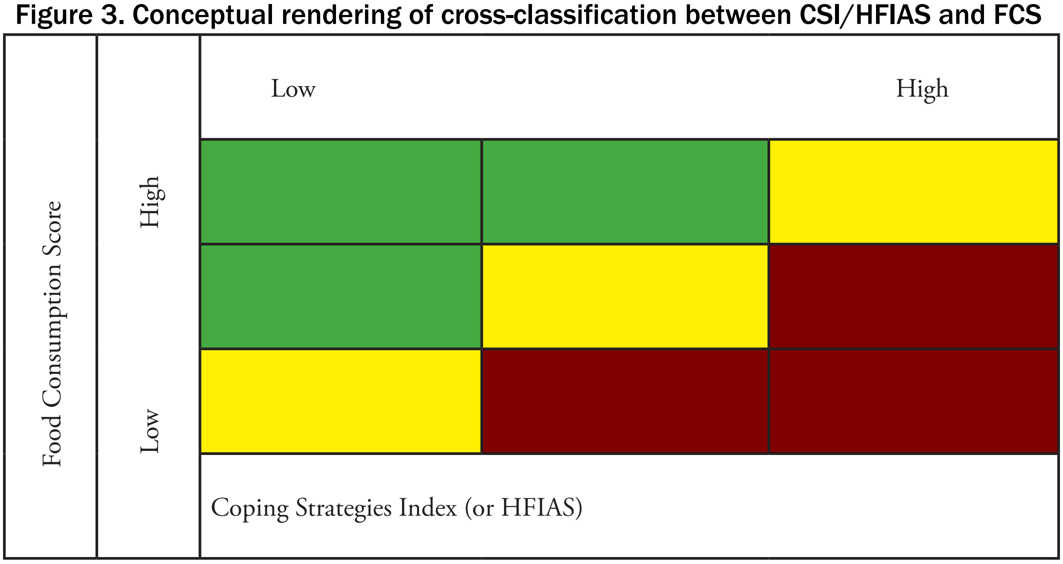


Table 10. Cross-classification of CSI and FCS categories, all rounds pooled

		FCS category			Total
		Acceptable	Borderline	Poor	
CSI category	Food secure (0–2)	39.1% (492)	2.4% (28)	0.3% (3)	41.7% (483)
	Mildly food insecure (3–12)	20.2% (234)	3.6% (42)	0.6% (7)	24.5% (283)
	Moderately/severely food insecure (13+)	22.8% (264)	8.3% (96)	2.7% (31)	33.8% (391)
Total		82.1% (950)	14.3% (166)	3.5% (41)	100.0%

Table 11. Cross-classification of HFIAS and FCS categories, all rounds pooled

		FCS category			Total
		Acceptable	Borderline	Poor	
HFIAS category	Food secure	33.4% (387)	1.0% (12)	0.1% (1)	34.5% (400)
	Mildly food insecure	5.2% (60)	0.9% (10)	0.2% (2)	6.2% (72)
	Moderately/severely food insecure	43.6% (506)	12.4% (144)	3.3% (38)	59.3% (688)
Total		82.2% (953)	14.3% (166)	3.5% (41)	100.0%

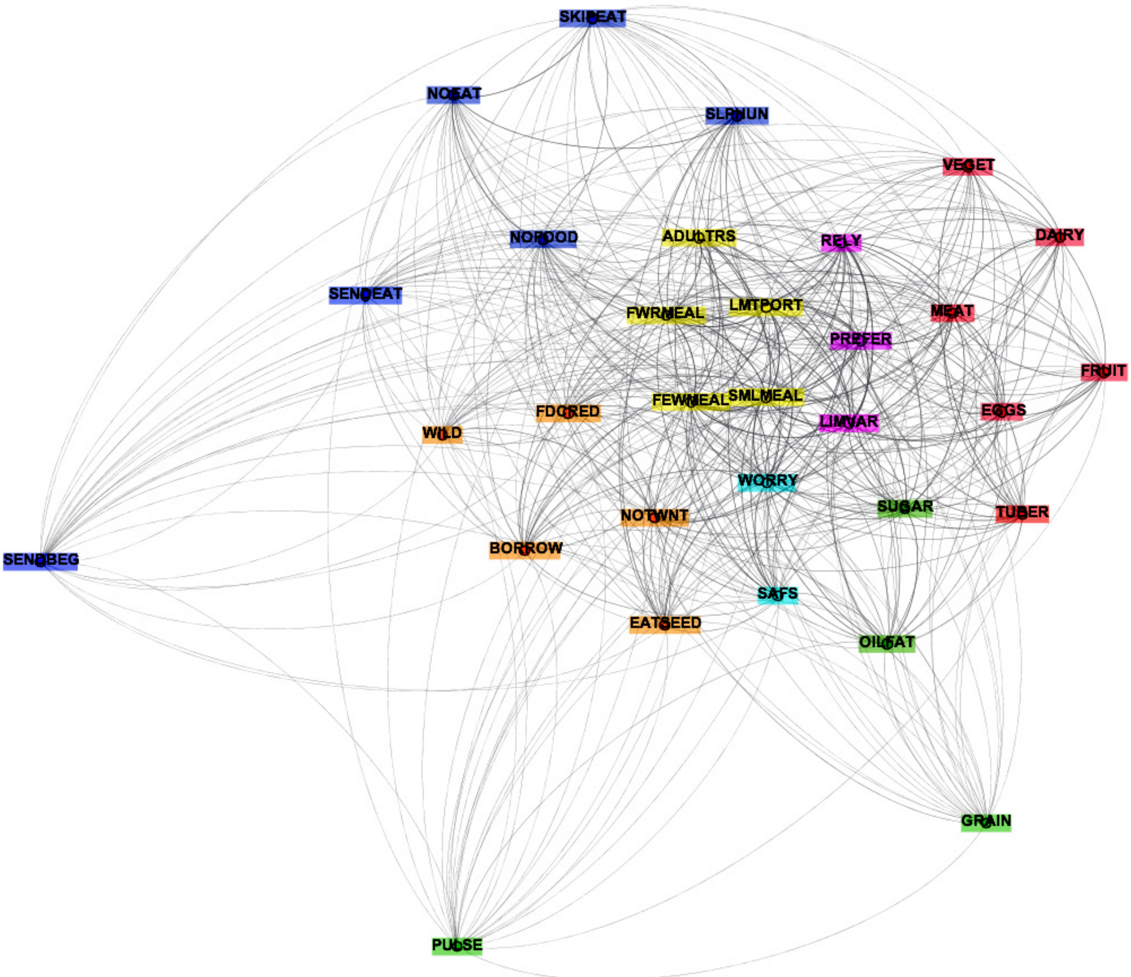
ment more consistent with the definition of food security) in identifying the food-insecure and food-secure groups by using two indicators. The practical implications of this for analysis and programming are discussed in the final section of the paper.

2. *Comparing with other aggregate indicators*

The obvious aggregate measure with which to compare the results of these indicators (individually and collectively) is the Integrated Phase Classification (IPC Partners 2013). The IPC maps corresponding to the times of the survey rounds are found in Annex K. For all rounds of the survey, the areas in which the survey was conducted are listed as Phase 2 (stressed). This would imply that IPC analysis noted no major change in food security status over time in the areas surveyed. Our results, on the other hand, show a relatively poor food security situation during Round 1, with improvements over time in the following three rounds, with a slight (and expected) reversal during the hungry season of Round 3. The IPC maps in fact show a poor situation in many of the surrounding woredas in Round 1, which would correspond to our findings. They also show a significant worsening of conditions in surrounding areas following the main 2012 harvest (Round 4), which our results do not. Our results are not intended to represent the whole woreda. They were randomly sampled; but being statistically representative of the two woredas was not an objective of the study. However, broadly speaking, our results are fairly similar to IPC analysis, if one considers that overall the situation was worse in the hungry season of 2011 (the period of the major crisis in the Greater Horn of Africa).

In terms of how these results would inform future iterations of the IPC (which includes various measures compared here: the CSI, FCS, and, in some cases, HFIAS, HHS, and rCSI), further research that is representative of different geographic areas would be necessary. Broadly speaking, however, the range of indicators shown here would seem to be indicative of "stressed" or Phase 2 (Rounds 2, 3, and 4) and "crisis" or Phase 3 (Round 1). To use these indicators to determine IPC categories in an algorithmic method would require decisions both about thresholds (what level of each indicator must be reached to qualify for Phases 2, 3, 4, and 5), and prevalence (what proportion of the population has to reach

Figure 4. Network diagram of correlation coefficients, color-coded by food security dimension



the threshold before that phase can be declared). This study did not have adequate empirical evidence on IPC phases to make such a determination.

3. *An algorithmic approach*

We constructed a multidimensional indicator using an approach that first extracted categories empirically using a network modularity approach and then refined these inductive results with theory.

A modularity algorithm with a pre-specified resolution of 0.77 was then run to group neighborhoods of nodes into clusters (Blondel et al. 2008, Labiotte et al. 2009.) From this, we returned to theory to interpret the clusters as dimensions of food security. In Figure 4, each dimension is coded by color: rationing strategies (yellow), food-increasing behaviors (orange), dietary change preferences (purple), rare and severe behaviors (blue), vulnerability (light blue), high value foods (red), and daily foods (green). Note again the outliers SENDBEG (sending household members to beg), PULSE, GRAIN, and, to a lesser extent, SKIPEAT (skipping entire days without eating).

How then can this diagram, depicting different dimensions of food security, be operationalized? Below, we propose a methodology for constructing a multidimensional food security measure.

First, we used the correlation matrix in Annex H, and depicted visually in Figures 3 and 4, to eliminate redundant questions—those that seemed to be picking up similar aspects of food security. To determine candidates for exclusion, we looked at each variable's average correlation coefficient—the simple mean of its correlation coefficients with all other variables; these are the variables whose information is most likely to be gathered by other variables. Annex I presents the results from strongest to weakest. Some candidates for elimination emerged. The pair SMLMEAL and LMTPORT

Table 12. Questions used in multidimensional indicator		
Indicator	Name	Food Security Dimension
SAFS	SAFS	Stability
HFIAS	WORRY	Stability
CSI/rCSI	LMTPORT	Rationing strategies
CSI/rCSI	FWRMEAL	Rationing strategies
HFIAS	PREFER	Dietary change preferences
HFIAS	LIMVAR	Dietary change preferences
HFIAS	NOTWNT	Dietary change preferences
CSI/rCSI	BORROW	Food-increasing behaviors
CSI	FDCRED	Food-increasing behaviors
CSI	WILD	Food-increasing behaviors
CSI	EATSEED	Food-increasing behaviors
FCS/HDDS	VEGET	High value foods
FCS/HDDS	FRUIT	High value foods
FCS/HDDS	MEAT	High value foods
FCS/HDDS	EGGS	High value foods
FCS/HDDS	DAIRY	High value foods
FCS/HDDS	GRAIN	Daily foods
FCS/HDDS	PULSE	Daily foods
CSI	SENDEAT	Rare and severe behaviors
CSI	SENDBEG	Rare and severe behaviors
CSI	SKIPEAT	Rare and severe behaviors
HFIAS/HHS	NOFOOD	Rare and severe behaviors
HFIAS/HHS	SLPHUN	Rare and severe behaviors
HFIAS/HHS	NOEAT	Rare and severe behaviors

and the pair FEWMEAL and FWRMEAL are somewhat redundant in their questions; we choose to eliminate the HFIAS questions SMLMEAL and FEWMEAL. RELY is strongly correlated with PREFER, so we excluded RELY as well. ADULTRS is strongly correlated with LMTPORT, SMLMEAL, and FWRMEAL, so we excluded ADULTRS as well. SUGAR, TUBER, and OILFAT were removed because of concerns over their cross-contextual applicability and nutritional importance in relation to the other dietary measures.

Table 12 presents the streamlined set of 25 questions, categorized by dimension. In future analyses, Principal Components Analysis or Rasch modeling or both will be used to compare the results presented here to those achieved by these traditional scaling methods.

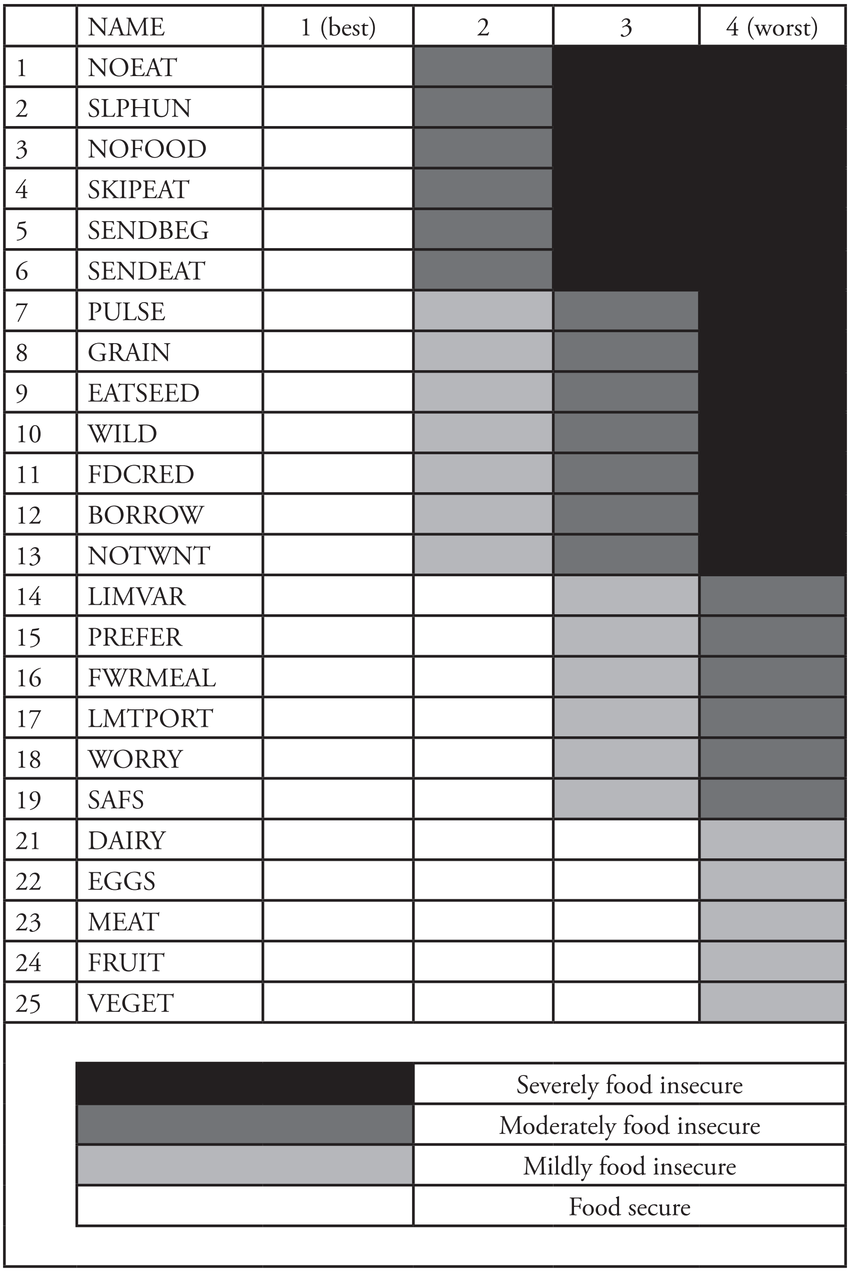
We now propose an algorithmic means of classification based on the HFIAS category construction methodology for using the questions to divide households into categories of "food secure," "mildly food insecure," "moderately food insecure," and "severely food insecure."

With the exception of SAFS, each question can be answered from a set of four responses, borrowed from the HFIAS: often, sometimes, rarely, and never, with the exact meaning of these responses differing by question. For SAFS, the responses would correspond to severely food insecure (4 or worst), moderately food insecure (3), mildly food insecure (2), and food secure (1).

The algorithmic method is depicted in Figure 5. The worst or second-worst response to any of the first six questions (from NOEAT to SENDEAT) results in an automatic classification into the overall "severely food insecure category"; any worst response to any of the next seven questions (from PULSE to NOTWNT) results in the same; and so on.

In this illustrative, exploratory approach, the classification of "food secure" can only result if the optimal response is given for questions 1–13, at least the optimal or second-best response for questions 14–19, and anything but the worst response for questions 20–25.

Figure 5. Algorithmic classification of food security



Using this algorithm, the Multi-dimensional Food Security Indicator (MFI) shows dynamics similar to the other food security measures, and is generally quite sensitive in the portrait it paints of food insecurity (Figures 6 and 7).

The advantages of this illustrative MFI over the other stand-alone measures are threefold. First, it is data-rich, capturing multiple dimensions of food security. Second, by taking this broad approach, its range of sensitivity to food insecurity is wider; it is able to pick up both milder food insecurity manifestations as well as severe outcomes. Third, by taking an algorithmic approach, it avoids the pitfalls of arbitrary quantitative food security cut-offs along a raw score. Though a choice of cut-offs is ultimately still subjective, an algorithmic approach demands careful conceptual attention (compared with the standard definition) to the implications of each answer for overall food security classification.

For this indicator to be even more operationally useful, each category currently labeled according to degrees of food security severity could, and should, be further described according to the types of conditions experienced by households classified into the category. This additional step would produce more-transparent results than those stemming from a purely aggregated classification system labeled according to status along a single dimension of severity.

Figure 6. MFI dynamics, by round

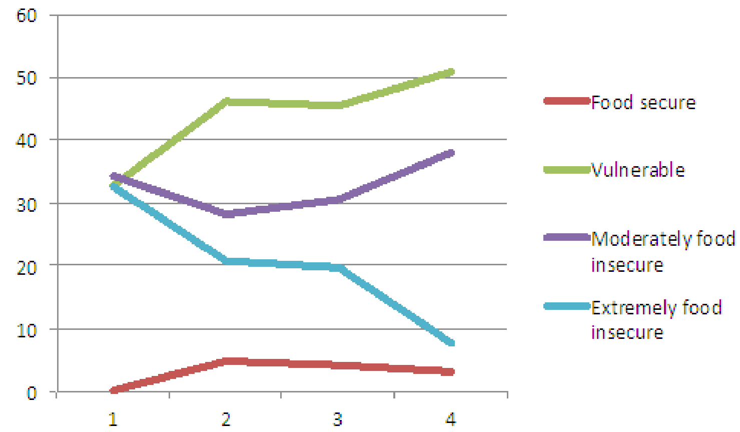
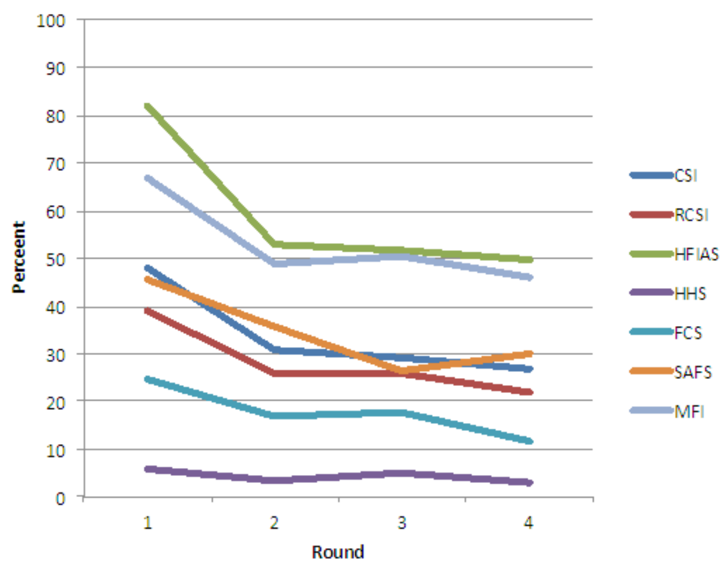


Figure 7. MFI dynamics in comparison to other measures



V. Conclusions and Discussion

Several conclusions can be drawn from this analysis. First, all of these indicators are well correlated, which suggests that they are all capturing some element of the multidimensional notion of food security (or food insecurity). But the network analysis suggests that not only does each existing indicator capture different elements of food insecurity, even the component questions of each, to some degree, capture differing elements of a more-nuanced understanding of food security. FCS and HDDS tend to capture quality and diversity (although FCS is calibrated to capture an element of quantity as well, its components are clearly more correlated with each other than with other elements that get at quantity). CSI and rCSI tend to capture the element of quantity or sufficiency. HFIAS captures a mix of sufficiency and psychological factors. HHS captures the most extreme manifestations of insufficiency.

In practice, these indicators are often used interchangeably—or one or another of the indicators is favored for reasons related more to their institutional evolution rather than to some characteristic of the method itself. But the evidence presented here suggests that indicators should be chosen to capture different elements of food security. The evidence presented here also strongly suggests that relying on only one measure of food security in analysis and program design runs the risk of serious misclassification by relying on a measure that captures some, but not all, of the dimensions of food insecurity inherent in the definition. However, in practice, it is often difficult to know which indicator is best suited to which situation. On balance, results from the CSI, rCSI, and HFIAS tend to be better corroborated by all the measures than the other indicators, but the categories derived from their scores also give very different estimates of prevalence due to the different methods used for classification.

It follows therefore that combining indicators can improve the measurement of food insecurity. As demonstrated by Tables 10 and 11, combining two measures—one that broadly captures diversity (quality) and one that broadly captures sufficiency (quantity) improves the classification of households. At a minimum, it reduces the number of potential false positives that would result from relying solely on the FCS measure, or the number of potential false negatives from relying solely on CSI or HFIAS. Those in an "intermediate" group that are food insecure according to one dimension but not another require further exploration of the causes and consequences of their particular situation as well as potential special operational consideration.

The field experience of this research strongly suggests that cross-classifying households through multiple measures does little to increase either costs of data collection and analysis, or the time required to request such information from respondents. Time and financial constraints are often cited as reasons for less-than-adequate analysis or impact assessment. The collection of CSI, rCSI, HFIAS, HHS, and the self-assessment measure all take only a few minutes per household, and recording the same data and converting it to the multidimensional indicators illustrated here also requires little time. Thus the marginal cost of adding such indicators to both assessments and monitoring or evaluation processes cannot be cited as a genuine justification for having poor indicators of food security.

A question arises from this discussion about whether indicators that give a higher prevalence estimate are more sensitive to changes or, alternatively, whether they simply consistently produce higher estimates of food insecurity. Several indicators routinely produced the highest estimates of food insecurity in this study, and they did so whether food insecurity was trending up (Round 1) or trending down (Rounds 2–4). In this study, the HFIAS and the MFI both produced the highest estimates of food insecurity, but in the absence of a gold standard for external validation it is impossible to say whether these were "over"-estimates. The closest thing to a gold standard is the food security definition; if one adheres to this standard when developing measures, individuals would experience food insecurity if they experienced any one of the manifestations described in the definition, at any time. However, a high prevalence of food insecurity as indicated by a measure may yield results that are much closer to the definition, but that may not be as operationally useful or as discriminating, especially in contexts where most households experience some form or another of food insecurity, even if only relatively mild, from time to time.

Much of the demand for indicators is to measure the impact of programs and policies, and to determine which households or groups are so badly off that they require some form of direct intervention (be it a social safety net or humanitarian response). Thus, the operational application of many of these measures may be towards the moderate to severe end of the spectrum. On the other hand, the HHS picks up only the most-severe behaviors and therefore tends to produce the lowest prevalence estimate for food insecurity—the HHS was an extremely important indicator for both assessment and monitoring during the worst phase of the Somalia famine in 2011 and 2012, but by 3–4 months into the response, HHS was estimating a very low prevalence of food insecurity, when dietary diversity measures were still showing a high prevalence (Hedlund et al. 2013). Unfortunately, other indicators reviewed here were not used in monitoring that crisis. All this implies the need to understand which indicators are best suited for which applications—in addition to using more than one.

Ultimately, however, the prevalence estimates that each indicator provides are a function not so much of the objective "truthfulness" of the indicators themselves (in their continuous quantitative formulation), as they are of the cut-off points assigned to the different categories and the ways in which categories are constructed. Assigning cut-off points to a continuous quantitative measure is usually a matter of analytical judgment and a matter of controversy—particularly about the extent to which such categorical cut-offs are universally applicable. Several studies reviewed above note that "global" categories may be seriously misestimating prevalence in certain local contexts (WFP 2010, Coates et al. 2007). For this reason, the CSI and rCSI field manual does not suggest generalized cut-offs—in fact, it states clearly that cut-offs should be based on location-specific criteria (Maxwell and Caldwell 2008).[10] Nevertheless, the operational demand is for indicators that can be put in place quickly and used without extensive contextualization. This paper has highlighted the dangers of such practices, but hasn't adequately addressed the conundrum of accurately categorizing continuous quantitative data. The way in which the algorithmic categorization was done for the HFIAS and the MFI may be one factor behind the observation that these two measures produce the highest prevalence estimates: a single instance of certain behavior would put the reporting household directly into the "food insecure" category, for example, even if other behaviors were not reported (see Figure 5). The categorization of food insecurity from continuous quantitative data—and the best method for conducting the categorization—remains an important area for further research.

Though the self-assessment questions were the least well-correlated to other measures, the SAFS provided results that were significantly correlated to the other indicators and trended in the same direction over time, raising the question of whether it might be simplest to merely ask people to categorize themselves. However, there are obvious disadvantages to relying solely on such an approach in assessments (or anything linked to assistance), especially the difficulty of measuring variance in both subjective understandings of food security and incentives for biased self-reporting.

This study also points out that conceptual and empirical approaches to identifying food security dimensions can be complementary. The cross-classification measure was based on a conceptual theorization of food security dimensionality while the network diagram approach utilized the data itself to separate variables into categories. Because the empirical method relies heavily on the user specifying the parameters for network modularity—

[10] And the cut-off points used in this paper are not intended for application in other contexts without location-specific considerations being taken into account.

i.e., how many categories are to be created—it cannot be used without an underlying conceptual framework. The differences between the empirical and conceptual approach, however, helped to suggest directions for future refinements.

The empirical results of this study suggest only general links to IPC analysis—they do not specifically answer the question about what thresholds within any of these indicators relate to different phase classifications. To empirically derive such thresholds—even for a given context such as the one studied here—would require measurement over a longer and more varied time period (i.e., a period over which different types and intensities of shocks were experienced) and a sampling strategy that would allow extrapolation of results to a specified geographic area. But more importantly, the data would have to capture a much broader range of food security outcomes—not a variable that can be controlled by the researchers. Further research and analysis over a much broader range of contexts will be required to empirically map out precise linkages between the thresholds in individual indicators and IPC phase classifications.

References

Barrett, Christopher. 2010. "Measuring Food Insecurity." Science. 327: 825–828.

Becquey, E., Y. Marin-Prevel, and P. Traissac. 2010. "The Household Food Insecurity Access Scale and an Index-Member Dietary Diversity Score Contribute Valid and Complementary Information on Household Food Insecurity in an Urban West-African Setting." Journal of Nutrition. 140: 2233–2240.

Blondel, Vincent D., Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. "Fast Unfolding of Communities in Large Networks." Journal of Statistical Mechanics P10008 doi:10.1088/1742-5468/2008/10/P10008.

Cafiero, Carlo. 2012. "Advances in Hunger Measurement: Traditional FAO Methods and Recent Innovations." Rome: FAO.

Coates, Jennifer. 2013. "Build it Back Better: Deconstructing Food Security for Improved Measurement and Action." Global Food Security Journal. Available online 24 June 2013. Print publication forthcoming.

Coates, Jennifer, Beatrice L. Rogers, Patrick Webb, Daniel Maxwell, Robert Houser, and Christine McDonald. 2007. "Diet Diversity Study." Final report to the World Food Programme. Medford, MA: Friedman School of Nutrition Science and Policy, Tufts University.

Coates, Jennifer, Anne Swindale, and Paula Bilinsky. 2006. "Household Food Insecurity Access Scale (HFIAS) for Measurement of Household Food Access: Indicator Guide (v. 2)." Washington, D.C.: Food and Nutrition Technical Assistance, Academy for Educational Development.

Coates, Jennifer, and Daniel Maxwell. 2012. "Reaching for the Stars?: Universal Measures Of Household Food Security." UN FAO International Scientific Symposium on Food Security and Nutrition Measurement, Rome, February 2012.

Cowell, Frank A. 2009. Measuring Inequality. Oxford: Philip Allan.

Deitchler, Megan, Terri Ballard, et al. 2010. "Validation of a Measure of Household Hunger for Cross-Cultural Use." Washington, DC, Food and Nutrition Technical Assistance II Project (FANTA-2), AED.

ELCSA Scientific Committee. 2012. "Escala Latinoamericana y Caribena de Seguridad Alimentaria (ELCSA): Manual de Uso y Aplicaciones." Rome, UN Food and Agricultural Organisation.

Fiedler, Jack. 2013. "Strengthening Household Consumption and Expenditures Surveys to Enable more Evidence-Based Nutrition Policies." International Scientific Symposium on Food and Nutrition Security information. FAO: Rome.

FANTA. 2006. "Household Dietary Diversity Score (HDDS) for Measurement of Household Food Access: Indicator Guide (Version 2)." Washington DC: Food and Nutrition Technical Assistance Project.

Food and Agriculture Organization. 2013. "Proceedings of the International Scientific Symposium on Food and Nutrition Security information: from Valid Measurement to Effective Decision Making." Rome, 17–19 January 2012. Rome: FAO.

Food and Agriculture Organization. 2010. "Guidelines for Measuring Household and Individual Dietary Diversity." 2010 (reprint 2013). Rome: UN Food and Agriculture Organization.

____________. 2006. "Food Security." Policy Brief No. 2. June. Rome: FAO.

Headey, Derek. 2011. "Was the Global Food Crisis Really a Crisis? Simulations versus Self-Reporting." IFPRI Discussion Paper. Washington DC: IFPRI

Headey, Derek. 2013. "The Impact of the Global Food Crisis on Self-Assessed Food Security." World Bank

Economic Review. Advance Access published January 7, 2013.

Headey, Derek, and Olivier Ecker. 2013. "Rethinking the Measurement of Food Security: From First Principles to Best Practice." Food Security. 5:327–343

Hoddinott, John, and Yisehac Yohannes. 2002. "Dietary Diversity as a Food Security Indicator." Washington, DC: IFPRI.

IFPRI. 2006. "Review and Validation of Dietary Diversity, Food Frequency and Other Proxy Indicators of Household Food Security." Rome: International Food Policy Research Institute.

IFPRI. 2008. "Validation of Food Frequency and Dietary Diversity as Proxy Indicators for Household Food Security." Report submitted to WFP. Washington DC: IFPRI.

IPC Partners. 2012. "IPC Technical Manual." Version 2.0. Rome: FAO.

Kennedy, Eileen. 2002. "Qualitative Measures of Food Insecurity and Hunger." Paper read at International Scientific Symposium on Measurement and Assessment of Food Deprivation and Undernutrition, 2002, Rome, Italy.

Lambiotte, R., J.C. Delvenne, and M. Barahona Laplacian. 2009. "Dynamics and Multiscale Modular Structure in Networks." arXiV:0812.1770 [physics.soc-ph].

Makdissi, Paul, and Myra Yazbeck. 2012. "Measuring Socioeconomic Health Inequalities in Presence of Multiple Categorical Information." Department of Economics Working Paper #1208, University of Ottawa.

Maxwell, Daniel, Clement Ahiadeke, Carol Levin, Margaret Armar-Klemesu, Sawadatu Zakaria, and Grace Mary Lamptey. 1999. "Alternative Food-Security Indicators: Revisiting the Frequency and Severity of 'Coping Strategies.'" Food Policy 24:411–429.

Maxwell, Daniel, and Richard Caldwell. 2008. The Coping Strategies Index: Field Methods Manual.

Maxwell, Daniel, Richard Caldwell, and Mark Langworthy. 2008. "Measuring Food Insecurity: Can an Indicator Based on Localized Coping Behaviors Be Used to Compare Across Contexts? Food Policy 33 (6):533–540.

Maxwell, Daniel. 1996. "Measuring Food Insecurity: The Frequency and Severity of "'Coping strategies.'" Food Policy 21 (3):291–303.

Smith, Lisa, Harold Alderman, and D. Aduayom. 2006. Food insecurity in Sub-Saharan Africa: New estimates from household expenditure surveys. Research Report 146. Washington DC: IFPRI.

Swindale, Anne, and Paula Bilinsky. 2006. "Development of a Universally Applicable Household Food Insecurity Measurement Tool: Process, Current Status, and Outstanding Issues." Journal of Nutrition 136:1449S–1452S.

Vaitla, Bapu, Girmay Tesfay, Megan Rousenville, and Daniel Maxwell. 2012. "Resilience and livelihoods change in Tigray, Ethiopia." Feinstein International Center report.

Watts, Michael. 1983. "Silent Violence: Food, Famine, and Peasantry in Northern Nigeria." Berkeley: University of California Press.

Webb, Patrick, Jennifer Coates, Edward A. Frongillo, Beatrice L. Rogers, Anne Swindale, and Paula Bilinsky. 2006. "Measuring Household Food Insecurity: Why It's So Important and Yet So Difficult to Do." Journal of Nutrition 136 (Supplement):1404S–1408S.

Weismann, Doris, John Hoddinott, Noora-Lisa Aberman, and Marie T. Ruel. 2006. Review and Validation of Dietary Diversity, Food Frequency and Other Proxy Indicators of Household Food Security. Washington, D.C.: International Food Policy Research Institute.

World Food Programme. 2009. Emergency Food Security Assessment Manual. Rome:WFP.

___________. 2010. "Validation Study of the WFP's Food Consumption Indicator in the Central American Context, with a Focus on Intra-household Sharing of Food." Rome: WFP.

Young, Helen, and Susanne Jaspars. 2006. "The Meaning and Measurement of Acute Malnutrition in Emergencies: A Primer for Decision-Makers." HPN Network Paper 56.

___________. 2009. Review of Nutrition and Mortality Indicators for the Integrated Food Security Phase Classification (IPC): Reference Levels and Decision-making. Inter-Agency Standing Committee Nutrition Cluster.

Feinstein International Center
Tufts University
114 Curtis Street
Somerville, MA 02144
USA
tel: +1 617.627.3423
fax: +1 617.627.3428
fic.tufts.edu