Food Policy 47 (2014) 107–116

Contents lists available at ScienceDirect

Food Policy

journal homepage: www.elsevier.com/locate/foodpol

How do indicators of household food insecurity measure up? An empirical comparison from Ethiopia

Daniel Maxwell [a,b,*], Bapu Vaitla [a], Jennifer Coates [b]

[a] Feinstein International Center, Tufts University, 114 Curtis Street, Somerville, MA, USA
[b] Friedman School of Nutrition Science and Policy, Tufts University, United States

ARTICLE INFO

Article history:
Received 11 November 2013
Received in revised form 21 March 2014
Accepted 14 April 2014

Keywords:
Food insecurity
Hunger
Indicators
Measurement
Ethiopia

ABSTRACT

Renewed emphasis on programs and policies aimed at enhancing food security has intensified the search for accurate, rapid, and consistent indicators. Measures of food security are urgently required for purposes of early warning, assessment of current and prospective status of at-risk populations, and monitoring and evaluation of specific programs and policies. Different measures are often used interchangeably, without a good idea of which dimensions of food security are captured by which measures, resulting in potentially significant misclassification of food insecure populations. The objective of this paper is to compare how the most frequently used indicators of food security portray static and dynamic food security among the same sample of rural households in two districts of Tigray State, Northern Ethiopia. Seven food security indicators were assessed: the Coping Strategies Index (CSI); the Reduced Coping Strategies Index (rCSI); the Household Food Insecurity and Access Scale (HFIAS); the Household Hunger Scale (HHS); Food Consumption Score (FCS); the Household Dietary Diversity Scale (HDDS); and a self-assessed measure of food security (SAFS). These indicators provide very different estimates of the prevalence of food insecurity, but are moderately well correlated and depict generally similar food security trends over time. We suggest that the differences in prevalence estimates, and in some cases the weaker than expected correlation, can be explained in three ways. First, the indicators differ in the underlying aspect of food security they attempt to capture. Second, each indicator is likely only sensitive within a certain severity range of food insecurity and these ranges do not always overlap. Third, categorization of the prevalence of food insecurity is strongly dependent on the choice of cut-off points. For valid reasons, "food insecurity" has no accepted gold standard metric against which individual indicators can be gauged, though without one it is difficult to say which indicator performs "best" in correctly and reliably identifying food insecure households. The implication is that using more than one indicator is advisable, and policy makers should be aware of what elements of food insecurity each indicator portrays.

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 once again prominent on the policy agenda today. This renewed emphasis has intensified the search for accurate, rapid and consistent indicators of food security for purposes of early warning, assessment of current and prospective status of at-risk populations, and for monitoring and evaluating specific programs and policies. Often these undertakings involve comparing food security status across dissimilar contexts.

The commonly accepted notion of food security follows from the 1996 World Food Summit definition: "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" (FAO, 1996). How to capture the full range of dimensions embedded within this definition—quality, quantity, acceptability, safety, and stability—has long been a subject of debate. Barrett (2010) reviews a variety of measures that capture "objective" aspects of the definition—dietary intake, expenditure, and health indicators—as well as a variety of measures that capture the more "subjective" aspects of the definition, such as perceived adequacy of consumption, exposure to risk and the cultural acceptability of foods. A newer generation of food security indicators has emerged from a demand for scientifically valid, direct, simple approaches to capturing one or more aspects of the "access" dimension of

* Corresponding author at: Feinstein International Center, Tufts University, 114 Curtis Street, Somerville, MA, USA. Tel.: +1 617 627 3410.
E-mail addresses: daniel.maxwell@tufts.edu (D. Maxwell), bapu.vaitla@gmail.com (B. Vaitla), jennifer.coates@tufts.edu (J. Coates).

http://dx.doi.org/10.1016/j.foodpol.2014.04.003
0306-9192/

food insecurity. Several of these are measures reviewed in this study.

Food security indicators are often used interchangeably, increasing the likelihood that the number of food insecure individuals or households may be misestimated by applying single or multiple measures without a solid understanding of which dimensions they are capturing and how variations in the indicators' construction affects the interpretation of the results (Coates, 2013). Major problems can ensue when different measures give very different prevalence estimates for food insecurity in the same place at the same time (WFP, 2012; Maxwell et al., 1999; de Haen et al., 2011). For example, reviews of food security interventions in the Somalia famine of 2011 showed extremely high levels of food insecurity across various different indicators or groups of indicators (Hedlund et al., 2013). Some of these prevalence estimates fell fairly quickly after humanitarian interventions began, but others continued to show high levels of food insecurity. This divergence of evidence raises important questions about how best to utilize such indicators to both allocate and target scarce resources in emergencies and non-emergency contexts alike.

Despite the important implications for advocacy, policy, and programming, very few studies have examined the degree to which common household food security indicators (i.e., measures of the "access" dimension of food insecurity) converge and under what circumstances. While there has been a great deal of work done in the process of developing and validating these indicators to examine how they relate to common comparators like income, expenditures, assets, nutritional status, and caloric adequacy (Maxwell, 1996; Maxwell et al., 1999; Maxwell et al., 2008; Coates et al., 2006; Deitchler et al., 2010; Weismann et al., 2006), there has been less overall attention to the degree to which these metrics, now in widespread use, are associated with one another.

The few studies that have sought to make these kinds of comparisons (Faber, 2009; Gandure et al., 2010; de Cock et al., 2013) have looked at the bivariate relationships between just two or three of these indicators at a time. Such studies have found that correlations tend to be significant, but there is also a great deal of unexplained variance. The existing literature is limited almost entirely to examining the relationships among indicators in their continuous form (i.e., the index/scale score rather than the categorical form used to generate prevalence estimates), and previous studies assessed these relationships at a single, cross-sectional time point.

The objective of this paper is to compare how seven of the most frequently used indicators of food access portray static and dynamic food security among the same sample of rural households. To what extent do they tell the same "story" about household food insecurity and classify households similarly? To our knowledge, no other study has yet performed an extensive comparison of multiple food security indicators from both a static and dynamic perspective. The results of this analysis are expected to shed light on the potential costs of relying on an arbitrary (or in some cases, institutionalized) selection of single indicators to classify the food-insecure.

Background

For many years, age-adjusted per capita caloric intake was considered the "gold standard" metric of access to food at the household level (Hoddinott and Yohannes, 2002; Weismann et al., 2006), and anthropometric measures of nutritional status were the gold standard at the individual level (Young and Jaspars, 1995). But 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), food safety, food preferences, and fluctuations and trends in consumption over time. Additionally, caloric intake can also be very time consuming and expensive to measure in developing country contexts, and so is rarely used outside of research applications. The anthropometric approach is problematic because nutritional status is determined by many factors, food security being only one (Young and Jaspars, 2006, 2009).

Several rapid, cross-contextual indicators of food security have been developed over the past decade or so. These are all relatively quick to use in field contexts; they are "indicators" of food access, not in-depth measures of dietary intake relative to a biological requirement, nor indicators of food availability (such as trade and production). Most of these indicators have been adopted in practice by agencies working in food security programs, but most agencies use only a small subset or a single preferred measure.[1]

These can be categorized as follows:

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. Foods and food groups are often weighted according to nutritional importance. The result is a score that represents the diversity of intake. In some case, diversity is well correlated with calorie consumption as well, although the magnitude of relationship between diversity and caloric intake varies across studies, as discussed below (IFPRI, 2006).

The Food Consumption Score (FCS) is a specific type of dietary diversity index used primarily by the World Food Programme (WFP, 2009). Most studies have confirmed a significant correlation between the FCS and caloric consumption, but to varying degrees across contexts (IFPRI, 2008; Coates et al., 2007; WFP, 2012). Hoddinott and Yohannes (2002) find correlation coefficients ranging from 0.15 to 0.5 between dietary diversity and calorie consumption. Wiesmann et al. (2006) also find significant correlations between FCS and calorie intake. The evidence is more mixed in other studies. One recent field validation test conducted in Latin America found that, although the FCS measure was generally correlated with other indicators, the proposed "universal" thresholds for the FCS were badly misclassifying food insecurity (defined in that study by caloric adequacy). In El Salvador, 0.2% of households were classified as having poor food consumption by FCS thresholds, but 19% of households were shown as having poor caloric consumption (WFP, 2012). A similar field validation in Nepal, Uganda, and Malawi (Mathiassen, 2013) noted the same lack of congruence between FCS and the benchmarks. That study suggested benchmarking the FCS to food basket prices rather than caloric intake (Mathiassen, 2013).

The Household Dietary Diversity Score (HDDS) is similar to the FCS, but with a 24-h recall period and without frequency information or weighted categorical cut-offs. A representative household study in five countries—Bangladesh, Nepal, Pakistan, Tanzania, and Uganda (Tiwari et al., 2014)—found that the FCS and HDDS were highly correlated and significant predictors of child nutritional status across the five countries studied.

FCS is primarily used by the World Food Programme, but – largely due to the influence of the WFP – has been adopted by

[1] The indicators reviewed here are mostly included in the acute food insecurity reference table for household group classification of Integrated Food Security Phase Classification (IPC), a tool that aggregates various kinds of data into a single food insecurity classification "phase" covering various degrees of severity, and used to compare the severity of food insecurity across dissimilar contexts. The IPC is now being used by governments, donors and agencies in 36 chronically risk-prone countries in Asia, Africa and Latin America, and is under consideration for adoption in some 17 more (IPC Partners, 2012). The attempt here is to compare the most commonly used indicators.

many other agencies. The HDDS has been widely promoted by the UN Food and Agriculture Organization and USAID (FANTA, 2006; FAO, 2010). Both are included in the acute food insecurity reference table for household group classification of the IPC.

Consumption behaviors

Some 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 measures the frequency and severity of behaviors in which people engage when they do not have enough food or enough money to buy food. A series of articles noted the correlation of the Coping Strategies Index with both caloric intake and other measures of food access (Maxwell, 1996; Maxwell et al., 1999; Maxwell et al., 2008). Christiaensen and Boisvert (2000) found that the CSI correlated as well with caloric consumption as dietary diversity/food frequency indicators, and was a better predictor of future consumption than either dietary diversity or current caloric intake, indicating that the CSI is a good measure of the element of vulnerability to future shocks.

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 et al., 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 Program/Vulnerability Analysis Mapping unit), FAO/FSNAU (UN Food and Agriculture Organization/Food Security and Nutrition Analysis Unit for Somalia), and the acute food insecurity reference table for household group classification of the IPC.

Experiential measures

Some measures combine items that seek to capture the different behavioral and psychological dimensions of the food insecurity experience along a spectrum of severity. The Latin America and Caribbean Food Security Scale (ELCSA) and the Household Food Insecurity Access Scale (HFIAS) are the best known and most widely used of these measures in international contexts (Coates et al., 2007; Swindale and Bilinsky, 2006). The HFIAS and ELCSA were designed to capture household behaviors signifying insufficient quality, quantity, acceptability, and anxiety over insecure access. The Household Hunger Scale (HHS) was derived from the HFIAS as a culturally invariant subset of questions. It includes three specific questions that tend to represent the most severe manifestations of restricted access to food.

Becquey et al. (2010) compared HFIAS and an individual dietary diversity score (IDDS) with a household "mean adequacy ratio" (MAR) measuring energy intake and a range of micronutrients. They showed that HFIAS and IDDS provided reasonable estimates of mean nutritional adequacy at the population level, but were not sufficient to be used as a means of targeting individual households. Hedlund et al. (2013) noted the use of HHS in monitoring the impact of cash transfers in Somalia during the famine of 2011–2012, and note that under circumstances of severe food deprivation, HHS is well adapted to capture the severity.

USAID, FAO, and others have adopted and promoted the HFIAS and HHS. HHS is included in the acute food insecurity reference table for household group classification of the IPC.

Self-assessment measures

Though highly subjective in nature and perhaps too easy to manipulate in programmatic contexts, a number of self-assessment measures have been introduced in recent years (Headey, 2011,2013). 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. For example, 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). Headey and Ecker (2013) critique self-assessment measures in a recent review of food security metrics.

There is strong evidence that all these measures are capturing something about the multidimensional nature of food security. Studies considering pairs of these indicators have shown strong correlations, but the entire set has not systematically been compared empirically in the same context over the same period of time. In this study we compare CSI, rCSI, FCS, HDDS, HFIAS, HHS, and self-assessed food security (SAFS) in the same sample over four rounds of survey data collection. While data on food expenditure was collected, we decided against inclusion of this variable in the present paper because of concerns about temporal inconsistency; the way this question was asked did not allow us to directly compare it in time to the other indicators.

Methodology

The data for this paper come from the Livelihoods Change Over Time (LCOT) study, a four-round panel survey conducted in two districts of northern Ethiopia between August 2011 and February 2013. The overall objective of the LCOT study was to assess household resilience in the face of an annually recurring shock: the "hunger season." At this time of year, diminishing household food stocks from the previous year's harvest lead to a fall in aggregate local supply, which combines with relatively inelastic demand to inflate grain prices in local markets. The higher prices during the hunger season are coupled with greater incidence of illness, especially malaria and acute respiratory infections, during the months immediately preceding the harvest. Rates of acute undernutrition and other forms of 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 (Devereux et al., 2008). The data in this study thus come from a chronically at-risk context, but not from an acute emergency.

To capture intra-annual as well as inter-annual livelihood dynamics, we collected panel data on a sample of 300 households two times a year for two years: at the height of the hunger season in August 2011 and 2012 and in the middle of the postharvest season in February 2012 and 2013 (between two and three months after harvest, when household access to income and food was expected to be relatively high). Each survey round not only gathered information on the situation prevailing at the time, but also asked retrospective questions about household decisions and experiences over the six-month period prior to the survey (i.e., since the last survey round). It should be noted that the first round of the survey (August 2011) corresponded in time to a larger food security crisis in the region, reaching the level of famine in Somalia. While Tigray was much less severely affected, there were signs of moderate to severe food insecurity stress at the time (FEWS NET, 2011).

The following food security indicators were assessed as part of the household survey: (1) the Coping Strategies Index (CSI), (2) the Reduced Coping Strategies Index (rCSI), (3) the Household Food Insecurity and Access Scale (HFIAS), (4) the Household Hunger Scale (HHS), (5) Food Consumption Score (FCS), (6) the Household Dietary Diversity Scale (HDDS) and (7) a self-assessed measure of food security (SAFS). For each of these indicators, respondents were asked about their experience of food security over the month preceding the survey. The exact wording of the questions, modified in some cases from standard usage to harmonize the time scales covered by each indicator, is shown in Appendix A of the supporting materials.

The 300-household survey sample was stratified by *woreda* (district), *kebele* (sub-district), and *kushet* (village). In each of two *woredas* in the state of Tigray, Tsaeda Amba (Eastern Tigray) and Seharti Samre (Southern Tigray), 150 households were selected, 75 from each of two *kebeles* within each woreda. The *kushet*-level sampling units were obtained by PPS (probability proportional to size) sampling with a random start. Within the village, sampling of households was done by transect walks, with the selection interval determined by the village size in order to allow all households within the village to have an equal probability of being selected.

Several types of analysis were done in this paper. First, the strength of correlations between indicators was assessed using Spearman's *r*, often employed for examining non-parametric bivariate relationships (Section 4.1). A higher score on the CSI, rCSI, HFIAS, HHS, and SAFS is indicative of *greater* food insecurity, whereas a higher HDDS and FCS score indicates greater dietary diversity and food frequency and, thus, *lesser* food insecurity. Thus, inverse correlations among some of these indicators were expected. Second, the scale values of the indicators were converted to categorical values using commonly used cut-offs, and the categorical values of each indicator were compared to each other indicator (Section 4.2). Third, using a binary classification of food secure/food insecure for each indicator, we looked at "dynamic correlation": how different indicators portrayed the movement of households into and out of food insecurity between Rounds 1 and 4 of our survey, over approximately an 18-month period (Section 4.3). We then made some general observations about the trend of (Section 4.4) and variance in (Section 4.5) food security outcomes across all rounds, as suggested by each indicator.

Results

How well did the different measures correlate?

The Spearman's *r* correlations among the seven measures were generally quite strong and in the expected direction, and were all significant at the $p < 0.01$ level, as shown in the table of pooled correlations (all rounds) below (Table 1). The CSI, rCSI and HFIAS were particularly well correlated.

As would be expected, the FCS, in measuring dietary diversity, was less well correlated with either the HFIAS or CSI. The SAFS was moderately well correlated with HFIAS, CSI and rCSI but less so with the other measures (although its association with all variables was statistically significant and the correlation coefficient was never lower than about 0.3, which is above the range reported in other studies reviewed above).

The rCSI and HDDS measures maintained the same strength of correlation with other indicators as their expanded/weighted counterparts, CSI and FCS, but HHS did not maintain the same degree of correlation with other measures as its "parent" indicator, HFIAS. This is expected, as the HHS captures the most extreme consequences of food insecurity, while the HFIAS captures a greater range of the food security spectrum.

Table 1
Spearman's rho correlations between food security measures, all rounds pooled.

	CSI	rCSI	HFIAS	HHS	FCS	HDDS	SAFS
CSI	1						
rCSI	0.95	1					
HFIAS	0.85	0.84	1				
HHS	0.44	0.42	0.48	1			
FCS	−0.51	−0.48	−0.57	−0.34	1		
HDDS	−0.56	−0.53	−0.63	−0.34	0.92	1	
SAFS	0.45	0.46	0.46	0.23	−0.24	−0.29	1

All correlations significant at the $p < 0.01$ level.
This matrix is symmetric.

Interestingly, the strength of correlation between the various food security measures did not change in any consistent pattern when disaggregated by rounds (Appendix B). Thus interrelationships among the measures were stable across times of food scarcity and sufficiency. Overall, all of the measures were significantly correlated, and yet there was enough unexplained difference to suggest that the various metrics might be capturing different aspects of food security (Maxwell et al., 2013).

How were households classified at a single point in time (static classification) by each measure?

Most of these measures can be used in a continuous form (as a scale or index) or as a categorical indicator (for estimating prevalence). Some of these indicators (HFIAS, HHS and FCS) have commonly-used cut-offs to determine categories, although different assumptions and methods underlie each measure's cut-offs. CSI and rCSI do not have universal thresholds for different categories of food insecurity; for these two, thresholds specific to this local context were developed for this study. SAFS is simply analyzed according to the self-assessed categories in the indicator itself. Table 2 below summarizes the categories and qualitative labels for each indicator. (HDDS does not have established categorical cutoffs and is analyzed only as a scale measure; for this reason, it is excluded from Table 2 and from this section's analysis).

Appendix C presents cross-tabulations of the categorical food security measures (Appendix D in the supporting materials provides histograms of every categorical measure, as well as the HDDS continuous measure, by round). Table 3 below summarizes Appendix C by presenting the extent (i.e., the number of categories) to which the *row* food security measure over- or under-calculates food insecurity relative to the *column* measure[2] using data pooled across all rounds. For example, in Table 3 we see that CSI classifies fewer households as food insecure than the 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 is about half a category less food insecure (i.e., more food secure) than the corresponding HFIAS classification. 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 three-quarters 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 emerged from these results:

- First, it is clear that while these indicators are closely correlated in continuous quantitative analysis, the categorical comparisons show a much different picture. We discuss this more below, as it has major implications for policy, practice and measurement.

[2] The procedure of 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 agree on food security status – that is, 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 one and summed); when the over-calculation is two categories, the number of observations are multiplied by two 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).

Table 2
Classification systems of food security measures.

Indicator	Category number	Category description	Range (internal to indicator)
CSI	1	Food secure	0–2
	2	Mildly food insecure	3–12
	3	Moderately food insecure	13–40
	4	Severely food insecure	>40
rCSI	1	Food secure	0–3
	2	Mildly food insecure	4–8
	3	Moderately food insecure	9–18
	4	Severely food insecure	>18
HFIAS	1	Food secure	(N/A; algorithmic classification process)
	2	Mildly food insecure	
	3	Moderately food insecure	
	4	Severely food insecure	
HHS	1	Little to no hunger	0–1
	2	Moderate hunger	2–3
	3	Severe hunger	4–6
FCS	1	Acceptable	0–21
	2	Borderline	21.5–35
	3	Poor	>35
SAFS	1	Food secure	(N/A; categorical measure)
	2	Slightly food insecure	
	3	Moderately food insecure	
	4	Very food insecure	
	5	Extremely food insecure	

Table 3
Degree of difference in classification of food insecurity by 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	–

This matrix is symmetric but the signs change.

- **CSI/rCSI.** As would be expected, CSI and rCSI classified households similarly. The CSI tended to portray a slightly worse food security situation relative to the rCSI (−0.12), as seen in the cell where the CSI row measure and the rCSI column measure intersect (i.e., the second cell in the first row). The CSI and rCSI both tend to depict a worse food security status than the HHS and FCS, but better than the HFIAS, and slightly better than the SAFS measure.
- **HFIAS/HHS.** HFIAS gave the highest prevalence estimates of food insecurity, as might be expected given its inclusion of less severe manifestations of insecurity, including psychological anxiety and food consumption preferences (although this may also be due to the algorithm for classification, a topic we discuss further below). In contrast, the HHS depicted the lowest prevalence of food insecurity. Households were only identified as food insecure if their situation was quite severe, as the measure only counts the most extreme behaviors.
- **FCS.** After HHS, FCS was the most likely to identify a household as food secure.
- **Self-assessment.** SAFS, the food security self-assessment measure, was more likely to classify households as food insecure than all measures but the HFIAS.

Overall, we see that different measures classified food insecure households quite differently. A static classification of food insecurity by the different measures thus implies a hierarchy with respect to the estimation of food insecurity. The categorical indicator related to HFIAS provided the highest estimate of food insecurity prevalence, followed by the categorical SAFS, CSI, rCSI, FCS and HHS indicators (again, HDDS was not evaluated, as it lacks a standard categorical counterpart). 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 vs. relief) of data collection and intervention. At the same time, this presents a puzzle for analysis that relies on more than one indicator: how should the different prevalence estimates be interpreted and acted upon? And alternatively, if only one indicator is used, which one should it be? These choices have very real consequences in situations where the results of assessments using these indicators are used for resource allocation, targeting, and program management.

How are households classified across rounds ("dynamic correlation") by each measure?

In addition to static classification, we can also look at "dynamic correlation": whether different measures portrayed movement of households in and out of food insecurity similarly across the four rounds. We convert the classification systems in Table 2 into binary categories of "food secure" and "food insecure."[3] We do this for simplicity of presentation purposes, but it should also be noted that in operational practice such binary classification is more likely to be used than more nuanced multi-category classification.

Tables 4 and 5 below depict this issue in two different ways. Table 4 shows "net movement into food security", as identified by each measure. To calculate this, we first determine the number of households who moved from food insecurity between Round 1 and Round 4 (Column A, expressed as a percentage of the original number of food insecure households). We then determine the number who were originally food secure but fell into food insecurity (Column B, as a percentage of the number who were originally food secure). We then subtract the second figure from the first, and again present the result as a fraction of the number of households who were originally food insecure (Column C).

For example, in Table 4 below, CSI identifies 247 households as food insecure in Round 1; of these, 117, or 47.4% (Column A), were food secure in Round 4: they thus moved into food security.[4] Conversely, 14 of the 53 households who were originally food secure, or 26.4% (Column B), fell into food insecurity by Round 4. Thus the net number of households who moved into food security is 117−14 = 103; that number represents 41.7% of the original 247 food insecure households (Column C).

Table 4 above shows that there are differences in how the measures estimate food security dynamics. We see that with respect to Column A's gross movement into food security, SAFS has the lowest rate (i.e., it is the indicator most likely to depict persistent food insecurity), followed by HFIAS, CSI, rCSI, FCS and finally HHS. The converse – the measurement of movement into food insecurity (Column B) – is somewhat different, with HHS showing the least movement, followed by FCS, HFIAS, rCSI, CSI and SAFS. The size of the net differences roughly resemble the earlier static analysis,

[3] With reference to Table 2: for CSI, rCSI and HFIAS, category 1 is considered "food secure" and categories 2–4 "food insecure". For HHS, category 1 is considered "food secure" and categories 2 and 3 "food insecure". For FCS, category 1 is considered "food secure" and categories 2 and 3 "food insecure". For SAFS, categories 1 and 2 are considered "food secure" and 3–5 "food insecure".

[4] Note that Round 1 was that a hunger season and Round 4 a harvest season, and so some of the movement into food security is likely to be temporary. However, the Round 3 (hunger season) food security prevalence estimates – much lower than those of the Round 1 hunger season – suggest that most of the movement is *not* due to hunger/harvest season dynamics.

Table 4
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 movement into food security as % originally food insecure
CSI	47.4% (117/247)	26.4% (14/53)	41.7% (103/247)
rCSI	58.2% (128/220)	22.2% (18/81)	50.0% (110/220)
HFIAS	36.0% (95/264)	13.2% (5/38)	34.1% (90/264)
HHS	100.0% (18/18)	2.8% (8/284)	55.6% (10/18)
FCS	74.3% (52/70)	9.5% (20/211)	45.7% (32/70)
SAFS	27.3% (57/209)	52.7% (49/93)	3.8% (8/209)

with some exceptions; for instance, FCS is closer to the coping strategies indices when measuring food security dynamics rather than statics.

Table 5 below presents a more detailed comparative picture of the 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 of 34.2% of them (second to last column). Similarly, of all the households that HFIAS reports as having moved into food security, FCS corroborates just 16.8% of these. Thus the matrix provides a rough "dynamic correlation matrix" between Rounds 1 and 4 of the binary outcome of food security vs. insecurity.

Table 5 also provides convenient summary measures in the final column and final row. The totals in the final column could be seen as "movement into food security" implied by a row measure: *the extent to which all other measures corroborate the claim of movement into food security by a given row measure*. For example, when HFIAS claims that a household moved into food security, in a total of 30.4% of possible instances the other indicators corroborate this claim, 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 13.9% of the time. Note that the HHS and FCS measures are again at the other end of the range of estimates of food insecurity prevalence. Relative to the other indicators, movement into food security by HHS and FCS is less frequently corroborated.

The totals in the final row of Table 5, meanwhile, capture the "total dynamic correlation" of a given column measure to the full set of other 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 other measures combined. Thus we see that CSI, rCSI and HFIAS tend to corroborate the full set of measures more than the others – corroborating 44.3%, 42.6% and 37.0%, respectively, of the instances of movement into food security identified by the other measures. HHS is at the other extreme, corroborating just 3.2% of the claims of movement into food security by other measures.

The import of corroboration is this: underlying each indicator is a particular concept of food security, embedded in the choice of its questions. The movement of food insecurity into food security is a particularly important aspect of that concept, especially for operational planning: when do indicators claim that previously insecure households are now in an acceptable state and no longer require assistance? Indicators that are strongly corroborated by others can be assumed to have similar underlying notions of food security and dynamics of movement into food security. Planners may thus choose to consider a combination of "least-like" indicators as a part of an overall dashboard of measures to ascertain whether food security has been achieved in all of its aspects. Conversely, if funding or time is limited and multiple types of indicators cannot be used, an indicator that is known to be well corroborated by others can be assumed to best capture the majority of those indicators' conceptions of food security.

The cut-offs traditionally used to construct the categorical indicator for each of these measures were derived through a combination of "expert opinion" and judgment; all categorization schemes currently used for such indicators are ultimately a matter of judgment. We thus performed sensitivity analyses of the cross-corroboration matrices by altering the binary cut-points. The results across 16 different combinations of cut-points were substantively similar, suggesting that the conclusions above are not overly sensitive to how categories are constructed. However, it should be noted such insensitivity to cut-points is likely to be a particular feature of binary classification, and less likely to hold if, as in the original form of all the indicators, more than two categories are considered. Further research should consider the sensitivity of food insecurity classification across indicators when the original categories are considered, although this will be complicated by how the categories' qualitative labels are interpreted and compared across indicators.

What do the measures tell us about comparative changes in food insecurity across all four rounds?

We now compare trends in food security across all 4 rounds, as shown by the different measures. Appendix E in the supporting materials presents more detailed results. Fig. 1 below suggests that, as measured by any of the indicators, food security improved across the timeframe of the survey. The differences in prevalence of food insecurity between rounds are statistically significant across all 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 of the same type, 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 was damping the potentially negative within-year seasonal effects. Note again that the prevalence figures below were calculated using our binary system of categorization.

Table 5
Dynamic correlation—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	–	87.2	69.2	2.6	17.1	34.2	29.3
RCSI	79.7	–	58.6	3.9	13.3	32.0	25.9
HFIAS	85.3	78.9	–	1.1	16.8	25.3	30.4
HHS	16.7	27.8	5.6	–	33.3	5.6	13.9
FCS	40.4	32.7	30.8	11.5	–	23.1	19.2
SAFS	70.2	71.9	42.1	1.8	21.1	–	34.5
Corroboration of all other measures by column measure	44.3	42.6	37.0	3.2	12.6	25.3	

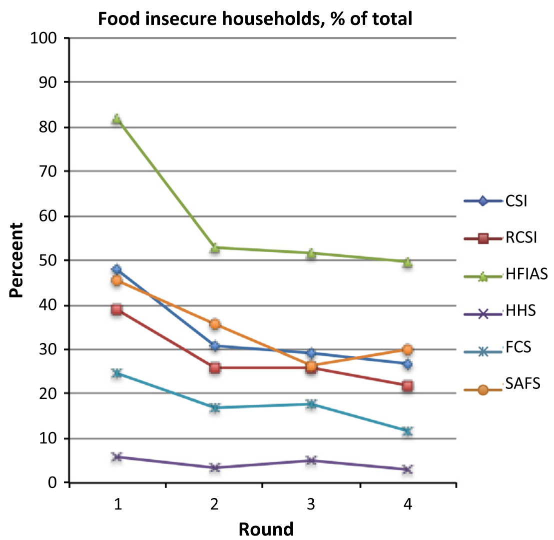

Fig. 1. Food insecure households as % of total, by round and indicator.

The HFIAS measure shows the highest prevalence of food insecurity in the survey areas across all 4 rounds. The CSI and rCSI showed similar dynamics to each other, suggesting that the questions in the rCSI are adequately picking up the same kind of change in insecurity detected by the CSI, in this context. 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.

Overall, the trends are very similar, but the prevalence of food insecurity is quite different across indicators. This can be explained in three (non-exclusive) ways. First, it is possible that the choice of cut-offs within the indicators is not harmonized; that is, one indicator's cut-off for food insecurity reflects a much different situation than another's. However, in the absence of a gold standard linking cut-offs to all of the aspects of food security – quantity, quality, acceptability, safety, and stability – such harmonization cannot be done. Second, it is possible that food insecurity cut-offs across indicators *do* reflect the same food security situation (again, in relation to some theoretical, unobserved gold standard), but different indicators are measuring different aspects of food security. Finally, it is likely that these indicators are sensitive only at certain "severities" of food insecurity; HHS, for example, is almost certainly only applicable in extreme situations. Outside of a given indicator's range of highest sensitivity, the variation in score is likely to be dampened because the indicator has hit one extreme (e.g., maximally food insecure) or the other (e.g., maximally food secure) of its diagnostic capability.

What do the measures tell us about the variance in food security outcomes across indicators?

The different measures imply different distributions of food security scores among households in a single survey round; Appendix F in the supporting materials shows the distribution of food insecurity as indicated by each index/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/scale scores or categorical measures (Makdissi and Yazbeck, 2012). More traditional measures of inequality such as the Gini and Thiel indices cannot be calculated, particularly when individual household scores cannot be represented as a proportion of the sample, as with the food security indicators we use in this study. However, the coefficient of variation (CV) – the standard deviation normalized by dividing by the mean – for each measure by round can give a general comparative sense of variance across indicators (Cowell, 2009). Table 6 below presents the CV of each measure by round.

In the absence of a variance "target", the numbers themselves cannot suggest absolutely "high" or "low" variances, but we can look at some general comparisons. The table suggests that the greatest variance exists with the HHS measure. The CSI, rCSI and to a lesser extent HFIAS also display high variance in this dataset, although much less than HHS. The dietary diversity (FCS and HDDS) and self-assessment (SAFS) measures show much less variance, implying that how food security is measured matters greatly for evaluating dispersion of food security outcomes in a population. In this case, variance in food security outcomes is lesser 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.

This suggests that knowing the distributions of food security outcomes when using these latter indicators is important; means alone may obscure the presence of a high food insecure portion of the population.

With respect to our specific dataset, the distributions in Appendix F of the supporting materials, summarized by the coefficients of variation in Table 6, might be explained by the following factors:

- **CSI/rCSI.** The relatively high variances of the CSI and rCSI are largely due to the high number of households reporting zero values (meaning no coping behaviors reported), especially in the latter three rounds. Table 7 below 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 et al., 2008). 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.
- **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. 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 discussed further below.

Table 6
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 7
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

- **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 studied.
- **FCS/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 1–3 are the only measures that pass a formal Jarque–Bera test for normality (i.e., fail to reject the null hypothesis of normality at $p < 0.1$). The shape of the distribution largely explains these measures' low variance relative to the others.
- **SAFS.** The relatively low relative standard deviation for the self-assessment measure is largely an artifact of the few discrete values available as responses (1 through 5).

It is also worth noting that variance in the sample increases as food security outcomes improve across rounds: dispersion of values within the population appears to worsen when times are better for the population as a whole, and vice-versa. Thus a food secure situation at the population level may not reflect the fact that many households remain food insecure. This is especially a concern with respect to safety net programs targeted at the district or higher levels. Food security outcomes measured through district mean values should not be used as justification for termination of programs without knowing underlying distributions of outcomes.

Discussion

Given the strong correlations among the measures as well as the depiction of generally similar food security trends over time, and given that most of these measures have been tested elsewhere for content validity (Coates et al., 2006; Deitchler et al., 2010; Maxwell et al., 2008), it is fair to say they are all measuring elements of food security. They do also have significant differences, however, reflected both in the correlation coefficients and in the distributions.

The indicators reviewed are strongly correlated, and in the expected directions in all cases. Perhaps predictably, CSI, rCSI and HFIAS tended to cluster together in terms of correlation, as did HDDS and FCS. HHS and SAFS generally had lower correlation coefficients with the other indicators, but are still robustly correlated to the others.

These results, however, pertain to the indicators in their continuous quantitative form. In practice, all of these indicators—with the exception of HDDS, which does not have specified cut-offs for category construction—are more frequently used as categorical measures and probably most frequently simply as binary outcome variables: "food secure" and "food insecure" (especially in an operational context—assessment, targeting, monitoring and evaluation, etc.). When viewed in this categorical manner, some clear differences between the indicators emerge.

Statically, food insecurity is categorized quite differently by the different measures (Table 3). The HFIAS, even though highly correlated with the CSI and rCSI, gives a much higher prevalence estimate than the others. HHS, even though derived from HFIAS, provides the lowest prevalence estimate.

Dynamically, the same story is reproduced. HFIAS and SAFS produce the highest estimates of food insecurity across all four rounds; but all the indicators demonstrate a similar pattern overall, with Round 1 results showing considerably higher levels of food insecurity, the situation improving rapidly by Round 2, and only a slight worsening in the hunger season of Round 3. HHS and FCS produce the lowest estimates of prevalence, but depict almost exactly the same trend. Correlation between the different measures did not change when disaggregated by rounds—the inter-relationships among the measures were stable across times of food scarcity and sufficiency, but there was enough difference to suggest that different measures were capturing different aspects of food security. This will be explored further in a separate forthcoming analysis.[5]

Some of the differences in prevalence estimates can be explained fairly easily: HHS focused only on the most extreme experiences included in the HFIAS, and thus tends to produce very low estimates because, in the surveyed areas during these seasons, such experiences were rare. HFIAS on the other hand (along with CSI and rCSI) picks up a wider range of experiences or behaviors. The difference between HFIAS and the CSI/rCSI is likely explained by the way in which the categories are constructed. The CSI typically has *not* had thresholds or cut-off points included in the methodology because it is intended to be context specific. In this study, cut-off points specific to the Ethiopian context were devised for comparative purposes, but these are not intended to be universal. The HFIAS, on the other hand, relies on an algorithmic approach to constructing the categories of food insecurity. According to the algorithm selected through an "expert consensus" exercise, even a single occurrence of certain experiences can put a household into a more severely food insecure category. This is likely part of the reason for the higher prevalence estimates.

FCS produces fairly low estimates of food insecurity in our survey context, even when considering the lower two categories (termed "poor" and "borderline") both as food insecure, as shown in Table 2 (typically, FCS is most commonly used with three outcome categories, with some ambiguity whether the middle "borderline" measure should be considered food insecure). But even our preferred approach to a binary FCS, which should raise estimates of food insecurity, yields a relatively low prevalence of food insecurity. This may in part be explained by choice of the global (universal) cut-off points suggested in the guidance documentation for FCS, and may suggest the need for more context-specific cut-off points (as indeed some of WFP's internal documentation has also argued; see WFP, 2012).

The self-assessment measure deserves some additional consideration. After the HFIAS, it provided the highest prevalence estimates of food insecurity, but for the most part showed the same pattern over time as the other indicators. Self-assessment measures are sometimes dismissed as "subjective" and for being too easy for respondents to "game"—particularly if respondents perceive them to be tied to some form of assistance or entitlement

[5] See Maxwell et al. (2013) for some preliminary conclusions.

(though we took great pains to ensure respondents that the study was for research purposes only, not a needs assessment, and that answers were not tied to the provision of assistance). The SAFS has not been previously validated, and one might expect an indicator constructed from a single question to be less reliable than an indicator scaled from a range of questions. For these reasons, the SAFS should probably not be used as a stand-alone measure, but it is a useful method of crosschecking, and provides some hints about self-perception that other measures do not. For instance, while all of the other measures show improvement in food security status from Round 3 (hunger season) to Round 4 (post-harvest), the SAFS actually depicts a deteriorating status. There are a couple of possible (and interconnected) explanations for this. First, improvements in food security as defined by these indicators could be less quickly apparent to households. Second, and more likely, there are aspects to households' own notions of food security that are not adequately captured by these indicators, perhaps most importantly psychological states of uncertainty and anxiety around medium-term livelihood planning. In any case, had the collection of data been tied to real-time analysis (as it might in an operational situation), deteriorating or stagnating food security trends could have been identified very quickly with a self-assessment measure, and follow-up surveys could be proposed.

Conclusion

"Getting measurement right" is critical to improve humanitarian and development response to food insecurity. Gaps in measurement can result in serious exclusion error – and the attendant loss of lives, damage to health and livelihoods, and waste of scarce assistance resources. Using a panel dataset from northern Ethiopia, we have shown in this paper that the most commonly used indicators of food insecurity overlap to an extent, but that the correlation is far from perfect, and is sometimes quite weak, especially when the indicators are examined in their categorical form as a prevalence estimate.

We suggest three major reasons for this lack of fit. First, the various indicators reflect different dimensions of the complex notion that "food security" represents. In practice, however, 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 if indicators are capturing different elements of food security, relying on only one measure of food security in analysis and program design runs the risk of serious misclassification by capturing some, but not all, of the dimensions of food insecurity inherent in the definition (Coates, 2013). 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 constructed from their scores also give very different estimates of prevalence due to the different methods used for classification.

The general consensus in the literature is that there is will never be a single measure that adequately and accurately captures the complexity of food security. The best way forward is to identify and rely on a range or "suite" of indicators that can capture the different elements of food security (Cafiero, 2012; FAO/WFP/IFAD, 2013; Coates, 2013), and offer a sufficiently wide range of alternatives that will fit different data/measurement needs. The next challenge is to unpack exactly what these aspects are in a conceptual sense and to understand the overlap and independence of the seven indicators used in this study in measuring these aspects of food security.

The second reason for weaker than expected correlation is that the indicators are applicable only at certain levels of food security severity. Outside of this "high sensitivity" range, variation in scores will fall, dampening correlation with indicators that still show variation at the given range. Operationally, indicators are much in demand for the measurement of program and policy impact, and to determine which households or groups are so badly off that they require some form of direct intervention (be it in the form of social protection programs or humanitarian response). Thus the operational application of many of these measures may be towards the moderate to severe end of the spectrum. This can be problematic. For example, 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). All this implies the need to understand which indicators are best suited for which applications, in addition to using more than one. It also suggests that different indicators may be useful in the same locations but at different times in relation to a shock of acute crisis.

Finally, the prevalence estimates that each indicator provides are a function not only of the objective accuracy of the indicators themselves – and in fact, there is no gold standard of food security measurement by which "truthfulness" could be judged – but also 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 a matter of analytical judgment and can often be controversial, 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, 2012; 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).[6] 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 has not adequately addressed the conundrum of accurately categorizing continuous quantitative data.

Overall, in the absence of a gold standard or some kind of "clinical assessment" against which to measure the individual indicators, we can simply note that all these indicators perform as expected, but it is difficult to say which performs "best" in terms of correctly and reliably identifying food insecure households. Operationally, this raises several questions: (a) which indicator to use, (b) how to interpret results where more than one is used and (c) whether the use of multiple indicators in a single application of context can reduce the likelihood of misclassification of food insecurity status. For example, if policymakers wish to assure that interventions target the broadest set of food security needs possible, analysts should use multiple indicators and in any given context select either the one that shows the highest prevalence or employ some type of additive process by which a household is considered food insecure if *any* measure identifies it as being so. Alternatively – and in the more realistic scenario where resources are limited – claims of food insecurity could be qualified by noting what aspect of food security is likely being measured by the indicator in use, and interventions targeted towards households that are food insecure in the dimension deemed of greatest priority. As the results of this paper demonstrated, choosing any given

[6] 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.

indicator to the exclusion of others may miss a large fraction of the population that is food insecure in some important sense, though not in the sense captured by that indicator. The question of which dimensions of food insecurity are being captured by which indicators will be addressed more directly in a second forthcoming paper, to be published using these same study data.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.foodpol.2014.04.003.

References

Barrett, Christopher, 2010. Measuring food insecurity. Science 327, 825–828.

Becquey, Elodie, Marin-Prevel, Yves, Traissac, Pierre, Dembèlè, Bernard, Bambara, Alain, Delpeuch, Francis, 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. J. Nutr. 140, 2233–2240.

Cafiero, Carlo, 2012. Advances in Hunger Measurement. In: Presentation to the International Scientific Symposium on Food and Nutrition Security Information. FAO, Rome.

Christiaensen, Luc, Boisvert, Richard, 2000. On Measuring Household Food Vulnerability: Case Evidence from Northern Mali. In: Working Paper. Department of Agricultural, Resource, and Managerial Economics. Cornell University, Ithaca, New York.

Coates, Jennifer, 2013. Build it back better: deconstructing food security for improved measurement and action. Global Food Security 2 (2013), 188–194.

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

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

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

de Cock, N., D'Haese, M., Vink, N., van Rooyen, C.J., Staelens, L., Schönfeldt, H.C., D'Haese, L., 2013. Food security in rural areas of Limpopo province, South Africa. Food Security J. 5, 269–282.

de Haen, Hartwig, Klasen, Stefan, Qaim, Matin, 2011. What do we really know? metrics for food insecurity and undernutrition. Food Policy 36 (6), 760–769.

Deitchler, Megan, Ballard, Terri, 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.

Devereux, Stephen, Vaitla, Bapu, Hauenstein-Swan, Samuel, 2008. Seasons of Hunger. Pluto Press, London.

Faber, Mieke, 2009. Dietary diversity in relation to other household food security indicators. Int. J. Food Saf., Nutr. Public Health 2 (1), 1–15.

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

Famine Early Warning System Network (FEWS NET), 2011. East Africa Seasonal Monitor (various months). FEWS NET, Washington DC.

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

Food and Agriculture Organization (FAO), World Food Program (WFP), and International Fund for Agricultural Development (IFAD). 2013. The State of Food Insecurity in the World 2013: The Multiple Dimensions of Food Security. Rome: FAO/WFP/IFAD.

Food and Agriculture Organiztion, 1996. The Rome Declaration on World Food Security. The World Food Summit. FAO, Rome.

Gandure, Sithabiso, Drimie, Scott, Faber, Mieke, 2010. Food security indicators after humanitarian interventions including food aid in Zimbabwe. Food Nutr. Bull. 31, 513–523.

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

Headey, Derek. 2013. The impact of the global food crisis on self-assessed food security. World Bank Economic Review (Advance Access published 01.07.2013).

Headey, Derek, Ecker, Olivier, 2013. Rethinking the measurement of food security: from first principles to best practice. Food Security 5, 327–343.

Hedlund, Kerren, Majid, Nisar, Maxwell, Daniel, Nicholson, Nigel. 2013. Final Evaluation of the Unconditional Cash and Voucher Response to the 2011–12 Crisis in Southern and Central Somalia: UNICEF.

Hoddinott, John, Yohannes, Yisehac, 2002. Dietary Diversity As a Food Security Indicator. IFPRI, Washington, DC.

IPC Partners. 2012. Integrated Food Security Phase Classification. Technical Manual. Version 2. IPC Technical Team, Rome.

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

IFPRI. 2008. Validation of food frequency and dietary diversity as proxy indicators for household food security. Report submitted to WFP. IFPRI, Washington DC.

Makdissi, Paul, Yazbeck, Myra. 2012. Measuring socioeconomic health inequalities in presence of multiple categorical information. In: Department of Economics Working Paper #1208, University of Ottawa.

Mathiassen, Astrid, 2013. Validation of the Food Consumption Score Based on LSMS Data. WFP, Rome.

Maxwell, Daniel, 1996. Measuring food insecurity: the frequency and severity of coping strategies. Food Policy 21, 291–303.

Maxwell, Daniel, Ahiadeke, Clement, Levin, Carol, Armar-Klemesu, Margaret, Zakaria, Sawadatu, Lamptey, Grace Mary, 1999. Alternative food-security indicators: revisiting the frequency and severity of 'coping strategies'. Food Policy 24, 411–429.

Maxwell, Daniel, Caldwell, Richard, 2008. The Coping Strategies Index: Field Methods Manual. Tufts University, Medford, USA and CARE, Atlanta, USA.

Maxwell, Daniel, Caldwell, Richard, Langworthy, Mark, 2008. Measuring food insecurity: can an indicator based on localized coping behaviors be used to compare across contexts? Food Policy 33, 533–540.

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

Swindale, Anne, Bilinsky, Paula, 2006. Development of a universally applicable household food insecurity measurement tool: process, current status, and outstanding issues. J. Nutr. 136, 1449S–1452S.

Tiwari, Sailesh, Skoufias, Emmanuel, Sherpa, Maya. 2014. Shorter, cheaper, quicker, better: linking measures of household food security to nutritional outcomes in Bangladesh, Nepal, Pakistan, Uganda, and Tanzania, World Bank Policy Research Working Papers 6584.

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

World Food Programme, 2009. Comprehensive Food Security and Vulnerability Analysis Guidelines. WFP, Rome.

World Food Programme, 2012. Validation Study of the WFP's Food Consumption Indicator in the Central American Context, with A Focus on Intra-Household Sharing of Food. WFP, Rome.

Young, Helen, Jaspars, Susanne, 1995. Nutritional assessments, food security and famine. Disasters 19, 26–36.

Young, Helen, Jaspars, Susanne. 2006. The meaning and measurement of acute malnutrition in emergencies: a primer for decision-makers. HPN Network Paper 56.

Young, Helen and Susanne Jaspars. 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.