Measurement 116 (2018) 146–152

Contents lists available at ScienceDirect

Measurement

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

Food security measurement in a global context: The food insecurity experience scale

Carlo Cafiero*, Sara Viviani, Mark Nord

Food and Agriculture Organization of the United Nations, Statistics Division (ESS), Viale delle Terme di Caracalla, 4, 00153 Rome, Italy

ARTICLE INFO

Keywords:
Food security measurement
Rasch model
Scale equating
Sustainable development goals

ABSTRACT

The ability of households and individuals to access food (one of the key aspects of 'food security') is an important welfare dimension that poses important challenges for objective measurement. This paper describes the Rasch model-based procedures developed to define the eight-item Food Insecurity Experience Scale (FIES) as a contribution towards the establishment of an indicator for global monitoring of food insecurity. Experiential food insecurity survey data, collected by FAO from nationally representative samples of the adult population, once every year in 2014, 2015 and 2016 from 153 countries or territories, are used to develop methods to estimate cross-country comparable prevalence rates of moderate and severe food insecurity. A Rasch model-based scale was estimated separately for each country and data were assessed for consistency with model assumptions. To ensure cross-country comparability, a procedure based on the median normalized severities of each of the eight FIES items was used to define a global reference scale, against which measures obtained in each country can be separately calibrated. Calibration is obtained by equating the mean and standard deviation of the severity parameters of the items that appear to be common between the national and the reference scale, and thus used as anchoring points for the metric. Data showed sufficient consistency with the Rasch model assumptions to produce reliable measures of the prevalence of food insecurity in each country. Calibration was possible using 4 or more items as anchoring points in 151 of 153 (98.7%) of the cases, and 6 or more items in the vast majority of them (121 cases). Concurrent validation of the estimates of prevalence of food insecurity at national level was obtained by comparing the FIES-based indicator with other established indicators of social (under) development. National prevalence rates of moderate-or-severe food insecurity obtained by FAO correlate well with the prevalence of undernourishment and with several widely used indicators of national income, health, and well-being. The proposed calibration method can be applied to other existing experience-based food security scales that use similar items, thus affording the possibility to use data collected with those instruments to produce internationally comparable measures of the prevalence of food insecurity. Pending broader adoption of the FIES or compatible experience-based food security scales worldwide, countries could choose to use the 2014–2016 results obtained using the data collected by FAO as the baseline to monitor progress towards Target 2.1 of the recently established 2030 Agenda for Sustainable Development.

1. Introduction

Food security is said to exist when all people, at all times, have physical, social and economic access to sufficient safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life [1]. Although food security is inherently multi-dimensional, one critical dimension is continued access to adequate food. The United Nations Food and Agriculture Organization (FAO) has undertaken a project called Voices of the Hungry (VoH) to develop and support a survey-based experiential measure of access to food, called the Food Insecurity Experience Scale (FIES). The approach to measuring households' ability to access food is similar to that of other experience-based food security scales such as the US Household Food Security Survey Module (HFSSM), the *Escala Brasileira de Insegurança Alimentar* (EBIA), the *Escala Latinoamericana y Caribena de Seguridad Alimentaria* (ELCSA), the *Escala Mexicana de Seguridad Alimentaria* (EMSA) and the Household Food Insecurity Access Scale (HFIAS) used in the United States, Brazil, Canada, Mexico and several other countries to monitor food security in the context of large national programs [2]. The innovation brought about by the VoH project is the possibility to calibrate the measures with the FIES or with any of these other scales and the thresholds used for classification, against a standard reference scale,

* Corresponding author.
E-mail address: carlo.cafiero@fao.org (C. Cafiero).

https://doi.org/10.1016/j.measurement.2017.10.065
Received 17 April 2017; Received in revised form 26 October 2017; Accepted 29 October 2017
Available online 06 November 2017

Table 1
English version of the food insecurity experience scale.

N.	Short reference	Question wording
1	WORRIED	During the last 12 MONTHS, was there a time when you were worried you would not have enough food to eat because of a lack of money or other resources?
2	HEALTHY	Still thinking about the last 12 MONTHS, was there a time when you were unable to eat healthy and nutritious food because of a lack of money or other resources?
3	FEWFOODS	Was there a time when you ate only a few kinds of foods because of a lack of money or other resources?
4	SKIPPED	Was there a time when you had to skip a meal because there was not enough money or other resources to get food?
5	ATELESS	Still thinking about the last 12 MONTHS, was there a time when you ate less than you thought you should because of a lack of money or other resources?
6	RANOUT	Was there a time when your household ran out of food because of a lack of money or other resources?
7	HUNGRY	Was there a time when you were hungry but did not eat because there was not enough money or other resources for food?
8	WHOLEDAY	During the last 12 MONTHS, was there a time when you went without eating for a whole day because of a lack of money or other resources?

thus ensuring proper comparability of the estimated prevalence rates and the possibility to compute consistent estimates at regional and global level, an essential feature for an indicator to be used in the context of global monitoring frameworks. Following a very broad consultation with many stakeholders, the FIES was chosen as the basis to compile indicator 2.1.2, one of the two indicators included in the global SDG indicator framework put forth by the Interagency and Expert Group on SDG indicators (IAEG-SDG) of the United Nations Statistical Commission to monitor Target 2.1 of the recently adopted 2030 Agenda for Sustainable Development [3].

The FIES measures the severity of food insecurity modelled as a latent trait, broadly conceptualized as the condition of not being able to freely access the food one needs to conduct a healthy, active and dignified life. The measure is based on conditions and behaviors reported by responding to an 8-item questionnaire, the Food Insecurity Experience Scale Survey Module (FIES-SM; see Table 1), resulting from the inability to access food due to lack of money or other resources. These conditions have been selected, among the many possible ones that could be meant to be a direct consequence of the latent condition, as those holding the greater promise to be empirically valid in many different contexts.

The dichotomous ("yes"/"no") responses to the FIES-SM questions, provide information sufficient to construct a one-dimensional measure, using the Rasch model. Based on the measured severity of food insecurity, each respondent in a representative sample is assigned a probability of being beyond a specified threshold of severity to compile an estimate of the prevalence rate of food insecurity in the reference population. Thresholds used for classification and, thus, prevalence rates of food insecurity, are made comparable across countries by calibrating the measures obtained from estimating the Rasch model parameter separately on each dataset, against a common, global reference scale.

The next sections describe the data used, the statistical modeling and the procedures developed to form the global reference scale and to calibrate the measures, and address validation of the food insecurity prevalence rates estimated in 153 countries for 2014–16.

2. Data

In proposing the FIES as the basis to compile an SDG indicator, FAO expects that national prevalence rates of food insecurity for monitoring progress toward SDG Target 2.1 will eventually be based on data from national surveys conducted by national statistical agencies in each country, in accordance with the principles that govern the definition of the global SDG indicator framework by the UN Statistical Commission. To develop methods for making prevalence rates across countries comparable, however, it was necessary to process data obtained from as large a set of different countries as possible while controlling for the survey vehicle used. To that aim, in 2013 FAO contracted with the Gallup Organization as a data collection service provider. The 8-question FIES-SM was added as a client module to the Gallup World Poll (GWP) and data were collected in 153 countries, areas, and territories in 2014, 2015 and 2016. The GWP is a worldwide survey conducted annually since 2006 in about 150 countries interviewing nationally representative samples of the adult population (aged 15 and older) in each country. It covers a range of topics including family economics, employment, human development, and well-being [4].

A dataset of about 1000 records for each country (3000 for India and 5000 for China) was provided by Gallup to FAO containing the responses given to the 8 FIES items by the people reached by the GWP. Most of the countries/territories were covered in each of the three years 2014, 2015 and 2016, with a few exceptions. Responses are coded as 1 for "yes", 0 for "no" and N/A otherwise. Percentages of N/A are very low. Post-stratification weights were provided to project sample estimates to the reference population of individuals aged 15 years or more.

3. Statistical modeling of FIES data to produce estimates of the prevalence of food insecurity at comparable levels of severity

The statistical model used for FIES data assessment and scale construction is the single-parameter logistic measurement model, commonly known as the Rasch Model [5]. The Rasch model assumes that the position of a respondent and that of the items can be located on the same one-dimensional scale and postulates that the log-odds of respondent r saying "yes" to item i is a linear function of the difference between the severity of the food insecurity condition experienced by r and the severity of item i. By coding $x_{r,i}$ (the answer given by respondent r to item i) as 1 for "yes" and 0 for "no", we have:

$$p \equiv \text{Prob}(x_{r,i} = 1) = \frac{e^{(a_r - b_i)}}{1 + e^{(a_r - b_i)}} \Leftrightarrow \ln\left(\frac{p}{1-p}\right) = a_r - b_i$$

where a_r is the position of the respondent and b_i that of the item on the same scale. In applying to food insecurity, we define the scale as a scale of "severity", that is, the degree of negative impact on the household or individual welfare of the inability to freely access the food one needs.

While more flexible IRT parameterizations, such as the 2 or 3-parameter logistic models used in the field of educational testing might increase the fit of the data to the model, in choosing the Rasch model as the basis for a globally valid food insecurity measurement scale we have been guided by the desirable invariance property of Rasch measures [5,6]. To illustrate the concept of invariant measurement, it may be useful to refer to the relationship between the location of the respondent and the probability of reporting an item, described by the so-called item characteristic curve (ICC), which plots the probability of affirmation against respondent severity (Fig. 1). The severity of an item is thus implicitly defined as the severity experienced by respondents for which the probability to affirm or deny the item is the same (i.e., 0.5). Items whose ICC are located on the left along the scale of severity are thus less severe than items whose ICC are on the right. One important assumptions of the Rasch model is that all items discriminate equally—that is, that the slopes of the ICCs are equal for all items for any given level of probability, as it is the case for items labeled A and B in

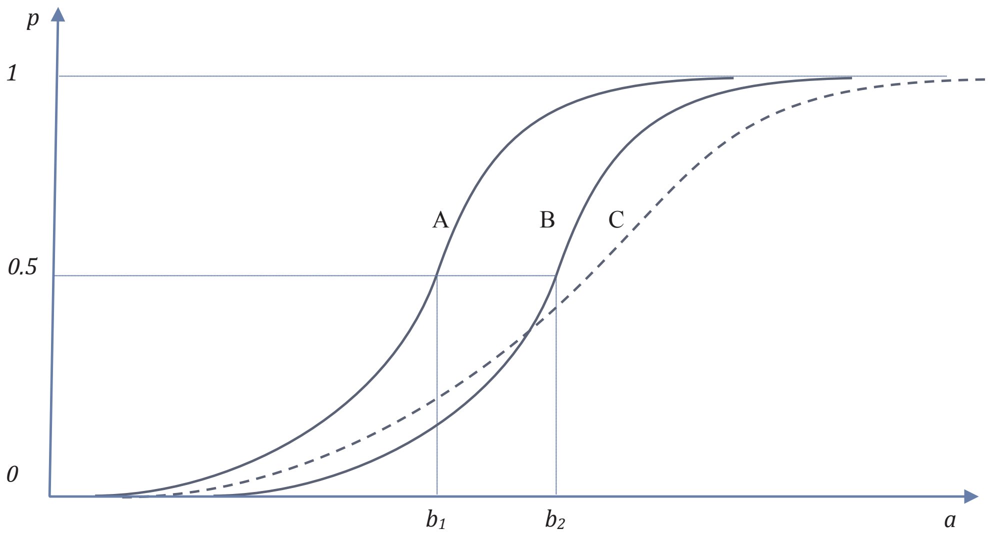

Fig. 1. Item characteristics curves.

Fig. 1. This implies that, independently from the level of severity of the respondent's condition, the probability to affirm less severe items is always higher than that of affirming more severe ones. Allowing for items with different discriminating power (such as item C in Fig. 1) would result in the possibility that the implicit order of severity of the items might differ, depending on the severity of the respondent, thus violating the "invariance" property.

Testing adherence of the data to the Rasch model assumptions is of essence, as only measures obtained with data that do not reject the Rasch model's assumption can be considered invariant with respect to the specific sample of respondents used to estimate the parameters of the model, an important feature of a globally valid measurement tool.

Maximum likelihood methods are used to estimate the severity of each item and the severity experienced by each respondent [7]. Model estimation also includes calculation of item and respondent fit statistics, conditional correlations across items, and measurement reliability of the scale. Item severity parameters, item-fit statistics, respondent-fit statistics, and inter-item conditional correlations are all based on conditional maximum likelihood (CML) estimation procedures. A customized R-package has been created that implements Rasch model estimation by also allowing for complex survey design and that produces a number of additional statistics useful to analyze FIES data [8]. The package produces estimates of item severity parameters by maximizing the likelihood function conditional on the raw score, and using, in the estimating process, only cases with non-extreme response patterns (that is, with raw score between 1 and 7). Respondent severity parameters are then estimated by maximizing the likelihood function given item parameters. Finally, the extent of uncertainty (standard error of estimation) around the respondents' severity parameters is measured as the square root of the inverse of the slope of the test characteristic curve at the point corresponding to the raw score parameter.

In assessing the extent to which data is consistent with the Rasch model assumptions, various statistics are used. The assumption of equal discrimination of the eight items is assessed primarily by the item infit statistic, an information-weighted chi square-type statistic that compares observed with expected misfit of each item. The expected value of all infits is 1.0 and values between 0.7 and 1.3 are considered to be reasonably consistent with the assumption of equal discrimination [6]. To verify the conditional independence of the responses to the eight items, a matrix of correlation among residuals across the items is computed that can be analyzed to detect the presence of any residual structure. Finally, Rasch reliability is computed as the proportion of total variance in the population that is accounted for by the measurement model. As the standard Rasch reliability statistic is affected both by model fit and by the distribution of severity of food insecurity in the sample, a modified Rasch reliability, in which calculations of error variance and variance accounted for by the model are weighted equally across raw scores, was used in this study. This equal-weighted reliability statistic avoids the confounding effect of differences in the distribution of severity of food insecurity and ensures comparability to of the reliability assessment across countries.

For additional details on the application of Rasch model to food insecurity measurement, see Nord 2015 [9].

In spite of the wide range of countries in which the FIES-SM was administered, the attendant challenges of translation and adaptation to local cultures and languages, and the relatively small size of the samples, the fit of all the items to the Rasch measurement model was remarkably good in almost all countries, confirming that the 8 items chosen to be part of the FIES survey module do indeed possess global validity. In 94 percent of countries, infit statistics for every item were in the acceptable range. Only 9 countries in 2014 had any item with an infit higher than 1.4. With one exception, those were countries with small numbers of non-extreme cases and therefore sizeable estimation errors for item-infits. Conditional correlations among residuals were always found not to be excessive for any possible pairs of items in all countries with sufficient sample size of non-extreme cases to produce reliable assessments. Mean 'flat' Rasch reliability across countries (using the modified statistic) was 0.740, with a range from 0.68 to 0.83. Reliability was greater than 0.70 for 88 percent of countries in 2014. These levels of reliability for a scale comprising just eight items reflect reasonably good model fit, and result in measurement errors in national prevalence estimates that are small compared with sampling errors.

4. Development of a global reference and scale calibration

Use of a measure of food insecurity to inform indicators used in a global monitoring framework must ensure that estimated prevalence rates are comparable over time and across countries. To do so, severity thresholds for classification should be defined on a common reference scale and kept constant during the monitoring period, while prevalence rates computed by ensuring that severity measures and thresholds are expressed in the same metric. This can be done either by mapping the national measures on the global scale, or vice versa, mapping the thresholds defined on the reference scale on the national ones. Such mapping is obtained by the linear transform that equates the mean and the standard deviation of the severity values for items that can be considered equivalent in the two scales and therefore used as anchoring points. In principle, all eight FIES items should be equivalent when comparing the application of the FIES in a country with the global reference scale, as they are all obtained from translation and adaptation of the same eight questions reported in Table 1. In practice, we had to confront the possibility that in some cases the same reported experience might be associated with different relative levels of severity in different countries due to differences in cultures or contexts, and therefore that not all eight FIES items could be used as “anchoring” points.

The definition of the global reference scale and of the mapping that allows calibration of the national measures on the global scale were initially obtained through an iterative process, using the data collected in 147 countries or territories in 2014. The iterative procedure worked as follows: First, the FIES scale was estimated separately for each country, and item severity measures normalized to have zero mean and unit standard deviation in each of the countries. Then, the median values of the severities for each item were taken as the severity levels associated with a provisional global 8-item FIES reference scale. The set of item severities for each country was then compared with the provisional global scale, to identify items that differed in severity by more than a specified tolerance. Any such item was marked as 'unique', and all measures rescaled by equating the mean and standard deviations of the remaining common items only. Once the procedure was applied to all countries, a new provisional global scale was formed by taking again the median values of the newly computed item severities. The procedure was repeated, identifying – at each iteration – possible new items that would be 'unique', and the process continued until no further change was observed in the set of common items in any of the 147 countries. The global scale resulting from this final iteration (Fig. 2) was thus adopted as the global standard scale, and the transformation that equated the mean and standard deviation of the common items at the latest iteration provided the needed calibration mapping for each country. It is useful to note that, even if not used for calibration, 'unique' items are retained in the countries’ scale as they contribute to the overall measure of severity, provided they reveal adequate infit statistics.

Once the global FIES standard is established, measures produced with other existing experience-based food security scales that include a sufficient number of equivalent questions to be used as anchoring points can be calibrated against it following the same procedure we used for FIES data. After normalization, the severity levels of the candidate anchoring item, that is the items that have cognitive content similar to that of one of the eight FIES items, are compared to the severity level of the corresponding items on the global FIES. Items whose severity differs by more than a small set tolerance, are excluded from the calibrating set, and the procedure is repeated until no further item is excluded. Calibration is considered acceptable if at least 5 items can be identified as anchoring points.

Using the scales estimated in each country with the pooled samples from 2014, 2015 and 2016 (see more below), calibration against the global reference scale was very robust: in 121 out of 153 cases calibration was achieved using at least 6 items, and in only 2 cases it was not possible to identify at least 4 common items (see Table 2).

Table 2
Number of cases by number of items used in the calibration of the national scale against the global reference standard.

Number of common items	Number of cases	Percent
3	2	1.31
4	10	6.54
5	20	13.07
6	32	20.92
7	62	40.52
8	27	17.65
	153	100.00

In preparation for the 2014 Voices of the Hungry Technical Report, measures based on data from the US collected through the HFSSM, from Brazil using the EBIA, from Mexico using the EMSA and from Guatemala using the ELCSA were successfully calibrated against the global FIES standard, thus allowing to use the thresholds defined by FAO to produce estimates of the prevalence of food insecurity in those countries that are comparable with those obtained using the GWP FIES data [10]. Similarly, in 2015, national HFSSM data from Canada and from Israel have been used to produce FIES comparable food security estimates. Other datasets obtained using the HFIAS in Bangladesh and the FIES in Swaziland and in Lesotho were successfully calibrated against the global FIES in the context of research aimed at promoting the use of FIES-based measures in the context of the Integrated Food Security Phase Classification (IPC) [11], as were ELCSA data collected in Ecuador. Finally, FIES data collected by national institutions in Burkina Faso in 2014 [12] and in St. Lucia [13] and the Seychelles [14] in 2016 have been successfully used by the responsible statistical units to produce estimates of food insecurity using the international FIES thresholds.

5. Scale stability and estimates of food insecurity prevalence rates with small samples

We were concerned that GWP effective sample sizes of non-extreme cases (i.e., after omitting cases that denied or affirmed all items, which provide no information for parameters estimation when using CML) might be too small, in many countries, to provide sufficiently precise parameter estimates. If that is the case, estimation of the scale using data from only one year could be rather unstable for those countries. With data sets from the GWP rounds of 2014, 2015 and 2016, stability over time was tested by comparing the scale estimated using 2014 data

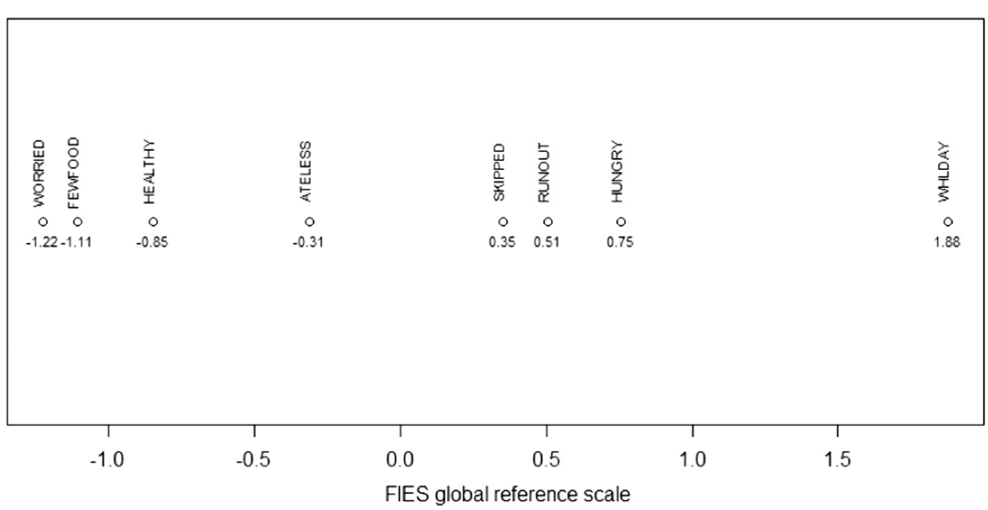

Fig. 2. The food insecurity experience scale global reference scale.

with the one obtained using 2015 data as well as the 2015 with the 2016 one, for all countries. Item severity parameters estimated for each country in each year were then compared. If, for a country, the difference between the estimated severities of the same item in the two years was larger than a set threshold, we defined that item as "unique" and concluded that, for that country, it might have functioned differently in the two years. This could be due to differences in the realized samples or to poor fit of the Rasch model in one of the two or in both years. Comparing the 2014 with the 2015 assessment, we found that 93% of countries had either all items common or only one item unique. Comparing 2015 to 2016, the corresponding percentage is 95%. Overall stability of the scale is thus rather good, considering the small size of each year effective sample for many countries and the fact that the impact of the detected instability on the severity level estimated for the raw scores would be almost negligible.

Even if small, the additional variability induced by the instability of scale estimation on the estimated food insecurity prevalence rates adds to the effect of sampling variability in confounding measures of changes in food insecurity prevalence from one year to the next. To reduce the impact of such potential confounding, annual estimates of the prevalence of food insecurity at country level, using the small GWP sample, are best presented as three-year moving averages, especially if these values are to be taken as baseline reference for monitoring progress against Target 2.1 of the SDGs. Estimates of the prevalence of food insecurity at regional and global level, on the other hand, obtained by averaging results across several countries, can be deemed sufficiently reliable and thus presented separately for each year, as the potential bias due to sampling variability and scale instability is reduced.

Thanks to the invariance property of the Rasch model, estimates that are more robust against potential differential item functioning between samples collected in different years, can be obtained by estimating each country's specific item and raw score parameters using the data from all available years for that country. In 2017, parameters estimated from these 'pooled' samples of 2014, 2015 and 2016 data are used to revise the global reference scale and to define the parameters needed to map the common thresholds on each of the national scales; the only time-varying information that is needed to estimate prevalence rates separately for each year is then that year's distribution of respondents across raw scores.

6. Setting thresholds and estimating prevalence rates for global SDG monitoring

For the specific purpose of monitoring progress against Target 2.1 of the SDGs, two thresholds have been set: one that identifies the level of severity beyond which a respondent would be classified as having moderate or severe food insecurity, and one that identifies severe levels only. The definition of a threshold of severity for the latent trait is, to a certain extent, arbitrary, as the only requirement for consistent classification is that whatever threshold is chosen, it is kept constant over the relevant monitoring scope and period. However, to associate a substantive meaning to the statistics reported, the two thresholds to be used in the context of the global SDG monitoring have been chosen to correspond to the severity of two specific items on the global FIES scale. The first threshold is set to correspond to the severity of the 'ATELESS' item, while the second to the severity of the 'WHOLEDAY' item. Individuals classified as having experienced moderate or severe food insecurity could be thus described as having eaten less than they thought they should at times during the year because they lacked sufficient resources for food, and most of them will have experienced more severe conditions. Those classified as having experienced severe food insecurity might have had high chances of going for whole days without eating, at times during the reference period, for the same reason of lack of sufficient resources to procure food.

The international thresholds defined on the global FIES scale can be mapped to the metric obtained in each country following the procedure described in the previous section. In this way, internationally comparable prevalence rates of food insecurity at the specified levels of severity (i.e., moderate or severe, $FI_{mod+sev}$, and severe, FI_{sev}) can be computed. The most accurate estimate of the prevalence of food insecurity in a given population is obtained by probabilistically assigning respondents to each of the food security classes. The probability of belonging to a given food insecurity class is computed by assuming that respondents reporting a certain raw score come from a population in which the severity of food insecurity is normally distributed, with mean equal to the value of the severity parameter and standard deviation equal to the estimated standard error for the same parameter. A probability to be at or beyond any given level of severity defined by a threshold can be thus associated with each of the 9 different raw score values and the prevalence of food insecurity in the population at that level of severity or beyond computed as the weighted sum of these probabilities, across all raw scores, using the weighted number of cases by raw scores as weights. The conditional maximum likelihood methods used to estimate the Rasch model parameters do not provide severity parameters or measurement errors for the extreme scores of 0 or 8, therefore, special considerations had to be made to classify respondents with such extreme raw scores. Respondents with raw-score 0 were classified as food secure or, at most, mildly food insecure, assuming that the probability to be beyond the threshold for moderate food insecurity is zero. For respondents with raw-score 8, an approximate severity parameter was calculated based on a pseudo raw score ranging from 7.5 to 7.7 (with the actual value being higher in samples with a higher proportion of raw score 8).

7. Results and assessment of the consistency between FIES-based measures of the prevalence of food insecurity in the world and other development indicators

Based on the above procedure, $FI_{mod+sev}$ and FI_{sev} were computed for all countries for which FIES compatible data were available. National prevalence of moderate-or-severe food insecurity in 2014–16 in the adult population ranged from 2.3% to 94%. Severe food insecurity rates ranged from below 0.5% to 83%. Regional prevalence rates were calculated as the population-weighted averages of the prevalence rates of the countries included in each region. Across the continents, food insecurity is found to be most prevalent in Africa and least prevalent in Europe. Analysis of the results at regional levels shows a statistically significant increase of the prevalence of food insecurity in 2016 compared to 2015 in Africa, in the Americas (particularly due to changes in a few countries in South America) and in Asia (due to the contribution of the South-eastern Asia sub-region) (Fig. 3).

The possibility to conduct validation of the FIES measure against other individual-level measures will be greatly enhanced once the FIES will be used in other surveys, such as Demographic and Health Surveys or Living Standard Measurement Surveys. Indirect validation of the FIES based measure against other individual level measures collected with the Gallup World Poll has been recently provided by Smith et al. [15], who find that low levels of education, limited social capital, and living in a country with low GDP per capita are the three determinants associated with the largest increase in the likelihood of experiencing food insecurity in Latin America and Caribbean countries, and by Jones [16] who finds food insecurity to be associated with poorer mental health and specific psychosocial stressors across regions.

Of course, these are only indirect proofs of the validity of the food insecurity measures obtained with the FIES. To conduct a proper validation of the estimated food insecurity prevalence rates generated with the FIES one would need alternative measures of the same phenomenon, computed using other methods/definitions, for the same set of countries and years for which FIES-based estimates have been produced. Unfortunately, no such references exist at a global scale. Many indicators that are described as generic indicators of "food insecurity", are based on concepts that are different from the "inability to access

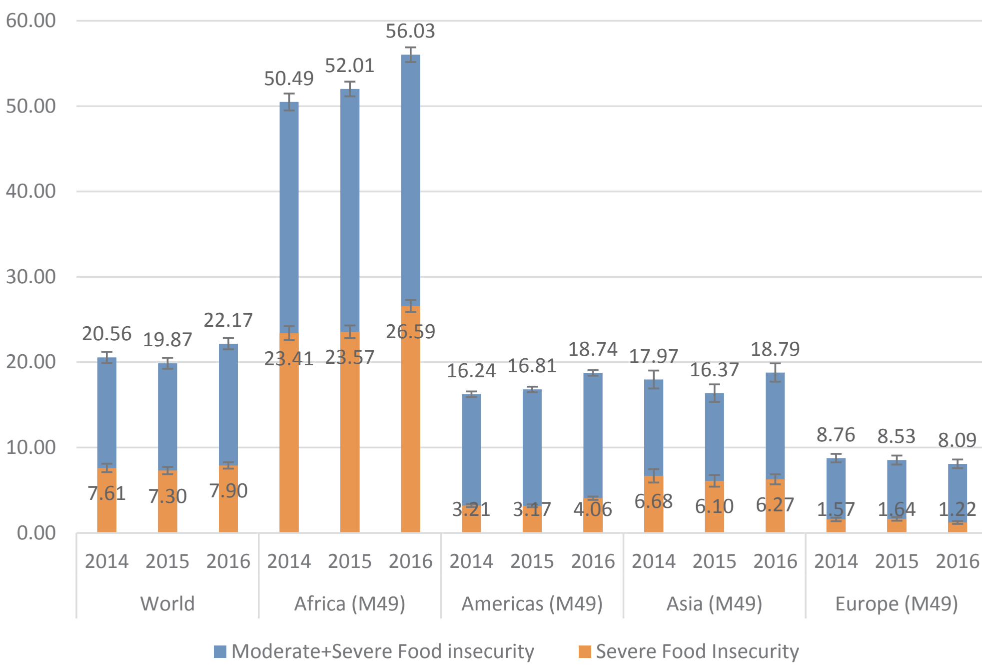

Fig. 3. Prevalence of food insecurity among adults. World and regions.

Table 3
Linear regression of national level indicators of under-5 mortality rate against $FI_{mod+sev}$, PoU and extreme poverty headcount, for 92 countries.

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	0.13	0.06	2.17	0.031
PoU (2014)	0.20	0.09	2.38	0.020
$FI_{mod+sev}$ (2014–16)	0.34	0.09	3.99	0.000
Extreme poverty (2014)	0.35	0.08	4.24	0.000

Note: Standardized regression coefficients, i.e. the variables are scaled in such a way that the reported coefficients indicate the percentage of variability in child mortality explained by each of the three indicators included in the regression.
N = 92, Adjusted R-squared: 0.7412.

food" concept that defines the latent trait measured by the FIES. Even the two indicators used in the past to monitor the "hunger" target of the Millennium Development Goals (MDGs), namely, the Prevalence of Undernourishment (PoU) and the percentage of children under five years of age who are underweight (child malnutrition) [17], are based on the concepts of adequacy of the caloric intake and of nutritional status of children, respectively, which can be considered possible *consequences* of the inability of households to access food. Moreover, as they are derived from data from household surveys conducted in different years in different countries, matching them with the FIES-based indicators is problematic. These indicators are only available as model-projected values for 2014 or later years, which means that real cross-country differences for these indicators for 2014–16 are confounded (and likely smoothed) by the assumptions made in constructing the respective projection models.

While it is not possible to conduct a validation of the indicator in the strict sense of the term, it may still be informative to verify the extent to which the ranking of countries obtained using FIES-based indicators for 2014–16 matches that obtained with the latest available series of PoU and child malnutrition. Results of such comparison show a value of the Spearman's rank correlation coefficient of $FI_{mod+sev}$ of 0.79 against the FAO Prevalence of Undernourishment and of 0.60 against child malnutrition. By broadening the comparison to include the rankings of countries provided by other general indicators of development, we find that national prevalence rates of food insecurity measured with the FIES correlate in the expected direction with the World Bank Poverty Rate (Pearson r = 0.84), the Human Development Index (r = −0.86), and the World Health Organization Under-5 Mortality Rate (r = 0.87).

One final test of the extent to which the prevalence of food insecurity defined and measured through the FIES provides useful information to monitor a global development agenda, was to verify whether or not the significant correlations noted above are simply the result of spurious correlation (i.e. due to the fact that the indicators are simply reflecting a same underlying socioeconomic phenomenon, such as monetary poverty.)

The result of a regression analysis conducted over the 92 countries for which the four indicators are available, shows that the prevalence of moderate and severe food insecurity remains a significant predictor of differences in child mortality rates across countries, explaining about 35% of the variability even when controlling for the Prevalence of Undernourishment and the prevalence of extreme poverty (Table 3).

8. Conclusions

The analysis of FIES data collected over three years in more than 150 countries worldwide confirms that self-reported evidence on the occurrence of conditions typically associated with the inability to access food due to lack of money or other resources, gathered through simple interviews, can indeed inform the construction of a valid measurement scale of the severity of the food insecurity condition. Rasch model-based analytic procedures, consistent with the item anchoring and scale equating procedures common in other applications of Item Response Theory (e.g., educational testing) has allowed for the definition of a global reference scale that can be used to ensure comparable estimates of the prevalence of food insecurity in different populations. The possibility to calibrate measures against the global standard is preserved when using other experience-based scales that share a sufficient number of items equivalent to those included in the FIES. In advocating for the adoption of the FIES or compatible scales, the inclusion of additional questions can be encouraged. Whether such additional questions effectively contribute to the measurement of the underlying latent trait may be explored in specific countries or cultures by examining their relationship with the standard FIES items using the proposed IRT statistical methods.

Pending broader adoption of the FIES or of compatible experience-based food security scales by countries worldwide, FAO offers the 2014–2016 results obtained using the data collected through the Gallup

World Poll as a possible baseline, against which to monitor progress towards Target 2.1 of the recently established 2030 Agenda for Sustainable Development.

Acknowledgements

Financial support for the Voices of the Hungry project has been provided by the United Kingdom Department for International Development (DfID) and by the Kingdom of Belgium through the FAO Multipartner Programme Support Mechanism (FMM). The comments, opinions and judgments expressed in this paper are those of the authors and do not imply any official position by FAO, the FAO Statistics Division, or the funding partners.

References

[1] Declaration of the World Summit on Food Security, World Summit on Food Security, Rome, 16–18 November 2009.

[2] A. Coleman-Jensen, M. Rabbit, C. Gregory, A. Singh, Household Food Security in the United States in 2014, Economic Research Report No. ERR-194, U.S. Department of Agriculture, Economic Research Service, Washington, DC, 2015.

[3] United Nations, Transforming Our World: the 2030 Agenda for Sustainable Development, United Nations, Resolution adopted by the General Assembly on 25 September 2015, 2015.

[4] Gallup World Poll, < http://www.gallup.com/services/170945/world-poll.aspx > .

[5] G. Engelhard Jr., Invariant Measurement: Using Rasch Models in the Social, Behavioral, and Health Sciences, first ed., Routledge, New York and London, 2013.

[6] T. Bond, C.M. Fox, Applying the Rasch Model: Fundamental Measurement in the Human Sciences, third ed., Routledge, New York and London, 2015.

[7] G.H. Fischer, I.W. Molenaar, Rasch Models: Foundations, Recent Developments, and Applications, Verlag Springer, New York, 1995.

[8] S. Viviani, C. Cafiero, M. Nord, RM.weights: Weighted Rasch Modeling and Extensions using Conditional Maximum Likelihood, < https://CRAN.R-project.org/package=RM.weights > .

[9] M. Nord, Introduction to Item Response Theory applied to Food Security Measurement: Basic Concepts, Parameters, and Statistics. Technical Paper, FAO, Rome, 2014 (available at < http://www.fao.org/economic/ess/ess-fs/voices/en >).

[10] United Nations Food and Agriculture Organization, Methods for Estimating Comparable Rates of Food Insecurity Experienced by Adults throughout the World, FAO, Rome, 2016.

[11] C. Cafiero, S. Viviani, F.N. Yassin, Using experience-based food security scales in the context of the Integrated Food Security Phase Classification (IPC). Unpublished Report, 2017.

[12] D. Serme, Use of the FIES in Burkina Faso, 2017 (unpublished report).

[13] E. St.Catherine, Use of the FIES in St. Lucia, 2017 (unpublished report).

[14] H. DeLetourdis, Use of the FIES in the Seychelles, 2017 (unpublished report).

[15] M.D. Smith, W. Kassa, P. Winters, Assessing food insecurity in Latin America and the Caribbean using FAO's food insecurity experience scale, Food Policy 71 (2017) 48–61.

[16] A.D. Jones, Food insecurity and mental health status: a global analysis of 149 countries, Am. J. Prevent. Med. 53 (2) (2017) 264–273.

[17] Millennium development goals indicators, < https://mdgs.un.org/unsd/mdg/default.aspx > .